UNIVERSITEIT GENT

FACULTY OF SCIENCES

Department of Organic and Macromolecular Chemistry
Research group NMR and Structure Analysis

Effects of fluorination on the conformation of proline: an NMR study

Thesis submitted to obtain
the degree of Master of Science in Chemistry by

Emile OTTOY

Academic year 2015 - 2016

Promoter: dr. Davy Sinnaeve
Copromoter: prof. dr. José Martins

UNIVERSITEIT GENT

FACULTY OF SCIENCES

Department of Organic and Macromolecular Chemistry
Research group NMR and Structure Analysis

Effects of fluorination on the conformation of proline: an NMR study

Thesis submitted to obtain
the degree of Master of Science in Chemistry by

Emile OTTOY

Academic year 2015 - 2016

Promoter: dr. Davy Sinnaeve
Copromoter: prof. dr. José Martins

List of abbreviations

Ac-Pro-OMe	Methyl acetyl-L-prolinate
Ac-(4*R*)-FPro-OMe	Methyl acetyl-L-(4*R*)-fluoroprolinate
Ac-(4*S*)-FPro-OMe	Methyl acetyl-L-(4*S*)-fluoroprolinate
Ac-4,4-F_2Pro-OMe	Methyl acetyl-L-4,4-difluoroprolinate
CD	Circular dichroism
COSY	Correlation spectroscopy
DFT	Density functional theory
EXSY	Exchange spectroscopy
Flp	(4*R*)-fluoroproline
flp	(4*S*)-fluoroproline
FRET	Förster resonance energy transfer
GUI	Graphical user interface
HMBC	Heteronuclear multiple bond correlation
HOESY	Heteronuclear Overhauser enhancement spectroscopy
HSQC	Heteronuclear single quantum correlation
NMR	Nuclear magnetic resonance
nOe	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser enhancement spectroscopy
PPI	Polyproline I helix
PPII	Polyproline II helix
Pro	Proline
PRR	Proline rich region
PSYCHE	Pure shift yield by chirp excitation
PSYCHEDELIC	PSYCHE to deliver individual couplings
RMSD	Root mean squared difference
TOCSY	Total correlation spectroscopy
Xaa	Any amino acid

Dankwoord

Alvorens deze thesis te beginnen zou ik graag de tijd nemen om een aantal mensen te bedanken, te beginnen met de mensen, zonder wiens inbreng, deze thesis niet mogelijk was geweest. Mijn dank gaat eerst en vooral uit naar Davy, om zo een goede begeleider te zijn en me ten allen tijde met raad en daad bij te staan. Jouw enthousiasme werkt zeer aanstekelijk en ik ben blij dat ik het fluoroproline team mogen versterken heb. Ook professor José Martins zou ik willen bedanken om me de kans te geven om mijn thesis in zijn groep te doen. Voorts zou ik ook Gert-Jan willen bedanken voor de synthese van de onderzochte moleculen, die anders niet onderzocht hadden kunnen worden. Mijn dank gaat tevens uit naar Krisztina, om me vertrouwd te maken met de spectrometers, en naar Ilya Kuprov, voor het ter beschikking stellen van zijn computationele berekeningen.

Bedankt aan alle leden van de NMR groep: Tim, voor de technische hulp en voor me op te lappen als dat nodig was; Matthias en Dieter, voor me een aantal labohandelingen bij te brengen; Niels, voor antwoord te geven op mijn (waarschijnlijk pietluttige) vragen; en Kim, Jonathan en Emile, voor de sfeer en gezelligheid. Jullie humor, vriendelijkheid en professionaliteit zorgen ervoor dat ik graag naar “mijn bureau” kwam.

Omdat de thesis ook het sluitstuk is van deze vijfjarige opleiding, zou ik graag ook van de gelegenheid gebruikt willen maken om een aantal van mijn medestudenten de bedanken. Eerst en vooral, Olieux, Matthi, Hannes en Laurent, voor de vijf onvergetelijk jaren en de vele epische kaartnamiddagen. Een betere vriendengroep kan men zich niet wensen. Ook bedankt Yann, om mijn persoonlijke taxichauffeur te spelen het laatste half jaar. En als laatste, maar zeker niet de minste, Jurgen en Anne-Mare, zonder wiens leuke gezelschap het maar een eenzaam thesisjaar zou geweest zijn.

Als laatste wil ik nog graag mijn ouders bedanken, om me de kans te geven deze opleiding aan te vatten, en om me steeds te steunen ook in de moeilijke tijden.

Table of contents

Samenvatting

NMR is een krachtige techniek voor het ophelderen van structuur en dynamica van peptiden en eiwitten. Wanneer men evenwel te maken krijgt met prolinerijke regio's (PRR's), ontstaan er moeilijkheden door de afwezigheid van amide protonen in deze regio's, die normaal gebruikt worden voor de toekenning via TOCSY. De problemen die dit met zich meebrengt, zouden verholpen kunnen worden door fluorering van proline. De aanwezigheid van fluor, dispergeert niet alleen de waarden van de chemische verschuivingen van het opeengepakte ^{1}H spectrum, maar maakt ook het gebruik van ^{19}F NMR mogelijk. Als men conformationele analyse van gefluoreerde prolines doet, dient men er rekening mee te houden, dat de conformatie van proline sterk beïnvloed zal worden door de stereo-elektronische effecten die fluor uitoefent. Fluoroprolines kunnen dus ook gebruikt worden om de structuureigenschappen van PRR's te moduleren. Twee conformationele evenwichten zijn van belang voor proline, zijnde de vijfring pucker (letterlijk opplooiing), die voor proline derivaten varieert tussen C$^{\gamma}$ *endo* en C$^{\gamma}$ *exo* (beiden enveloppe conformaties waarbij C$^{\gamma}$ uit het vlak ligt), en de amidebinding *cis/trans* isomerie. Een grondige conformationele analyse van alle mogelijke fluoroproline derivaten is dus nodig alvorens men kan overgaan tot de incorporatie van fluoroprolines in peptiden of eiwitten.

Om zulk een grondige conformationele analyse te kunnen uitvoeren, werd recent een nieuwe NMR methodologie ontwikkeld of basis van pure shift, zijnde PSYCHEDELIC. PSYCHEDELIC laat ons toe om alle individuele ^{1}H-^{1}H scalaire koppelingen naar een geselecteerd proton, gemakkelijk te meten als doubletten in een 2D J-geresolveerd spectrum. Deze koppelingen kunnen vervolgens met behulp van de zogenaamde Karplus curves gerelateerd worden aan torsiehoeken, die op hun beurt nodig zijn om de ring pucker te bepalen.

Het doel van deze thesis was, om voor een aantal gekende (in termen van conformationele voorkeuren) fluoroproline derivaten, zijnde Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe, Ac-(4*S*)-FPro-OMe en Ac-4,4-F$_2$Pro-OMe, het potentieel van PSYCHEDELIC te onderzoeken, d.m.v. zo veel mogelijk gemeten ^{1}H-^{1}H en ^{1}H-^{19}F koppelingen een uitgebreide analyse van de ring pucker uit te voeren, en om de gerapporteerde conformationele voorkeuren in de literatuur de bevestigen. Ook de *cis/trans* isomerie van de amide binding werd heronderzocht.

Het merendeel van de scalaire koppelingen kan zonder probleem met behulp van PSYCHEDELIC gemeten worden. Echter een probleem ontstaat wanneer protonen sterk gekoppeld zijn. Twee protonen zijn sterk gekoppeld als hun koppelingsconstante van dezelfde grootteorde is als hun resonantiefrequentieverschil. Deze sterke koppelingen leiden tot artefacten die de interpretatie van de spectra zodanig bemoeilijken, dat het meten van individuele koppelingen vaak niet meer mogelijk is. In sommige gevallen was het toch nog mogelijk om sommen of gemiddeldes van koppelingen van de sterk gekoppelde protonen te meten.

Met de puckering analyse op basis van gemeten scalaire koppelingen en ^{1}H-^{1}H en ^{1}H-^{19}F Karplus relaties, was het mogelijk om de in de literatuur gerapporteerde resultaten te bevestigen voor drie van de vier onderzochte moleculen. Ac-Pro-OMe vertoonde een 7:3 evenwicht tussen C^{γ} *endo* en C^{γ} *exo* respectievelijk, het evenwicht in Ac-(4*R*)-FPro-OMe is sterk verschoven richting C^{γ} *exo* en het evenwicht in Ac-(4*S*)-FPro-OMe is sterk verschoven richting C^{γ} *endo*. Alleen voor Ac-4,4-F_2Pro-OMe kon niet besloten worden dat het puckering evenwicht gelijkaardig is aan dat van Ac-Pro-OMe. Om dit te kunnen doen, en om het mogelijk bestaan van een kleine fractie ander conformeer voor de gemonofluoreerde moleculen te kunnen bewijzen, zal een verdere analyse en verfijning van de ^{1}H-^{1}H en vooral ^{1}H-^{19}F Karplus relaties in de context van fluoroprolines nodig zijn, aangezien de betrouwbaarheid van deze relatie onbekend is.

Het merendeel van de bekomen thermodynamische en kinetische resultaten van de *cis/trans* isomerie zijn in overeenstemming met deze uit de literatuur en kunnen tevens gerationaliseerd worden. De grootste portie aan toekomstig werk op dit vlak zal gewijd worden aan het uitbreiden van de temperatuursafhankelijke datapunten zodat er een preciezere uitspraak bekomen kan worden in verband met reactie-enthalpie, reactie-entropie, activeringsenergie, activeringsenthalpie en activeringsentropie. Zeker voor beide entropieën, waarop de onzekerheden zo groot zijn dat niet met zekerheid kan gezegd worden of ze positief of negatief zijn, zou dit de interpretatie kunnen vergemakkelijken.

Effects of Fluorination on the Conformation of Proline: An NMR Study

E. Ottoy[a], G. J. Hofman[b], B. Linclau[b], J. C. Martins[a], and D. Sinnaeve[a]

[a] Department of Organic and Macromolecular Chemistry, Ghent University, Ghent, Belgium
[b] Department of Chemistry, University of Southampton, Southampton, UK

To verify the conformational preferences of four (fluoro)proline analogues (Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe, Ac-(4*S*)-FPro-OMe and Ac-4,4-F_2Pro-OMe) (1), puckering analysis was performed using a methodology developed by Hendrickx *et al.* (2), and *cis/trans* isomerism was investigated using the strategies from Eberhardt *et al.* and Bain *et al.* (3, 4). For this, the recently developed NMR technique for extracting scalar couplings, PSYCHEDELIC (5), was used. The majority of vicinal ^{1}H-^{1}H and ^{1}H-^{19}F couplings could be extracted at pure shift resolution. The results of the puckering analysis could confirm the literature observations for Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe, but were inconclusive for Ac-4,4-F_2Pro-OMe. Thermodynamic and kinetic studies of the *cis/trans* isomerism, yielded similar results as those reported in literature.

Keywords: fluoroproline, conformational analysis, NMR

Introduction

NMR is a powerful tool for performing site-specific structure and dynamics studies of protein sequences (6). However, when dealing with proline rich regions (PRRs), this technique is plagued by the absence of amide protons. A possible solution to this problem is offered by proline fluorination. The presence of the fluorine atom not only disperses the chemical shift values of the otherwise crowded proline ^{1}H spectrum, but also makes the use of ^{19}F NMR possible. The ^{19}F spectra will be easily interpretable if only a limited amount of fluorine atoms is introduced. However, it needs to be taken into account that fluorine induces strong stereoelectronic effects that will affect the proline conformation (7). Therefore, it is important to evaluate the conformational preferences of the incorporated fluoroproline, before studying the conformation of the whole segment, or before incorporating fluoroproline with the purpose of modulating the protein structure. A library, containing the conformational preferences of all possible fluoroproline analogues, would thus be very practical. There are two conformational equilibria that are of interest: first, the five-membered ring pucker, that is usually assumed to be an equilibrium between C^{γ} *exo* and C^{γ} *endo*, both envelope conformations in which the C^{γ} atom is out of plane (Figure 2), and second, the peptide bond *cis/trans* equilibrium (8). In order to perform an extensive analysis of the puckering equilibrium, new advanced NMR methodology for measuring coupling constants, based on pure shift NMR, was employed (9). The purpose was to evaluate this methodology by reinvestigating the puckering of already studied fluoroproline analogues, being Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe, Ac-(4*S*)-FPro-OMe and Ac-4,4-F_2Pro-OMe (see Figure 1). Also a reinvestigation of the *cis/trans* equilibrium was performed for comparison with reported values.

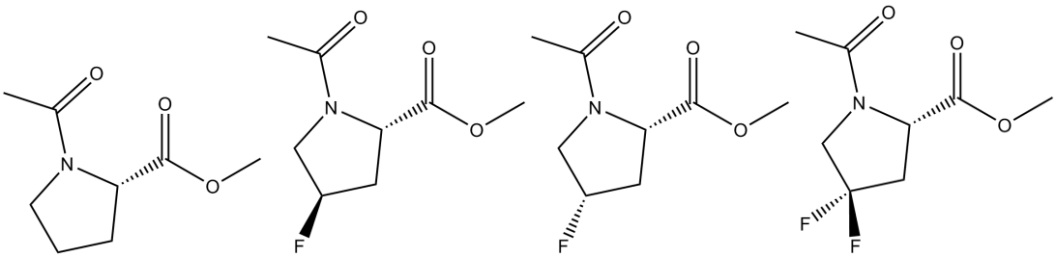


Figure 1. The investigated molecules. From left to right: Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe, Ac-(4*S*)-FPro-OMe and Ac-4,4-F_2Pro-OMe.

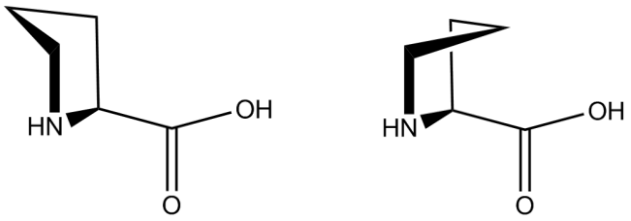


Figure 2. L-Proline. Left: C^{γ} *endo* pucker. Right: C^{γ} *exo* pucker.

Experimental

All molecules were studied in $CDCl_3$ (Euriso-Top, 99.8% D) and D_2O (Euriso-Top 99.9% D), at concentrations of approximately 20 mM. All NMR measurements were performed on either one of four different spectrometers: a Bruker Avance II spectrometer equipped with either a 5mm ^{1}H-^{13}C-^{31}P TXI-Z probe or a 5mm ^{1}H-^{13}C-^{19}F TXO-Z probe, operating at ^{1}H, ^{13}C and ^{19}F frequencies of 500.13 MHz, 125.76 MHz and 470.59 MHz respectively; a Bruker Avance III spectrometer equipped with a 5mm BBI-Z probe, operating at ^{1}H and ^{13}C frequency of 500.13 MHz and 125.76 MHz respectively; a Bruker Avance II spectrometer equipped with a 5mm ^{1}H-^{13}C-^{31}P TXI-Z probe, operating at ^{1}H and ^{13}C frequency of 700.13 MHz and 176.05 MHz respectively; or a Bruker Avance II spectrometer equipped with a 5mm ^{1}H-BB-^{19}F TBI-Z probe, operating at ^{1}H and ^{19}F frequency of 400.13 MHz and 376.50 MHz respectively. All measurements were performed at 298.0 K unless mentioned otherwise. Spectral assignment was performed using 1D ^{1}H, 1D ^{19}F, 2D ^{1}H-^{1}H COSY, 2D ^{1}H-^{13}C HSQC, 2D ^{1}H-^{13}C HMBC, 2D ^{1}H-^{1}H NOESY and 2D ^{1}H-^{19}F HOESY spectra. The main NMR technique used for measuring coupling constants was PSYCHEDELIC (Pure Shift Yielded by CHirp Excitation to DELiver Individual Couplings) (5). In this experiment, which is based on the PSYCHE pure shift experiment (10), a ^{1}H is selected, and generates a 2D J-resolved spectrum that contains only the couplings to this ^{1}H as simple doublets. In the 2DJ spectrum, the doublets are dispersed along a –45° pattern. This can be tilted by 45° to align the splitting fully along the F_1 dimension, leaving only chemical shift information along F_2, and providing a resolution similar to the PSYCHE pure shift experiment. One can excite multiple ^{1}H's at the same time, but to retain the simple doublet splittings, and thus unambiguous interpretation, these ^{1}H's may not have mutual coupling partners amongst the ^{1}H's to which the couplings are measured. If only one fluorine atom was present in the molecule, ^{1}H-^{19}F couplings could be measured using a 1D PSYCHE experiment (10). In this experiment, ^{1}H-^{1}H splittings have been eliminated, meaning only the ^{1}H-^{19}F couplings, will be present as doublets and thus easily measurable. For Ac-4,4-F_2Pro-OMe, a variant of PSYCHEDELIC was used that measures ^{1}H-^{19}F couplings in a similar way as ^{1}H-^{1}H couplings.

The extracted couplings can subsequently be used in the puckering analysis, which uses the Altona-Sundaralingam formalism (11) (equation [1]) for the description of five-membered ring

conformations based on endocyclic torsion angles (v_i). The fitting program starts with initial guesses of puckering parameters P and v_{max}, equation [1] is used to calculate the endocyclic torsion angles, these are converted to exocyclic torsion angles by adding 0° for cisoid couplings or by adding ± 120° for transoid couplings, which are then fed into the ^{1}H-^{1}H or ^{1}H-^{19}F Karplus relation (equations [2] and [3], coefficients are calculated as described in (12, 13), input parameters can be found in the supporting information) to yield calculated coupling values. When fitting two conformations, the initial guess of the minor fraction is used to calculate the weighted average scalar couplings, which are then compared to the experimental couplings. Goodness-of-fit is determined by the RMSD value (equation [4]), which is checked for convergence. Until convergence is reached, the described fitting routine is iterated (2).

$K_{trans/cis}$ was calculated by taking the ratio of the integral of a certain resonance (methyl ^{1}H or ^{19}F) of the *trans* form and the integral of the same resonance of the *cis* form. This was done using quantitative 1D spectra. For determining k_{ex} (= k_1 + k_{-1}), either a selective inversion recovery or a 2D EXSY experiment was performed. In the selective inversion recovery a $90°_{x} - \Delta - 90°_{-x}$ pulse sequence is applied, with Δ half of the inverse frequency difference between the involved resonances (*cis* and *trans* of a certain resonance). The off-resonance signal will be inverted, while the on-resonance signal will remain unaffected. By varying the recovery delays, the intensity change of the on-resonance signal (due to exchange with the inverted signal) can be followed. These data points are then fitted to equation [5], in which T_1 (determined via a regular inversion recovery experiment) serves as input and which yields k_{ex}. In a 2D EXSY, cross-peak and diagonal peak volumes of the exchanging spins (*I* and *S*), together with the mixing time, are used to calculate k_{ex} via equation [6], in which r_I is the ratio of the diagonal peak volume over the cross-peak volume (both for resonance *I*) and p_I and p_S are the equilibrium populations of spins *I* and *S* respectively. For both fitting procedures (puckering and k_{ex}), 95% uncertainties were calculated using a Monte Carlo analysis (14).

$$v_i = v_{max} \cos\left(P + \frac{4\pi}{5} i\right) \quad [1]$$

$$J_{HCCH} = C_0 + C_1 \cos(\varphi) + C_2 \cos(2\varphi) + C_3 \cos(3\varphi) + S_2 \sin(2\varphi) \quad [2]$$

$$J_{HCCF} = 40.61\cos^2(\phi) - 4.22\cos(\phi) + 5.88 + \sum_i \lambda_i[-1.27 - 6.20\cos^2(\xi_i\phi + 0.20\lambda_i)] - 3.72\left(\frac{a_{FCC} + a_{HCC}}{2} - 110\right)\cos^2(\phi) \quad [3]$$

$$RMSD = \sqrt{\frac{1}{N}\sum_i^N (J_i^{calc} - J_i^{expt})^2} \quad [4]$$

$$I(t) = I^0\left(1 - \left(ae^{-\frac{t}{T_1}} - be^{-t\left(\frac{1}{T_1} + k_{ex}\right)}\right)\right) \quad [5]$$

$$k_{ex} = -\frac{\ln\left(\frac{1 - \frac{p_I}{p_S} r_I}{1 + r_I}\right)}{t_m} \quad [6]$$

Results and Discussion

Measuring couplings

The measured vicinal ^{1}H-^{1}H and ^{1}H-^{19}F couplings can be found in Tables **7**, **8**, **9** and **10** in the supporting information. With PSYCHEDELIC, it was possible to measure the majority of couplings without any problems. The couplings developed as clearly resolved doublets and the precision is approximately 0.1 Hz. In some cases, the couplings were smaller than the linewidth of a single line (approximately 1 Hz), which means that no doublet was visible. Sometimes this could be overcome by applying Lorentz-to-Gauss resolution enhancement ('line broadening parameter' in the supporting information), but this was not always successful. In those cases, it could only be reported that the coupling has an upper limit of 1 Hz. A more serious issue occurs when protons are strongly coupled, *i.e.*, when their mutual coupling constant is of the same order of magnitude as their resonance frequency difference. Protons that are strongly coupled not only typically feature overlap between their multiplets, which makes selective excitation of the separate protons difficult, but it also leads to a series of artefacts, that complicate spectral interpretation. In some cases (*cis* Ac-Pro-OMe in $CDCl_3$, *trans* Ac-(4*R*)-FPro-OMe in $CDCl_3$, *trans* Ac-(4*S*)-FPro-OMe in D_2O), it was still possible to measure sums or averages of couplings for protons involved in strong coupling, but these results are less useful compared to knowing their individual couplings. In case of *trans* Ac-Pro-OMe in D_2O, 8 out of 10 possible vicinal ^{1}H-^{1}H couplings could not be measured. Sometimes, overlap hampered the unambiguous measurement of ^{1}H-^{19}F couplings in the 1D PSYCHE pure shift spectrum. Sometimes it was possible to measure these couplings in the 2D ^{1}H-^{13}C HSQC as the signals in question were resolved along the ^{13}C dimension.

Puckering analysis

Only the results of the samples where (nearly) all of the ^{1}H-^{1}H and ^{1}H-^{19}F couplings could be measured are discussed here, as these are the most reliable. The puckering parameters can be understood in the following manner. *P* indicates which atoms are out of plane, while v_{max} indicates how far these atoms are tilted out of plane. A C$^{\gamma}$ *exo* pucker has a *P* value of 18°, while a C$^{\gamma}$ *endo* pucker has a *P* value of 198°. The puckering parameters can be most conveniently presented in a polar plot, called a pseudorotation wheel, in which *P* is plotted on the polar axis and v_{max} on the radial axis. A selection of such pseudorotation wheels can be found in Figure 3.

Ac-Pro-OMe. According to literature (1), the ring pucker in Ac-Pro-OMe is predominantly C$^{\gamma}$ *endo* with a minor fraction C$^{\gamma}$ *exo* (30%). Fitting one conformer for Ac-Pro-OMe yields results with very high uncertainties and RMSD values that indicate that it is not possible to adequately explain the observed couplings based on one conformation only. The results from fitting two conformations for Ac-Pro-OMe, which yielded much lower RMSD values, can be found in Table **I**. These results confirm the observations from literature, as the minor fraction (1 in this case) corresponds to a C$^{\gamma}$ *exo* pucker and the major to a C$^{\gamma}$ *endo* pucker. For the *trans* conformer, the same fractions as reported in literature are obtained, but for the *cis* conformer, the fraction of C$^{\gamma}$ *exo* is only about half of the reported value. This can be explained when

bearing in mind that the puckering equilibrium and *cis/trans* equilibrium are coupled due to $n \rightarrow \pi^*$ interactions (15). In an $n \rightarrow \pi^*$ interaction, the lone pair (n) electron orbital of one carbonyl oxygen overlaps with the π^* orbital of another carbonyl group, providing an electronically stabilizing effect. These interactions can only occur in the *trans* conformer, and the C^γ *exo* pucker preorganizes the backbone in such a way that these interactions are favoured, while the C^γ *endo* pucker does not (16). Therefore, a *trans* conformer will have a preference for a C^γ *exo* pucker and vice versa.

TABLE I. Results of the puckering analysis of all samples for which (nearly) all individual couplings could be measured. Results for Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe listed here, are without use of ^1H-^{19}F couplings.

Ac-[X]-OMe	**RMSD (Hz)**	P_1 (°)	$\nu_{max,1}$ (°)	x_1 (%)	P_2 (°)	$\nu_{max,2}$ (°)
Pro *trans* $CDCl_3$	0.54	17.1 ± 7.1	49.7 ± 17.5	29.5 ± 6.2	187.9 ± 5.9	31.5 ± 7.5
Pro *cis* D_2O	0.81	13.4 ± 10.2	43.7 ± 20.5	16.9 ± 4.8	182.8 ± 3.7	36.4 ± 4.4
(4*R*)FPro *trans* $CDCl_3$	0.94	24.9 ± 2.1	48.0 ± 2.0	/	/	/
(4*R*)FPro *cis* $CDCl_3$	0.37	24.1 ± 1.7	40.7 ± 1.1	/	/	/
(4*R*)FPro *trans* D_2O	0.44	13.5 ± 1.7	43.5 ± 0.9	/	/	/
(4*R*)FPro *cis* D_2O	0.45	25.0 ± 1.7	41.1 ± 1.2	/	/	/
(4*S*)FPro *trans* $CDCl_3$	0.74	206.5 ± 2.9	37.1 ± 1.1	/	/	/
(4*S*)FPro *cis* $CDCl_3$	0.68	199.3 ± 3.0	38.0 ± 1.1	/	/	/
(4*S*)FPro *cis* D_2O	0.76	203.2 ± 2.6	41.6 ± 1.2	/	/	/

Ac-(4*R*)-FPro-OMe. The result listed in Table **I** is based on ^1H-^1H couplings alone, and assumes a single conformation. According to literature (1), the ring pucker of Ac-(4*R*)-FPro-OMe has a large preference for C^γ *exo*. All the fitted *P* values for this molecule are indeed close to 18°, while the RMSD values are quite low. This not only confirms that C^γ *exo* pucker is dominant in Ac-(4*R*)-FPro-OMe, but this also suggests that it is reasonable to assume only one ring pucker is adopted. When including ^1H-^{19}F couplings, the outcome did not change significantly. If one wants to fit two conformations, inclusion of the ^1H-^{19}F couplings is unavoidable, as the maximum number of available vicinal ^1H-^1H couplings (six) is insufficient to fit five unknown parameters. For the two *trans* conformers, the resulting uncertainties on the fraction were unreasonably high, so that it could not be proven that a second fraction is present. For the two *cis* conformers, uncertainties on the fractions were acceptable, and the major pucker coincides with the C^γ *exo* pucker. However, the puckering parameters of the minor pucker are unrealistic, leading to endocyclic torsion angles of over 100°. It is not surprising that attempting to obtain two conformers is less reliable in this case, since a very minor fraction conformer will be very sensitive to small coupling measurement errors and the limited accuracy of the Karplus relations. It is very likely that there exists an equilibrium between a C^γ *exo* and a C^γ *endo* form with a very small fraction of the latter, but this cannot be proven with the current methodology. For this, proper validation of the ^1H-^{19}F Karplus relation, used in the context of fluoroprolines, is in order.

Ac-(4*S*)-FPro-OMe. The result listed in Table **I** is based on ^1H-^1H couplings alone, and assumes a single conformation. In literature (1), C^γ *endo* is the observed dominant pucker for Ac-(4*S*)-FPro-OMe. The fitted *P* values are all close to 198°, so the observations in literature can again be confirmed. Again, the low RMSD values indicate that assuming one puckering conformer adequately fits the data. When including the ^1H-^{19}F couplings, roughly the same results were obtained. When fitting two conformations (including ^1H-^{19}F couplings, *vide supra*), in all cases, the uncertainties on the fractions were very large. An analogous conclusion as for Ac-(4*R*)-FPro-OMe can be drawn, that it is not possible to adequately fit two conformers to the data with the methodology in use.

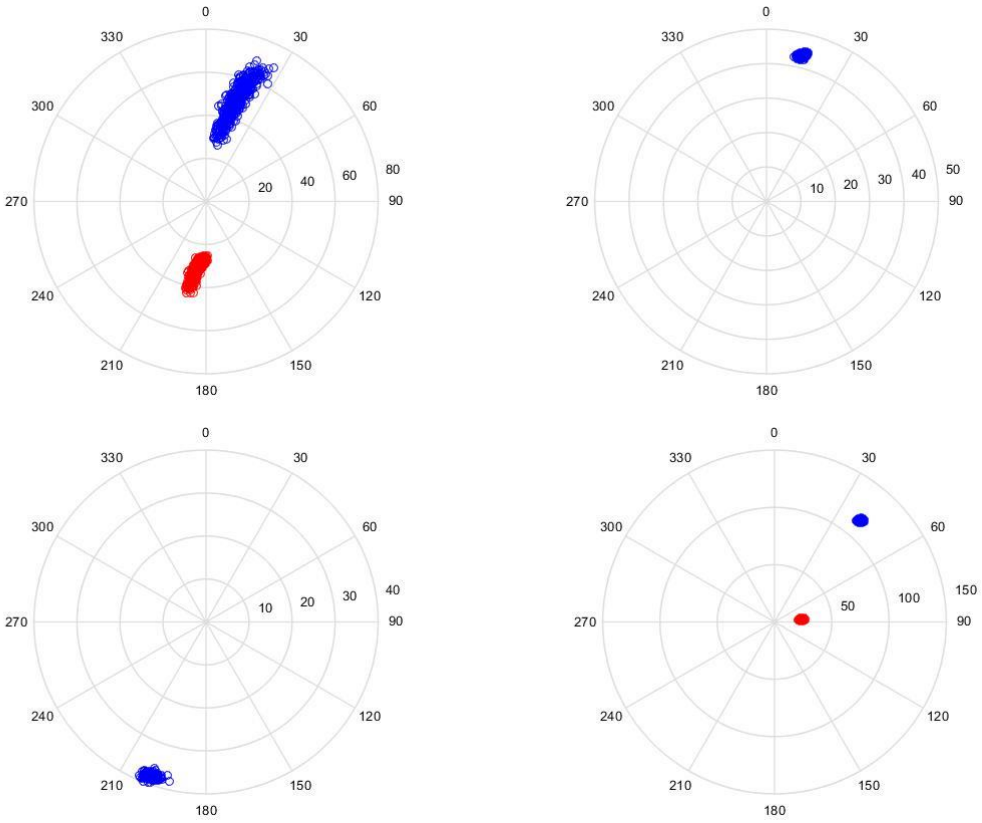


Figure 3. Selection of pseudorotation wheels. Top left: Two fitted puckers for *trans* Ac-Pro-OMe in $CDCl_3$. Top right: One fitted pucker for *trans* Ac-(4*R*)-FPro-OMe in D_2O without use of ^{1}H-^{19}F couplings. Bottom left: One fitted pucker for *cis* Ac-(4*S*)-FPro-OMe in $CDCl_3$ without use of ^{1}H-^{19}F couplings. Bottom right: Two fitted puckers for *trans* Ac-4,4-F_2Pro-OMe in $CDCl_3$ with use of ^{1}H-^{19}F couplings.

Ac-4,4-F_2Pro-OMe. According to literature (17), based on comparison of coupling constant patterns, and based on the assumption that the two counteracting stereoelectronic effects will cancel each other out, the conformational behavior of Ac-4,4-F_2Pro-OMe should be similar to that of Ac-Pro-OMe. Fitting for Ac-4,4-F_2Pro-OMe, always requires the use of ^{1}H-^{19}F couplings, as the amount of vicinal ^{1}H-^{1}H couplings (two) is insufficient, even when assuming only one conformer. Fitting one conformer yields results with high uncertainties and high RMSD values, which, like Ac-Pro-OMe, indicate that one conformer is insufficient to explain the data. This confirms the assumption that there will be an equilibrium between two or more puckers. When fitting two conformers, not only do the puckers not correspond to C^{γ} *endo* or C^{γ} *exo* (see Figure 3), but the results are also inconsistent between *cis* and *trans* forms and between the two solvents. While it is clear that the possibility of other conformations than C^{γ} *endo* and C^{γ} *exo* were never considered for this molecule, and that it was never analyzed so extensively before, care has to be taken not to jump to conclusions too rapidly. When comparing the two vicinal ^{1}H-^{1}H couplings of Ac-4,4-F_2Pro-OMe with those of the other molecules, it is indeed not unreasonable to assume that Ac-4,4-F_2Pro-OMe is similar to Ac-Pro-OMe. Furthermore, it is not clear to which extent the applied ^{1}H-^{19}F Karplus curves are accurate for fluoroprolines. The results for Ac-4,4-F_2Pro-OMe will indeed be the most sensitive to any inaccuracies in this respect. Yet again, it can be stated that a proper validation of this relation is necessary. It can thus be concluded that further research is needed in order to make a statement about the puckering preferences of Ac-4,4-F_2Pro-OMe.

Investigation of the *cis/trans* thermodynamics

The obtained $K_{trans/cis}$ values for all the samples can be found in Tables **II** and **III**. Three trends can be observed. First, it can be seen that $K_{trans/cis}$ decreases with increasing temperature. According to Le Châtelier's principle, this means that the reaction is exothermic. This confirms that the *trans* form is more stable than the *cis* form. Secondly, the observed $K_{trans/cis}$ values are always the largest for Ac-(4*R*)-FPro-OMe, and the smallest for Ac-(4*S*)-FPro-OMe, with the values for Ac-Pro-OMe and Ac-4,4-F_2Pro-OMe falling in between. There is a correlation between the puckering equilibrium and the *cis/trans* equilibrium by the $n \rightarrow \pi^*$ interactions (15) (*vide supra*). Since Ac-(4*S*)-FPro-OMe predominantly adopts the C^γ *endo* pucker, the favorability of the *cis* conformer increases (also because of electrostatic repulsion between fluorine and the carbonyl group), although the *trans* conformer is still more prevalent. For Ac-(4*R*)-FPro-OMe, which predominantly adopts the C^γ *exo* pucker, the *trans* conformer is more favored. Since the other two molecules are not biased towards one of the two puckers, their $K_{trans/cis}$ values lie intermediate. Lastly, $K_{trans/cis}$ is always larger in D_2O than in $CDCl_3$. A possible explanation was suggested by Siebler *et al.* (18), who also observed this trend. The increased charge separation induced by $n \rightarrow \pi^*$ interactions between adjacent carbonyl groups, only occuring in the *trans* conformer, is better solvated in D_2O than in $CDCl_3$.

TABLE II. $K_{trans/cis}$ values for the $CDCl_3$ samples at different temperatures.

***T*(K)**	**Ac-Pro-OMe**	**Ac-(4*R*)-FPro-OMe**	**Ac-(4*S*)-FPro-OMe**	**Ac-4,4-F_2Pro-OMe**
283	4.36	4.12	1.80	3.29
288	4.22	4.48	1.80	3.18
293	4.03	4.32	1.74	3.10
298	3.89	4.20	1.70	3.03
303	3.76	4.00	1.68	2.95
308	3.67	3.89	1.65	2.92

TABLE III. $K_{trans/cis}$ values for the D_2O samples at different temperatures.

***T*(K)**	**Ac-Pro-OMe**	**Ac-(4*R*)-FPro-OMe**	**Ac-(4*S*)-FPro-OMe**	**Ac-4,4-F_2Pro-OMe**
298	5.04	7.23	2.50	3.51
308	4.79	6.44	2.41	3.33
333	4.03	5.19	2.18	3.02

Investigation of the *cis/trans* kinetics

The obtained k_{ex} values for all samples can be found in Tables **IV** and **V**. Again, three trends can be observed. First, one can see that k_{ex} increases with increasing temperature, because the kinetic energy increases. Secondly, k_{ex} increases with increasing amount of fluorine atoms. This is because the electron-rich fluorine atom draws electron density away from the amide bond, decreasing its double bond character and therefore facilitating rotation around the bond (1, 7). Finally, a clear solvent effect is visible. Not only is k_{ex} a lot lower in D_2O than in $CDCl_3$, but also the values for Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe are clearly different in D_2O, while they are approximately the same in $CDCl_3$. There is a change in charge separation in the amide bond during rotation. This causes the carbonyl oxygen atom to have a larger partial negative charge in a planar amide (equilibrium state) than in an orthogonal amide (transition state). Therefore, the equilibrium state is more polar than the transition state. This explains why rotation over the amide bond is slower in D_2O, which shows better solvation for polar compounds than $CDCl_3$ and thus shows higher activation energy barriers to be overcome (19).

TABLE IV. k_{ex} values in s^{-1} for the $CDCl_3$ samples at different temperatures. Determined via EXSY: all temperatures for Ac-Pro-OMe and 303 K for the other samples. All others determined via selective inversion recovery.

***T*(K)**	**Ac-Pro-OMe**	**Ac-(4*R*)-FPro-OMe**	**Ac-(4*S*)-FPro-OMe**	**Ac-4,4-F_2Pro-OMe**
298	0.0973 ± 0.0073	0.2479 ± 0.0087	0.2058 ± 0.0205	0.7372 ± 0.0465
303	0.1849 ± 0.0181	0.4239 ± 0.0348	0.4327 ± 0.0105	1.4768 ± 0.0113
308	0.3295 ± 0.0389	0.7309 ± 0.0098	0.7544 ± 0.0125	2.5056 ± 0.0658

TABLE V. k_{ex} values in s^{-1} for the D_2O samples at different temperatures. Determined via EXSY: all temperatures for Ac-Pro-OMe and 308 K for the other samples. All others determined via selective inversion recovery.

***T*(K)**	**Ac-Pro-OMe**	**Ac-(4*R*)-FPro-OMe**	**Ac-(4*S*)-FPro-OMe**	**Ac-4,4-F_2Pro-OMe**
308	0.0354 ± 0.0080	0.1069 ± 0.0105	0.0415 ± 0.0063	0.1481 ± 0.0094
333	0.3693 ± 0.0086	0.8526 ± 0.0441	0.5200 ± 0.0269	1.7008 ± 0.0402

Summary

To summarize, PSYCHEDELIC has proven to be a very helpful tool for extracting ^{1}H-^{1}H and ^{1}H-^{19}F couplings. With the fitting methodology used, it was possible to confirm the observations in literature concerning the puckering preferences of Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe. Also thermodynamic and kinetic studies of the *cis/trans* equilibrium could be confirmed and could be rationalized. However, there are still some methodological challenges that need to be addressed. Handling strongly coupled protons in pure shift spectra is a first important goal, as these are prevalent within (fluoro)prolines and limit the number and accuracy of the measured couplings. A proper validation and refinement of the ^{1}H-^{1}H and certainly ^{1}H-^{19}F Karplus relations in the context of fluoroprolines is also needed in order to comment on the conformational preferences of Ac-4,4-F_2Pro-OMe, and in order to prove the existence of a second puckering conformer for Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe. In terms of thermodynamics and kinetics of the *cis/trans* equilibrium, extension of the temperature dependent data points is necessary in order to determine precise values for $\Delta H°$, $\Delta S°$, and the transition state activation-energy E_a, activation-enthalpy $\Delta H^{\ddagger}$ and activation-entropy $\Delta S^{\ddagger}$.

Acknowledgments

The 700 MHz is part of the interuniversitary NMR facility, jointly operated by UGent, UA and VUB. The 500 MHz infrastructure is funded by the Hercules Foundation. Dr. Ilya Kuprov is gratefully acknowledged for providing computationally optimized geometries for the fluoroprolines, required as input in the ^{1}H-^{19}F Karplus relations.

References

1. M. Salwiczek, E. K. Nyakatura, U. I. Gerling, S. Ye and B. Koksch, *Chem. Soc. Rev.*, **41**, 2135 (2012).
2. P. M. Hendrickx and J. C. Martins, *Chem. Cent. J.*, **2**, 20 (2008).
3. A. D. Bain, *Progress in Nuclear Magnetic Resonance Spectroscopy*, **43**, 63 (2003).
4. E. S. Eberhardt, N. Panisik, Jr. and R. T. Raines, *J. Am. Chem. Soc.*, **118**, 12261 (1996).
5. D. Sinnaeve, M. Foroozandeh, M. Nilsson and G. A. Morris, *Angew. Chem. Int. Ed. Engl.*, **55**, 1090 (2016).
6. N. G. Sharaf and A. M. Gronenborn, *Methods Enzymol.*, **565**, 67 (2015).
7. D. O'Hagan, *Chem. Soc. Rev.*, **37**, 308 (2008).
8. K. M. Thomas, D. Naduthambi, G. Tririya and N. J. Zondlo, *Org. Lett.*, **7**, 2397 (2005).
9. K. Zangger, *Prog. Nucl. Magn. Reson. Spectrosc.*, **86-87**, 1 (2015).
10. M. Foroozandeh, R. W. Adams, N. J. Meharry, D. Jeannerat, M. Nilsson and G. A. Morris, *Angew. Chem. Int. Ed. Engl.*, **53**, 6990 (2014).
11. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, **94**, 8205 (1972).

12. E. Diez, J. San-Fabian and J. Guilleme, *Molecular Physics*, **68**, 49 (1989).
13. C. Thibaudeau, J. Plavec and J. Chattopadhyaya, *J. Org. Chem.*, **63**, 4967 (1998).
14. J. S. Alper and R. I. Gelb, *J. Phys. Chem.*, **94**, 4747 (1990).
15. L. E. Bretscher, C. L. Jenkins, K. M. Taylor, M. L. DeRider and R. T. Raines, *J. Am. Chem. Soc.*, **123**, 777 (2001).
16. R. W. Newberry, B. VanVeller, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, **135**, 7843 (2013).
17. C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber and L. Moroder, *Angew. Chem. Int. Ed. Engl.*, **40**, 923 (2001).
18. C. Siebler, B. Maryasin, M. Kuemin, R. S. Erdmann, C. Rigling, C. Grünenfelder, C. Ochsenfeld and H. Wennemers, *Chem. Sci.*, **6**, 6725 (2015).
19. E. S. Eberhardt, S. N. Loh, A. P. Hinck and R. T. Raines, *J. Am. Chem. Soc.*, **114**, 5437 (1992).

1. Introduction

The title of this master thesis is ‘Effects of fluorination on the conformation of proline’. In order to gain an understanding about the research that was conducted during this thesis, an introduction to some concepts is necessary. First of all, the nature and biological importance of the amino acid proline will be discussed. Secondly, some basic concepts of fluorine chemistry will be introduced. Finally, a description of the use of fluorine in protein research will be given, as well as some more specific details in using fluoroprolines.

1.1. Proline

Proline is one of the 23 proteinogenic amino acids. Its most distinctive feature is that its side chain binds with the backbone amide nitrogen (see Figure **1**). This implies that proline incorporates its C^{α}, its backbone amide N, and its aliphatic side-chain into a five-membered ring, and that is has a secondary amine instead of a primary amine. This makes proline a rather unique amino acid. Furthermore, within a peptide, this five-membered cyclic side chain will impose steric constraints on the polypeptide backbone conformation, making proline less compatible for incorporation in regular secondary structures such as α-helices and β-sheets. Because proline lacks an amide hydrogen, it cannot provide the hydrogen bond donor that is necessary for the formation of α-helices and β-sheets. Often, proline is found in protein segments where unusual torsion angles are required such as turns, loops or polyproline helices, *vide infra* [1].

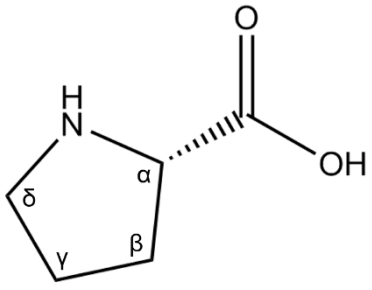


Figure 1 Structure of L-Proline

There are two equilibria that determine the conformation of a proline residue within a peptide: the *endo/exo* ring puckering equilibrium and the peptide bond *cis/trans* equilibrium [2]. Ring pucker denotes the fact that the ring in cyclic compounds is usually not flat, but adopts a certain conformation in which one or two of the ring atoms are located out of the plane formed by the other atoms. In five-membered rings, when one

atom is located out of plane, this is called an envelope conformation, and when two neighbouring atoms are located out of plane at opposite sides of the plane, this is called a twist conformation. Generally, the ring pucker in proline is found to adopt either a C^{γ} *exo* or a C^{γ} *endo* conformation, which are both envelope conformations where the C^{γ} atom is out of plane. In case of the C^{γ} *exo* conformer, the C^{γ} atom and the C' atom are oriented away from each other while in case of the C^{γ} *endo* conformer, the C^{γ} atom and the C' atom are facing towards each other (see Figure **2**) [3]. The *cis/trans* equilibrium around the peptide bond is common for all amino acids and depends on the ω dihedral angle (which is the angle defined by C^{α}-C'-N-C^{α}). The ω dihedral angle is 0° for the *cis* conformation and 180° for the *trans* conformation (see Figure **3**). In the vast majority of cases, a peptide bond will adopt the *trans* conformation, mainly because steric hindrance will occur between the two side-chains on the α-carbons when the peptide bond adopts the *cis* conformation (see Figure **4**). Again, proline confirms that it is a unique amino acid, since the prolyl peptide bond is the only peptide bond that will adopt the *cis* conformation in peptides at significant percentages. This is because there will also be steric hindrance in the *trans* conformation, between the prolyl side chain attached to the amide nitrogen and the side chain of the preceding amino acid (see Figure **5**), which causes the energy difference between both conformers to become less. However, the *trans* conformation will still be the most prevalent ($K_{trans/cis}$ of Ac-Pro-OMe = 4.6) [4].

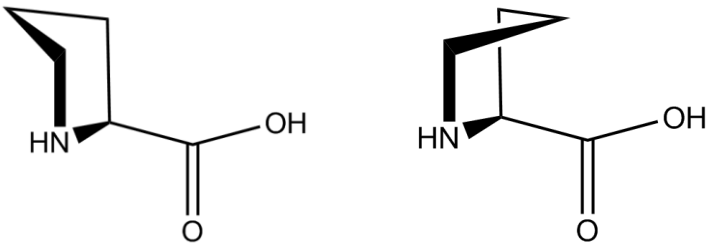


Figure 2 L-proline. Left: C^{γ} exo *conformation. Right: C^{γ}* endo *conformation. Based on a figure in [1]*

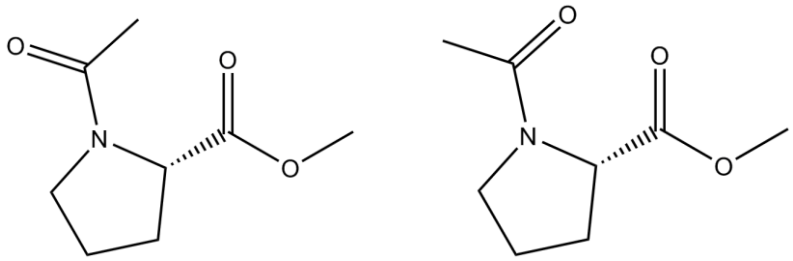


Figure 3 Methyl acetyl-L-prolinate. Left: cis *conformation. Right:* trans *conformation*

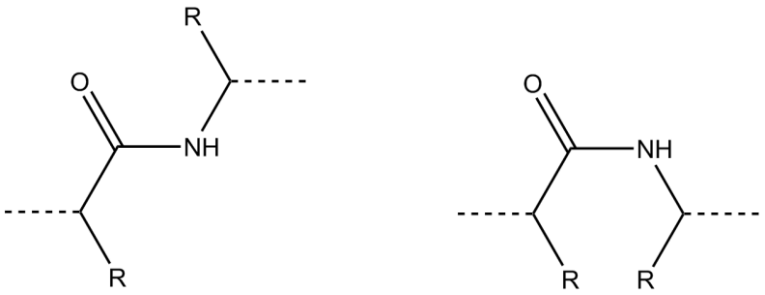


Figure 4 Regular peptide bond. Left: trans*. Right:* cis*. The dotted lines indicate the continuation of the peptide back bone.*

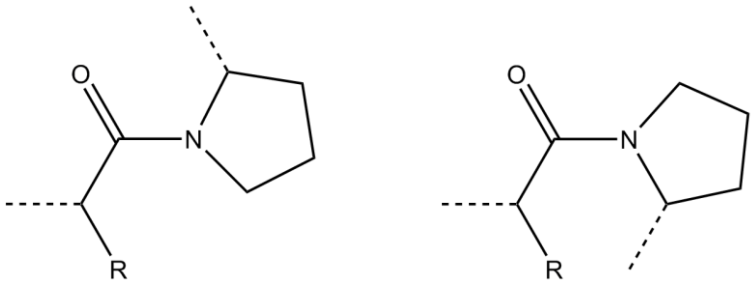


Figure 5 Prolyl peptide bond. Left: trans*. Right:* cis*. The dotted lines indicate the continuation of the peptide back bone.*

The *cis/trans* equilibrium appears to be correlated with the ring puckering. Crystal structure analysis of peptides and proteins revealed that *trans*-Xaa-Pro peptides are mostly associated with the C^{γ} *exo* pucker, while a *cis*-Xaa-Pro peptide bond is most often associated with a C^{γ} *endo* pucker [3]. It is assumed that this correlation is an effect of the so-called $n \rightarrow \pi^*$ interactions (*vide infra*) [5]. This suggests that if one would be able to shift the puckering equilibrium towards one of the two conformers, this could influence the *cis/trans* equilibrium. As will be discussed later, incorporation of fluorine in proline residues can bias the ring puckering towards one of the two forms [1].

In an $n \rightarrow \pi^*$ interaction, the lone pair (*n*) electron orbital of one carbonyl oxygen overlaps with the π^* orbital of another carbonyl group (see Figure **6**), providing an electronically stabilizing effect. Increasingly, it is realized that this non-bonding interaction could play a role in stabilizing proteins and peptides secondary structure [6, 7]. Because this interaction is a sub-van der Waals contact (the distance between the atoms involved is less than 3.22 Å), a release of energy will take place due to hyperconjugative delocalization. These interactions will polarize the electron density in carbonyl groups, which could increase the strength of subsequent $n \rightarrow \pi^*$ interactions in which these are involved, thus leading to a cooperative effect. Because of an increase of electron population of the π^* orbital, the hybridization of the carbon atom shifts from an sp^2-like character to a more sp^3-like character which effectively leads to

pyramidalization of the carbonyl group. This in turn causes the peptide bond to distort, moving away from planarity. Such distortions will be relatively small though, usually less than 1°.

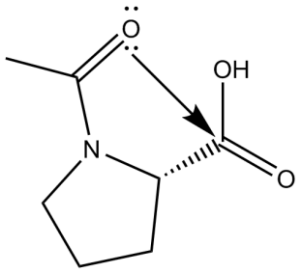


Figure 6 Methyl acetyl-L-prolinate. n → π* *interaction. Note that the amide bond adopts the* trans *conformation. Based on a figure in [7].*

$n \rightarrow \pi^*$ interactions can only take place when the peptide bond adopts the *trans* conformation (as can be seen in Figure **6**). This is the basis of the correlation of the puckering equilibrium and the *cis/trans* equilibrium. An $n \rightarrow \pi^*$ interaction will liberate most stabilization energy when the angle of approach of the oxygen lone pair is closest to the ideal Bürgi-Dunitz trajectory angle (τ_{BD}) of 109°, because these interactions can also be seen as a nucleophilic attack of the oxygen lone pair on the carbonyl carbon. The relative orientation of both is influenced by the φ torsion angle (defined by C'-N-C^{α}-C'). As the ring pucker will determine the values for φ, it will in turn determine the value of τ_{BD} (together with ψ). Therefore, it is clear the ring pucker will have an influence on the *cis/trans* isomerism. In case of a C^{γ} *exo* pucker, φ will be around -60°, leading to a τ_{BD} of approximately 100°, which is close to the ideal value of 109°, meaning that an $n \rightarrow \pi^*$ interaction will be more optimal (2.22 kJ/mol). This in turn, will contribute to a higher *trans/cis* ratio. In case of a C^{γ} *endo* pucker, φ will not change that much (-70°), due to the rigidity of the pyrrolidine ring, however, it will have a large influence on τ_{BD} leading to a value of approximately 90°. This is a suboptimal value for the Bürgi-Dunitz trajectory. This causes the stabilization energy of an $n \rightarrow \pi^*$ interaction to be rather low (0.59 kJ/mol), which will in turn decrease the *trans/cis* ratio [5, 7-9].

Next to α-helices and β-sheets, another important class of secondary structures are the so-called polyproline I and II helices [10]. As the name suggests, these are commonly found in peptides featuring repeated prolines, since these structures are not stabilized by hydrogen bonds. Nevertheless, other residues than proline can adopt this structural motif. The polyproline helix I (PPI) is a compact right-handed helix (φ = -75° ; ψ = 160° ; ω = 0° ; helical pitch = 5.6 Å/turn) in which all the peptide bonds adopt the

cis conformation. The polyproline helix II (PPII) is an extended left-handed helix (φ = -75° ; ψ = 145° ; ω = 180° ; helical pitch = 9.3 Å/turn) in which all the peptide bonds adopt the *trans* conformation (see Figure **7**) [11]. Because of the correlated equilibria mentioned above, proline derivatives that adopt the C^{γ} *endo* pucker have been found to stabilize a PPI type helix, but destabilize a PPII type helix when incorporated in those structures. Proline derivatives that adopt the C^{γ} *exo* pucker are found to stabilize a PPII type helix (*vide infra*). As will be discussed later, incorporation of fluorine can influence the ring pucker of proline residues [11].

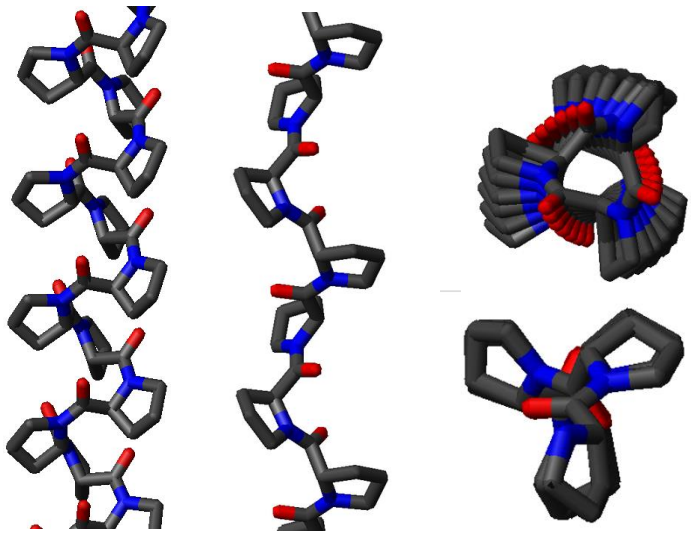


Figure 7 Polyproline helices. Left: PPI helix, side view. Middle: PPII helix, side view. Top right: PPI helix, top view. Bottom right: PPII helix, top view. Figures taken from www.wikipedia.org.

In organic solvents, PPI helices are often encountered, while in aqueous solvents, PPII helices are favoured [11]. This is because the carbonyl oxygen in a PPII helix is perpendicular to the helix axis, making it more exposed to the solvent, while in a PPI helix, the carbonyl oxygen is parallel to the helix axis and more shielded from the solvent [12]. In case of a PPII helix, these exposed carbonyl oxygens can be stabilized by hydrogen bonding in aqueous solvents but not or less in organic solvents. Since life on earth is water based, PPII helices are therefore the most commonly encountered in nature. Collagen and elastin, for example, which are two important fibrillar structural proteins of special interest in the biomedical sciences due to their mechanical properties, contain PPII helices [1]. PPII helices are also prevalent in proline rich regions (PRRs), which play a role in protein-protein interactions, molecular recognition, signal transduction and gene regulation, all typical processes to be targeted in rational drug design [13].

Gaining an understanding of aforementioned processes, certainly when PRRs are involved, is therefore of key importance. One thus needs to investigate the structure and dynamics of the PRRs, which has been done by spectroscopic techniques such as circular dichroism (CD), Förster resonance energy transfer (FRET) or Raman optical activity [14-16]. However, all these techniques have the disadvantage that they do not provide any residue-specific information. Nuclear magnetic resonance (NMR), the technique par excellence to provide residue-specific information, is unfortunately plagued by the absence of an amide proton in the proline residues, as well as the low chemical shift dispersion encountered in disordered and proline-rich proteins. This apparent incompatibility between proline residues and NMR leads us to consider the incorporation of fluorine in the proline residues (*vide infra*). For this reason, it is necessary to first discuss the structural effects that come with the introduction of fluorine.

1.2. Organofluorine chemistry

Fluorine is the element with the highest electronegativity (χ=4 on the Pauling scale), therefore the replacement of hydrogen (χ=2.1) with fluorine in organic molecules will have important consequences. First of all, the C-F bond will not only be more polarized than the C-H bond but will also be inversely polarized, *i.e.*, the partial negative charge will be on fluorine instead of carbon. This means that a C-F bond will have a more electrostatic character than a C-H bond, which is mostly covalent. Additionally, because of the high electronegativity of fluorine, the free electron pairs are quite tightly bound to the fluorine atom, rendering the C-F bond less polarizable than one would expect based on the electrostatic nature of the bond. For this reason, fluorine is generally considered not to be a good hydrogen bond acceptor, however it will interact with its environment via electrostatic-dipole interactions [17].

The C-F bond is the strongest known single bond in organic chemistry (bond dissociation energy in kJ/mol: C-F = 441.0 ; C-H = 413.4 ; C-C = 347.7 [17]) and the reason for this can again be found in the highly polarized nature of the bond. Fluorine will have the tendency to attract the largest portion of the electron density, which means that most of the electron density will be shifted to the fluorine atom. The high bond strength can thus be attributed to the significant electrostatic interaction between $C^{\delta+}$ and $F^{\delta-}$. This electrostatic attraction will also lead to bond shortening when multiple fluorine atoms are introduced on the same carbon atom, as the partial positive charge

of the carbon atom will increase. The polarization of the C-F bond not only has an influence on bond strength and bond length but also on bond angles. As we go from methane to fluoromethane, the H-C-H bond angle increases, because the C-F bond pulls electron density away from the otherwise electron rich C-H bonds, effectively relaxing electron repulsion between these bonds and causing them to spread out a little [17]. Contrary to what one might expect based on steric hindrance and lone pair repulsion, the F-C-H angle in fluoromethane is narrower than the H-C-H angle in methane and this narrowing trend even continues in difluoromethane where the F-C-F angle is the narrowest of all bond angles (H-C-H = 113.8° ; F-C-F = 108.4°). This trend can be understood as a consequence of the fluorine atoms pulling the carbon's p-orbital electrons towards them, as the fluorine atom has a lower energy 2p orbital than the carbon atom, giving more s-character to the carbon atom [17].

Another important effect introduced by fluorine is the so-called *gauche* effect. This effect finds its origin in the highly polarized C-F bond which leads to a low energy $\sigma^*_{C\text{-}F}$ antibonding orbital, susceptible to hyperconjugation. Normally, one would initially expect a 1,2-dihaloethane molecule to adopt the *anti* conformation due to the fact that steric repulsion of the two halogen atoms is minimized when their respective torsion angle is 180°. However, in the case of 1,2-difluoroethane, the lowest energy conformer will be the *gauche* arrangement rather than the *anti* arrangement. The reason for this is the stabilizing $\sigma_{C\text{-}H} \rightarrow \sigma^*_{C\text{-}F}$ hyperconjugative interactions (see Figure **8**). In the *anti* conformation, no such interactions can occur, while two such interactions can be supported in the *gauche* conformation [17]. This is a characteristic feature of fluorinated compounds, and cannot be observed in 1,2-dichloroethane and 1,2-dibromoethane. Likely due to the $\sigma^*_{C\text{-}Cl/Br}$ not being as low in energy, thus making orbital mixing less efficient, and due to a larger steric repulsion between the more voluminous Cl/Br atoms. Within 3,4-difluoroprolines (*vide infra*), the *gauche* effect will surely play an important role in determining their conformation.

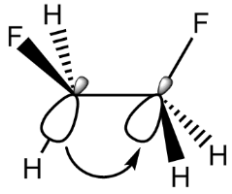


Figure 8 The stabilizing hyperconjugative interaction lies at the basis of the gauche *effect. Based on a figure in [17].*

From this point of view, one can understand that introducing fluorine in proline or polyproline helices will have an influence on the conformation of the amino acid and consequently on the global conformation of the peptide, since the *gauche* effect will influence the ring pucker.

1.3. Applications of fluorine in protein sciences

Proline is certainly not the only amino acid where fluorination is of interest. In fact, fluorination becomes more and more often exploited in protein sciences [18]. A short overview of the possible applications of protein fluorination is provided in this section, with particular attention to the use of fluoroprolines.

Fluorinated amino acids and fluorinated organic compounds in general are virtually absent in nature, only five natural fluorine containing metabolites have been discovered so far, only one of which is an amino acid, being 4-fluoro-L-threonine [19]. This implies that, in order to incorporate fluorinated amino acids in proteins, these have to be synthesized. Forming carbon-fluorine bonds is challenging, typically involving harsh reaction conditions (although specific classes of enzymes are known that can facilitate the process) [18]. Despite the synthetic challenges, incorporating fluorinated amino acids in proteins can be rewarding. Protein fluorination turns out to be a very effective way to increase their stability, creating more resistance against unfolding by heat, solvents, chemical denaturants or proteases. The reason is that protein folding is mainly driven by a hydrophobic effect, and fluorination generally increases hydrophobicity. Because of the stability of the carbon-fluorine bond (*vide supra*), the fluorine functional group is chemically inert. Perfluorinated proteins are not only very hydrophobic, but they also have interesting phase segregation behaviour, which is known as the “fluorous effect”. Perfluorinated proteins will often not be soluble in both aqueous and organic solvents, making them both hydrophobic and lipophobic. Perfluorinating proteins can thus be a way to force proteins to interact, but whether this is because of the “fluorous effect” or just because of the increased hydrophobicity, is still an issue of debate among protein scientists. [18, 20]

Fluorine is often regarded as an isostere of hydrogen since they have comparable atomic radii (1.2 Å for H and 1.47 Å for F). However, that claim is a bit too simplistic, since the C-F bond is considerably longer than the C-H bond (1.35 Å versus 1.09 Å), making fluorine more space-filling than hydrogen. Nonetheless, substituting fluorine for

hydrogen does not introduce much additional steric hindrance, which is confirmed by the observation that fluorinated compounds mimic their non-fluorinated analogues rather well. However, fluorination does introduce stereoelectronic effects (such as the *gauche* effect) that alter conformational preference, as described above [18].

Fluorine is also a very useful element in terms of NMR, which is the central analytical technique in this thesis. NMR is unrivalled when it comes to obtaining structural and dynamic information of proteins and peptides. Usually ^{1}H, ^{13}C and ^{15}N NMR techniques are used to study proteins in solution (often with ^{13}C or ^{15}N isotopic labelling), but ^{19}F NMR is gaining more and more popularity. The ^{19}F isotope is 100% naturally abundant and is an NMR active nucleus (spin ½) with a gyromagnetic ratio nearly as high as ^{1}H, leading to a very good sensitivity (83% relative to ^{1}H). Additionally, fluorine chemical shifts span a very wide range of values (from -500 up to 500 ppm) and are extremely sensitive to changes in local environment due to the large paramagnetic term in the shielding of the ^{19}F nucleus, making it an excellent probe for interaction studies. The virtual absence of fluorine in biomolecules makes ^{19}F an orthogonal probe, since spectra can be recorded without background interference [21].

As stated before, proline is not the most NMR friendly amino acid. Certainly when investigating PRRs, the lack of an amide proton is the most important drawback. Indeed, the amide proton signal is the starting point for the assignment, which typically involves TOCSY for peptides or ^{1}H-^{15}N HSQC for labelled proteins. A different issue is that there is typically a lot of spectral overlap of the proline side-chain ^{1}H resonances, even within one residue. As can be understood from the previous paragraph, introducing a limited number of fluorine atoms in one or multiple proline residues may offer an opportunity in this respect. On the one hand, the ^{19}F spectrum will be relatively easy to interpret, since only a few signals will be visible. On the other hand, the ^{1}H and ^{13}C chemical shifts of the fluorinated proline will be significantly altered by the fluorination, making it distinct from the other proline residues. A minor drawback is that the ^{1}H multiplets will generally display higher degrees of splitting due to the (often large) scalar couplings with ^{19}F. Fortunately, this can be overcome by the introduction of ^{19}F decoupling in the various NMR experiments. Thus, fluorination is a promising strategy to obtain residue specific information about the proline residues. However, one needs to take into account that fluorination will induce some conformational and

dynamical changes because of the stereoelectronic effects it imposes. These will be discussed in the following paragraphs.

As mentioned before, the ring pucker of proline is in equilibrium between the C^{γ} *exo* and C^{γ} *endo* forms. It has previously been reported that the ring puckering equilibrium will be biased towards one of the two forms depending on the position of the fluorine and the stereochemistry. The ring puckering equilibrium in (4*R*)-fluoro-L-proline (Flp) will be biased towards the C^{γ} *exo* form because of the *gauche* effect, which induces a preference for the fluorine substituent to adopt a *gauche* orientation relative to the amine (see Figure **9**). The opposite is true in the case of (4*S*)-fluoro-L-proline (flp), which will bias the ring puckering equilibrium towards the C^{γ} *endo* form for the same reason as in the case of (4*R*)-fluoro-L-proline (see Figure **9**). Since the proline ring pucker is correlated with the stability of the *cis* and *trans* peptide bond, and thus also the stability of polyproline helices (*vide supra*), introducing a (4*R*)-fluoro-L-proline in a PP-II type helix can be expected to stabilize that helix, while a (4*S*)-fluoro-L-proline destabilizes it [1]. Horng *et al.* [11] studied the behaviour of uniform decafluoroprolines – $(Pro)_{10}$, $(Flp)_{10}$ and $(flp)_{10}$ – in aqueous solutions by means of temperature dependent CD spectroscopy. The first thing that could be noticed, was the difference between the spectra of $(Pro)_{10}$ and $(Flp)_{10}$ on the one hand and the spectra of $(flp)_{10}$ on the other hand. The former two showed spectra that are a fingerprint of a PPII helix, while for the latter, the spectrum was that of a mixture of PPI and PPII helices. This already confirms that flp destabilizes the PPII conformation. To assess if Flp stabilizes the PPII conformation, the estimated T_m values for $(Pro)_{10}$ and $(Flp)_{10}$ were compared, which were 27°C and 53°C respectively. An increased melting temperature indeed indicates an increased stability. (Flp and flp nomenclature was taken over from [11])

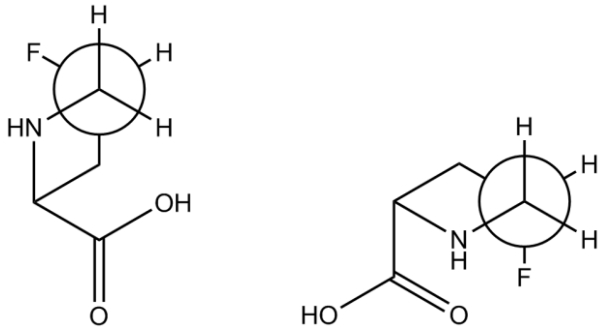


*Figure 9 Newman projections of fluoroprolines. Left: (4*R*)-fluoro-L-proline. Right: (4*S*)-fluoro-L-proline. When placing the F substituent* gauche *relative to the amine substituent, the C^{γ}* exo *and C^{γ}* endo *conformation can be found respectively.*

Furthermore, fluorination will have an influence on the extent with which $n \rightarrow \pi^*$ interactions occur in folded proteins. Since $n \rightarrow \pi^*$ interactions are only possible when the amide bond is in the *trans* conformation, introducing a fluoroproline that favours the C^{γ} *exo* pucker will preorganize the backbone in such a way that favours $n \rightarrow \pi^*$ interactions (*vide supra*) [1, 6, 7].

Lastly, fluorination does not only influence $K_{trans/cis}$, partly due to its influence on the ring pucker, and partly due to the electrostatic repulsion of the fluorine atom and the carbonyl oxygen, but it will also influence the kinetics of the *cis/trans* isomerism. The electron withdrawing effects of the fluorine atom will cause the C-N bond in the amide group to lose part of its double bond character, making rotation around this bond easier and thus accelerating the *cis/trans* isomerism [1].

So far, synthesis of the following fluoroprolines has been reported: (4*R*)-fluoro-L-proline, (4*S*)-fluoro-L-proline, 4,4-difluoro-L-proline [4, 5, 8, 22], (3*R*)-fluoro-L-proline, (3*S*)-fluoro-L-proline [23, 24], 3,3-difluoro-L-proline [25], (3*R*,4*R*)-difluoro-L-proline [26] and (3*R*,4*S*)-difluoro-L-proline [27] (see Figures **10**, **11** and **12**). Except for 3,3-F_2Pro and both 3,4-F_2Pro's, conformational analyses have been conducted by means of NMR and crystallography, the results of which can be seen in Table **1**.

In this research project, the studied (fluoro)prolines were derivatized as such: the carboxylic acid function was converted into a methyl ester and the amine was acetylized (Ac-(F)Pro-OMe derivatives). The following molecules will be investigated here: non-fluorinated (Ac-Pro-OMe), monofluorinated at the 4 position (Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe), and difluorinated at the 4 position (Ac-4,4-F_2Pro-OMe). These molecules were all already excessively studied in the past in terms of puckering and *cis/trans*. The goal of this thesis is to reinvestigate the ring pucker by measuring and analysing the ^{1}H-^{1}H and ^{1}H-^{19}F scalar couplings using advanced NMR methodology, that enables the extraction of individual coupling constants from crowded spectra. Also a systematic study of the *cis/trans* conformational equilibrium and its kinetics will be performed. The methodology that is explored in this thesis – as discussed in the following chapters – will be needed for future structural and dynamic studies of the much less well-studied 3-fluoroprolines, 3,3-difluoroproline and the 3,4-difluoroprolines.

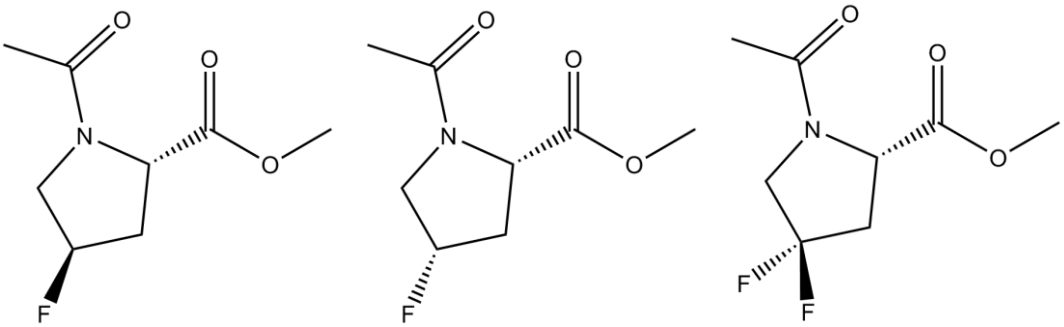


Figure 10 The 4-fluoro-L-prolines. Left: Ac-(4R)-FPro-OMe. Middle: Ac-(4S)-FPro-OMe. Right: Ac-4,4-F_2Pro-OMe.

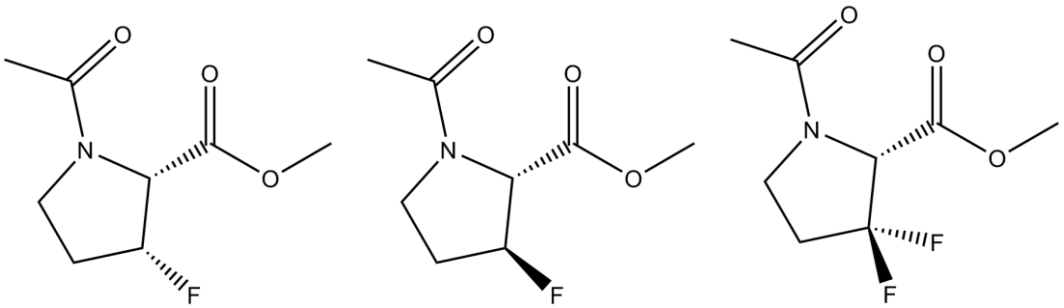


Figure 11 The 3-fluoro-L-prolines. Left: Ac-(3R)-FPro-OMe. Middle: Ac-(3S)-FPro-OMe. Right: Ac-3,3-F_2Pro-OMe.

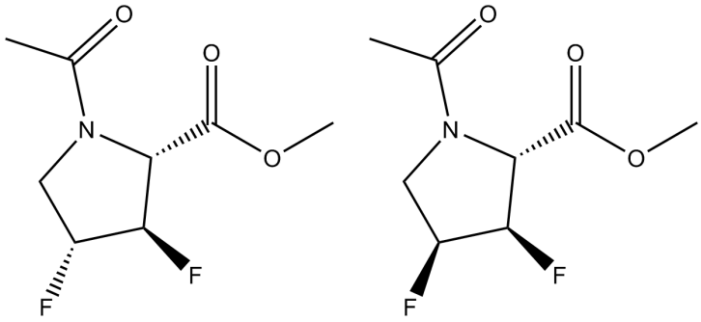


Figure 12 The 3,4-difluoro-L-prolines. Left: Ac-(3R,4R)-F2Pro-OMe. Right: Ac-(3R,4S)-F_2Pro-OMe.

Table 1 Literature values for conformational preferences of the Pro pyrrolidine ring in Ac-Pro-OMe derivatives [1].

Amino Acid	Predominant Pucker	$K_{trans/cis}$ (298K)
Pro	C$^{\gamma}$ *endo*	4.6
(4*R*)-FPro	C$^{\gamma}$ *exo*	6.7
(4*S*)-FPro	C$^{\gamma}$ *endo*	2.5
4,4-F_2Pro	C$^{\gamma}$ *endo*	3.4
(3*R*)-FPro	C$^{\gamma}$ *exo*	8.9
(3*S*)-FPro	C$^{\gamma}$ *endo*	4.3

2. Experimental Methods

In this chapter, methodologies for the assignment of the spectra, measuring coupling constants and performing kinetic studies will be discussed.

2.1. Introductory information

2.1.1. The scalar coupling

The basics of NMR can be found in the supporting information (section **6.1.1**) and will not be discussed here. However, since the concept of scalar couplings between spins will be central in this thesis, a brief description will be provided here. A scalar coupling is a through-bond interaction between nuclei, in which the nuclear spin of one nucleus affects the energy levels of the other. The mechanism of the coupling involves the electrons within the chemical bonds between the nuclei. The nuclear spin of one nucleus couples to the electron spin of the bonding electrons, which, in turn, are coupled to the nuclear spin of the second nucleus. This will lead to a decrease or increase of the nuclear spin's energy level depending on the spin-state (up or down) of the neighbouring nuclear spin. This causes the peaks of the involved spins – the effect is mutual – to split in two, with a splitting equal to the scalar coupling constant *J* (in Hz). The value of *J* is field independent. It is determined by the type of interacting nuclei, the number and type of bonds that separate them, as well as the molecular geometry (bond distances, bond angles, torsion angles). The latter makes the scalar coupling a very powerful tool for conformational analysis (see Chapter **3**). A scalar coupling will be denoted as $^{n}J_{AX}$, in which n is the number of separating bonds, and A and X are the types of interacting nuclei (e.g. $^{3}J_{HH}$).

2.1.2. Studied samples

The following molecules were studied: Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe, Ac-(4*S*)-FPro-OMe and Ac-4,4-F_2Pro-OMe (see Figure **13**), dissolved in either $CDCl_3$ or D_2O.

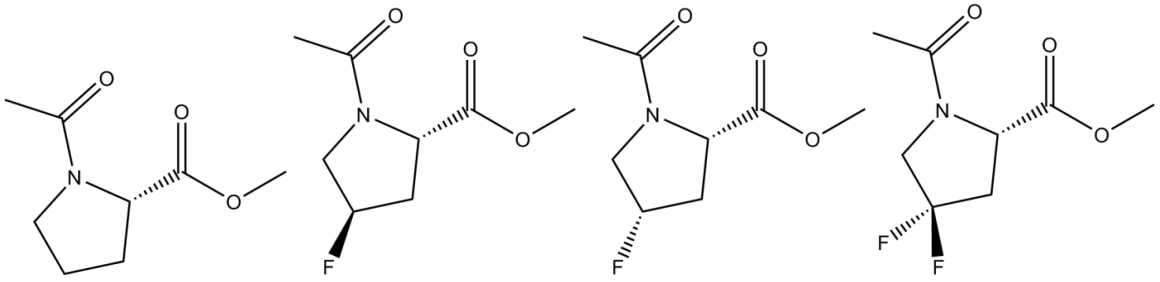


Figure 13 From left to right: Ac-Pro-OMe, Ac-(4R)-FPro-OMe, Ac-(4S)-FPro-OMe and Ac-4,4-F_2Pro-OMe.

2.2. Assignment of the spectra

The chemical shift assignment of each compound was carried out for three nuclei, being ^{1}H, ^{13}C and ^{19}F. For ^{1}H and ^{19}F, 1D spectra were recorded while for ^{13}C, the assignment was based entirely on 2D ^{1}H-^{13}C correlation spectra. The following 2D NMR techniques were used in the assignment of the spectra: 2D ^{1}H-^{1}H COSY, 2D ^{1}H-^{13}C HSQC, 2D ^{1}H-^{13}C HMBC, 2D ^{1}H-^{1}H NOESY and 2D ^{1}H-^{19}F HOESY. A short description of how these different kinds of spectra need to be interpreted can be found in the supporting information (section **6.1**). Before discussing the chemical shift assignment strategy, some general features of the spectra will be discussed. This will be necessary to understand some parts of the assignment process (*vide infra*).

2.2.1. General features of the spectra

First of all, as described in the previous chapter, two equilibria will play a role in the investigated samples. Because the nitrogen is acetylated, the resulting tertiary amide group can undergo *cis/trans* isomerism, similar as in peptides. The interconversion between the *cis* conformer and the *trans* conformer is slow on the NMR frequency timescale. This means that for each proton in the molecule, two resonances can be seen, one from the molecule that adopts the *trans* conformation, and another where the molecule adopts the *cis* conformation (see Figure **14**). The most intense resonance will be referred to as the major resonance, and the other as the minor. Based on literature we can expect the major and the minor to correspond with the *trans* and the *cis* conformers respectively. The ratio of the integrals of both peaks will reflect $K_{trans/cis}$ (see chapter **4**).

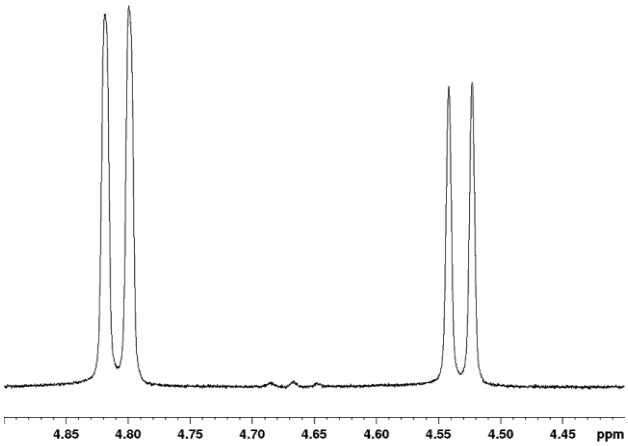


Figure 14 The α ^{1}H resonances of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

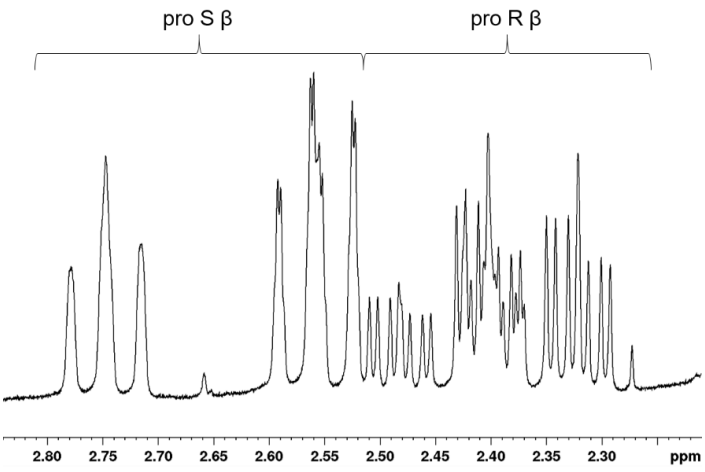


Figure 15 The β ^{1}H resonances of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

Secondly, the hydrogen atoms in the ring are diastereotopic, implying that two ^{1}H's on the same carbon atom, generally do not have the same chemical shift value (see Figure **15**). For this reason, some effort has to be invested into distinguishing the pro-*R* and pro-*S* resonances of the β and δ ^{1}H's. The same is true for the γ ^{1}H's in the case of Ac-Pro-OMe, and the γ ^{19}F's in the case of Ac-4,4-F_2Pro-OMe.

Lastly, except for Ac-Pro-OMe, where no fluorine atom is present, the resonances of all ^{1}H's in the vicinity of the fluorine atom will show peak splitting due to a scalar coupling with the ^{19}F nucleus. This scalar coupling is most striking for the γ ^{1}H of the monofluorinated analogues, which show a $^{2}J_{HF}$ splitting of ca. 50 Hz, as is evident from Figure **16**.

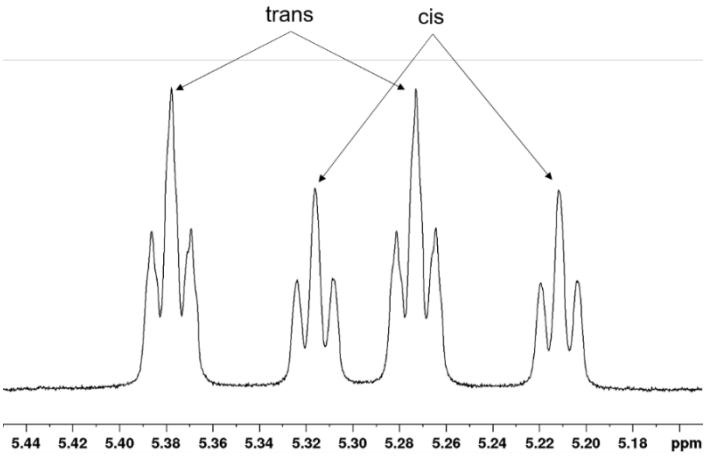


Figure 16 The γ ^{1}H resonances of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

These combined effects can make the spectrum rather crowded in some places, leading to a lot of spectral overlap, certainly in the β- and the δ-region (see Figure **17**).

2.2.2. General assignment strategy of the ^{1}H spectra

The methyl signals can easily be discerned in the ^{1}H spectra as four large singlet peaks (see Figure **18**). These correspond to the two different methyl groups, both showing a major and a minor form due to *cis/trans* isomerism (*vide supra*). The signals are intense since they should account for three protons, and because they are singlets. These methyl signals are always found around 2 and 3.75 ppm. The former correspond to the acetyl methyl, while the latter correspond to the methoxy methyl. This assignment is in agreement with the chemical shift values, since ^{1}H signals of a methyl adjacent to a

carbonyl typically exhibit chemical shifts between 1.90 and 2.25 ppm, while the chemical shifts of a methyl neighbouring an oxygen atom range from 3.5 to 4 ppm. The NOESY spectrum also provides confirmation of this assignment (*vide infra*).

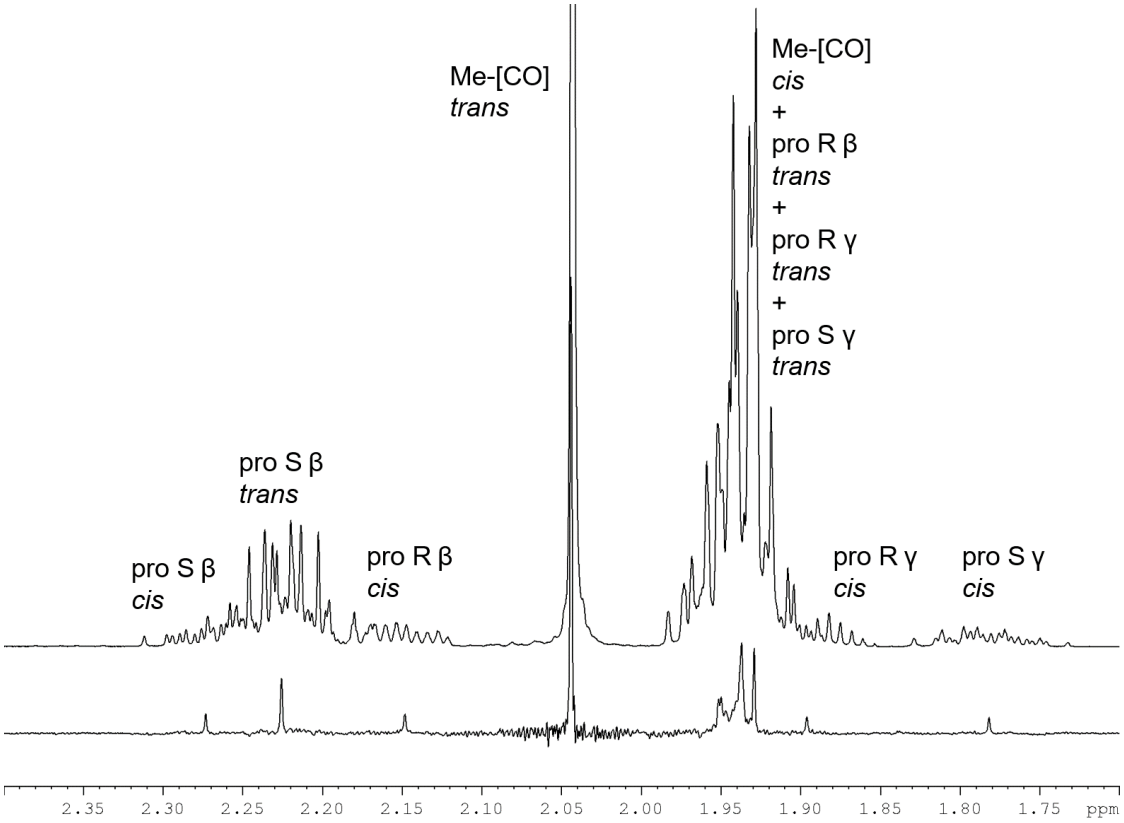


Figure 17 The β and γ region of the ^{1}H spectrum of Ac-Pro-OMe in D_2O. 298 K. 500 MHz. Top: 1D ^{1}H. Bottom: 1D PSYCHE. Note that the trans γ protons cannot be distinguished. There are also artefacts due to strong coupling between the trans γ and trans pro-R β protons at around 1.94 ppm.

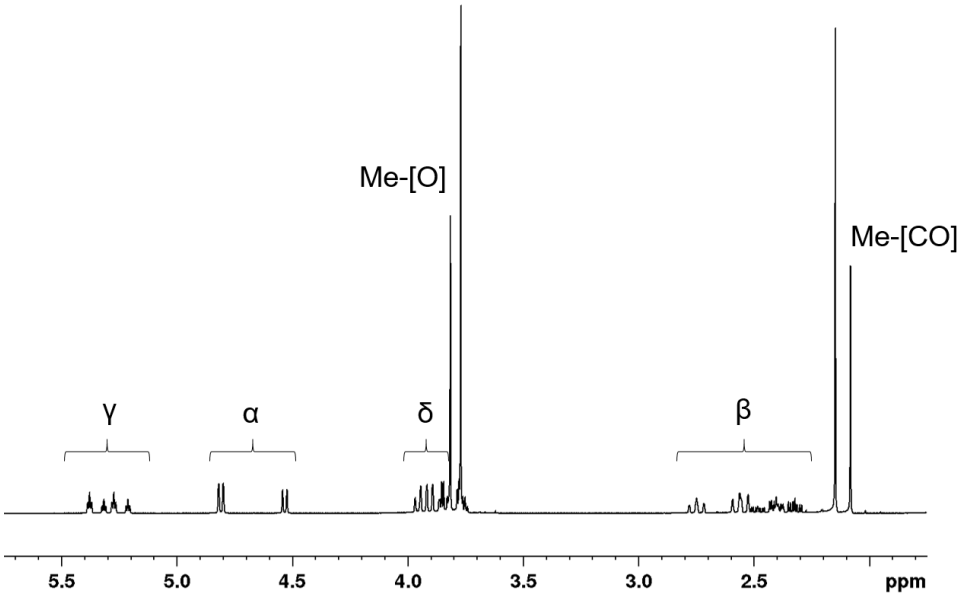


Figure 18 1D ^{1}H spectrum of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

The resonances of the α proton are also easily identified. For proline, these resonances are expected around 4.4 ppm [28], though for the fluorinated analogues, the resonances are a bit shifted to higher chemical shift values due to the presence of fluorine (see Figure **18**). They have the appearance of doublets of doublets, or triplets or doublets, depending on the value of the coupling constants (see next section). In the ^{1}H-^{13}C HSQC, these signals give positive (CH multiplicity) correlations, which is an additional confirmation that these are the α ^{1}H resonances.

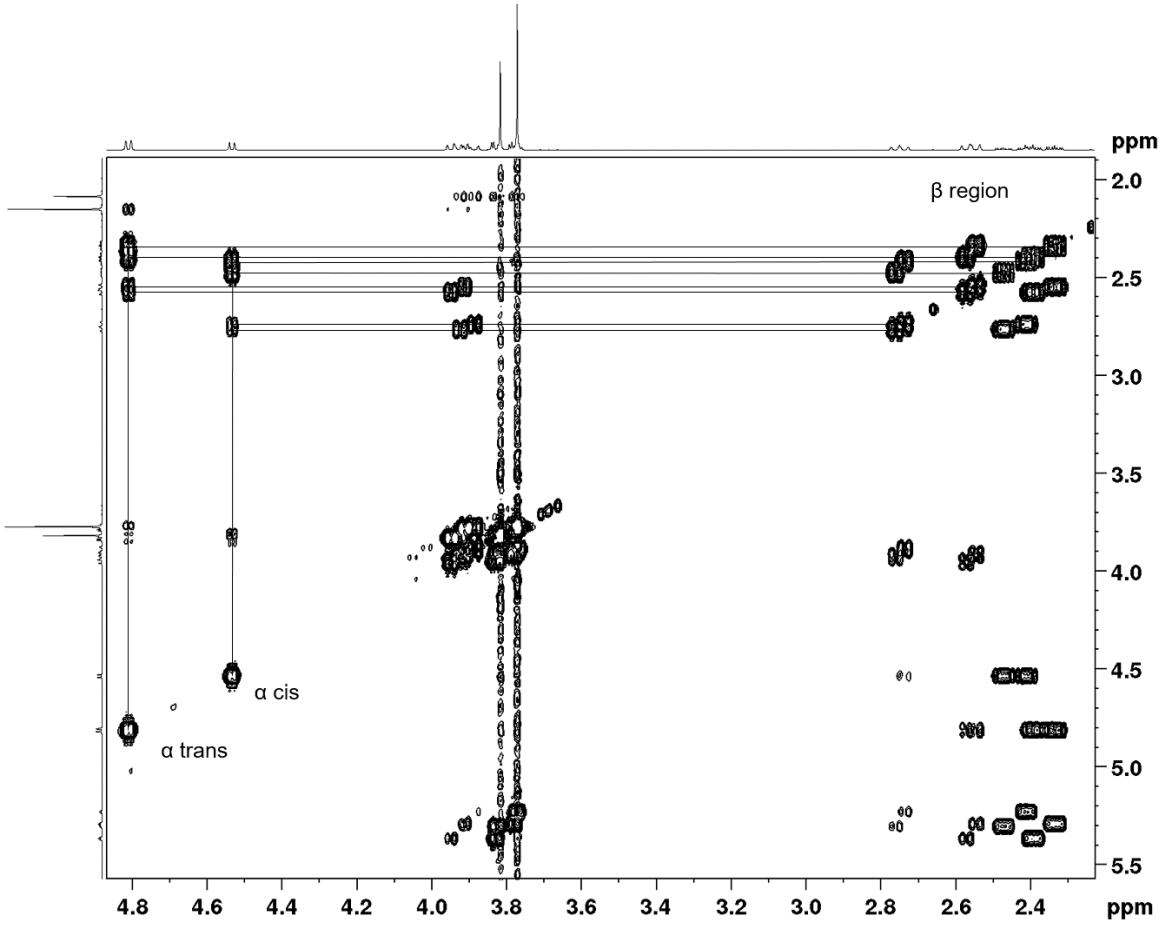


Figure 19 2D ^{1}H-^{1}H COSY spectrum of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 700 MHz. Start of sequential walk, correlations between α and β.

The β, γ and δ ^{1}H resonances can be assigned by making use of the COSY spectrum and a technique called sequential walking (see Figure **19**). The α resonance only shows correlations to the β resonances (both pro-*R* and pro-*S*) in the COSY. The major α resonance will only show correlations to the major β resonances, and the minor α resonance only to the minor β resonances, as the major and the minor are two different spin systems. This reasoning will be valid for all correlations and will no longer be explicitly mentioned from this point on. Next to a correlation with the α resonance, the

β resonances should also show a correlation to the γ resonance(s), and the γ resonance(s) should show a correlation to the δ resonances. This assignment is further validated by the ^{1}H-^{13}C HSQC. The resonances assigned to the β and δ ^{1}H's (and the γ ^{1}H's for the Ac-Pro-OMe samples), should give negative (CH_2 multiplicity) correlations in the HSQC, and both protons within a CH_2 fragment should correlate to the same ^{13}C chemical shift. In the case of the monofluorinated samples, the γ correlations will be positive, the chemical shift should be around 5 ppm, and a large $^{2}J_{HF}$ coupling should be visible (see Figure **20**).

In a last step, the stereospecific assignment (*i.e.*, pro-*R* and pro-*S*) needs to be performed for the β and δ resonances (and γ in case of the Ac-Pro-OMe samples). This can be done with help of the NOESY spectrum. For the monofluorinated samples, the strategy is rather straightforward. The β proton that shows the most intense nOe cross-peak with the α signal is assumed to be on the same side of the five-membered ring as the α proton. This is also verified by looking at the nOe cross-peaks with the γ resonance, where a similar reasoning holds (see Figure **20**). For the δ protons, one can check the intensities of the nOe cross-peaks to the γ signal and to the already assigned β signals. In all cases, it is again assumed that the most intense cross-peaks indicate that the corresponding pair of protons are on the same side of the ring.

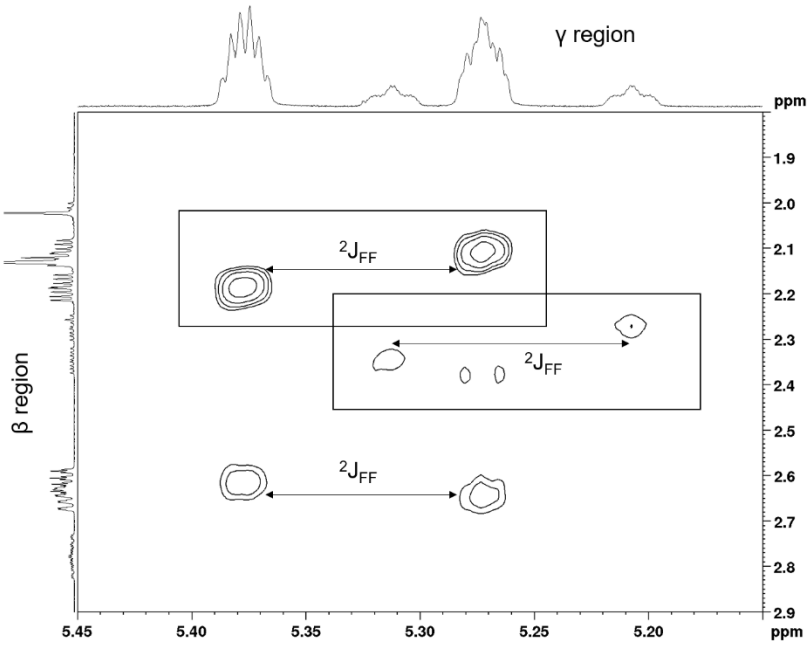


Figure 20 2D ^{1}H-^{1}H NOESY spectrum of Ac-(4R)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz. Most intense nOe cross peak indicates that this β hydrogen is on the same side of the ring as the γ hydrogen.

The NOESY also offers a possibility to confirm whether the major resonances are indeed originating from the *trans* conformer. The methyl resonance of the acetyl group of the major set of resonances shows an nOe cross-peak with the δ resonances, in agreement with what would be expected when the molecule adopts the *trans* conformation. For the minor set of resonances, no such cross-peak is observed, in accordance with the *cis* conformation.

2.2.2.1. Some special cases

2.2.2.1.1. The Ac-4,4-F_2Pro-OMe samples

The biggest nuisance in the spectra of Ac-4,4-F_2Pro-OMe is the fact that the large couplings with both ^{19}F's make the β (2.5–3 ppm) and δ (around 4 ppm) regions rather unintelligible (see Figure **21**). For this reason, ^{19}F decoupled spectra were recorded which greatly aided the assignment for these samples. Furthermore, the diastereotopic assignment of the δ hydrogens turned out to be unattainable based on the ^{1}H-^{1}H NOESY alone. Therefore, a 2D ^{1}H-^{19}F HOESY was recorded, in which the same strategy for distinguishing pro-*R* and pro-*S* as in the NOESY (*vide supra*) was used.

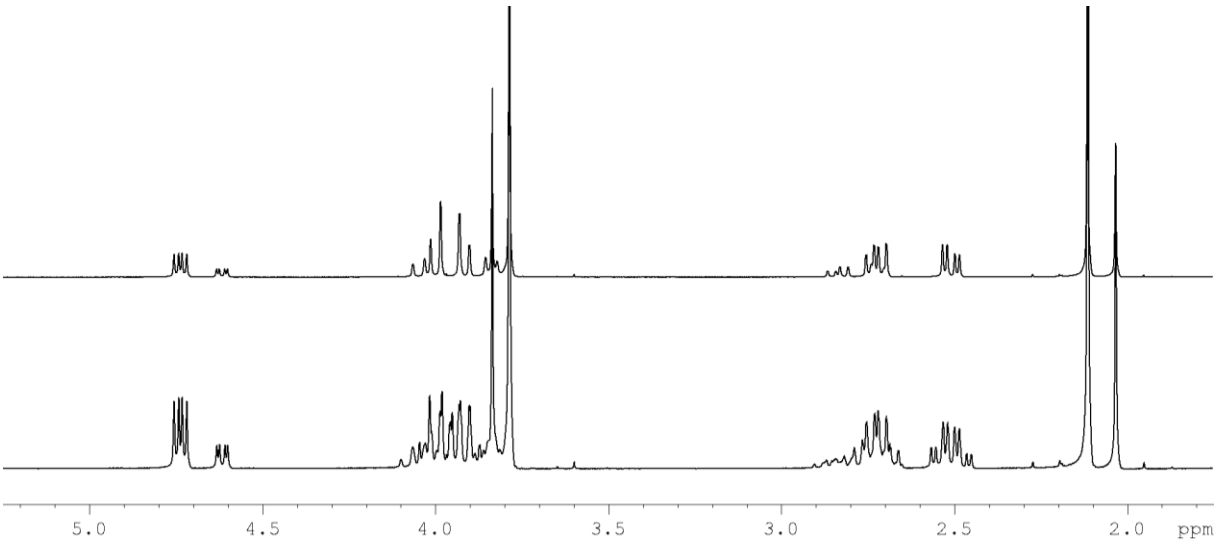


Figure 21 Ac-4,4-F_2ProOMe in $CDCl_3$, 298K, 400MHz. Bottom: 1D ^{1}H. Top: 1D ^{1}H with ^{19}F decoupling.

2.2.2.1.2. Overlap and strong coupling

In this part, the most difficult cases of spectral overlap will be discussed. This will be important for the next part, as this overlap may cause trouble in measuring couplings.

For Ac-Pro-OMe in $CDCl_3$, all four β and γ protons possess similar chemical shifts, leading to a very crowded region. At about 2.20 ppm, the pro-*R* β signal of the *cis*

conformer and the pro-*S* β signal of the *trans* conformer are found, with their multiplets overlapping. This complicates multiplet analysis in a later stage. At about 1.95 ppm, the signals of both the pro-*R* and pro-*S* γ ^{1}H of the *cis* conformer possess nearly the same chemical shift. Since these protons are coupled, this is a case of so-called strong coupling, *i.e.*, whereby the difference in resonance frequency is of the same order or smaller that the scalar coupling. This introduces second-order effects on the multiplet structures, severely complicating multiplet analysis, or, as is the case for these γ ^{1}H's, making the protons spectroscopically indistinguishable. At about 2.02 ppm, the multiplets assigned to the pro-*R* β and pro-*R* γ protons of the *trans* conformer are found – and show some degree of strong coupling – as well as the acetyl methyl signal of the *cis* conformer.

Also in D_2O, spectral overlap and strong coupling complicate matters (see Figure **17**). At 1.94 ppm, signals from four protons overlap, being pro-*R* β, pro-*R* γ and pro-S *γ* of the *trans* conformer and one of the methyl's of the *cis* conformer. The two *trans* γ protons and the pro-*R* β proton are thus strongly coupled. Also both δ protons of the *trans* conformer have close chemical shifts.

The ^{1}H spectra of Ac-(4*R*)-FPro-OMe do not pose many challenges. The only case of complications by strong coupling can be found in $CDCl_3$, where the multiplets of both the δ hydrogens of the *trans* conformer are overlapping.

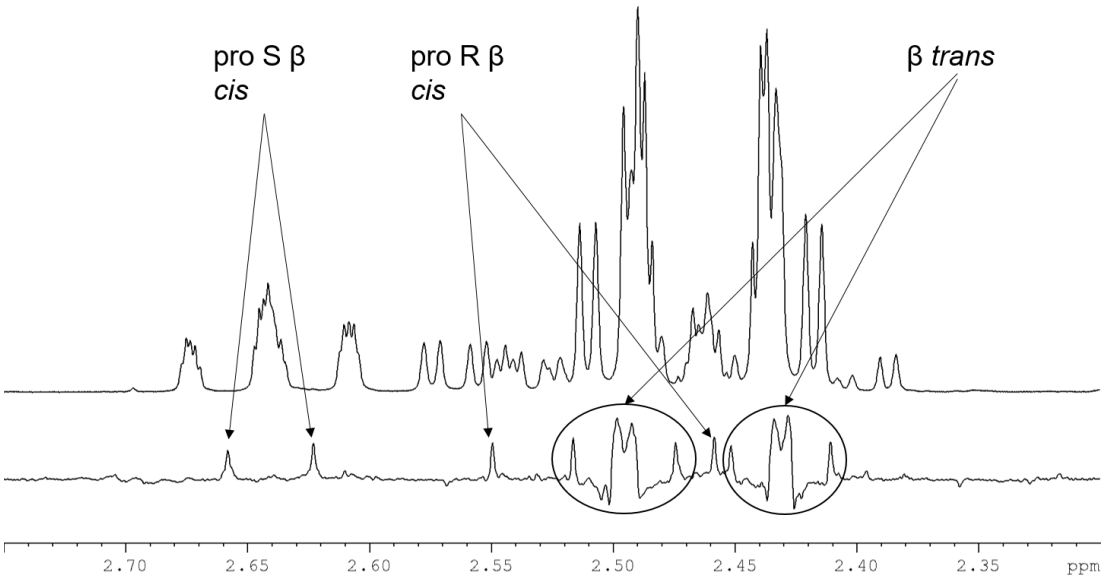


Figure 22 The β ^{1}H region of Ac-(4S)-FPro-OMe in D_2O. 298 K. 500 MHz. The two large signals are coming from the trans *conformer. Top: 1D ^{1}H. Bottom: 1D PSYCHE. Note the strong coupling artefacts (circled) in the PSYCHE.*

For Ac-(4*S*)-FPro-OMe, there is a big contrast between the two solvents in terms of overlap. In $CDCl_3$, there is overlap between the δ protons of the *trans* and *cis* conformers, but fortunately there are no strong coupling complications, since the pro-*R* and the pro-*S* within one conformer are always resolved. However, in D_2O, the β protons of the *trans* conformer are very strongly coupled and nearly indistinguishable, even in the pure shift spectrum (*vide infra*). (see figure **22**).

For Ac-4,4-F_2Pro-OMe in D_2O, the pro-*R* and pro-*S* δ signals appear to have practically the same chemical shift (see Figure **23**). In $CDCl_3$, there are no significant overlap problems.

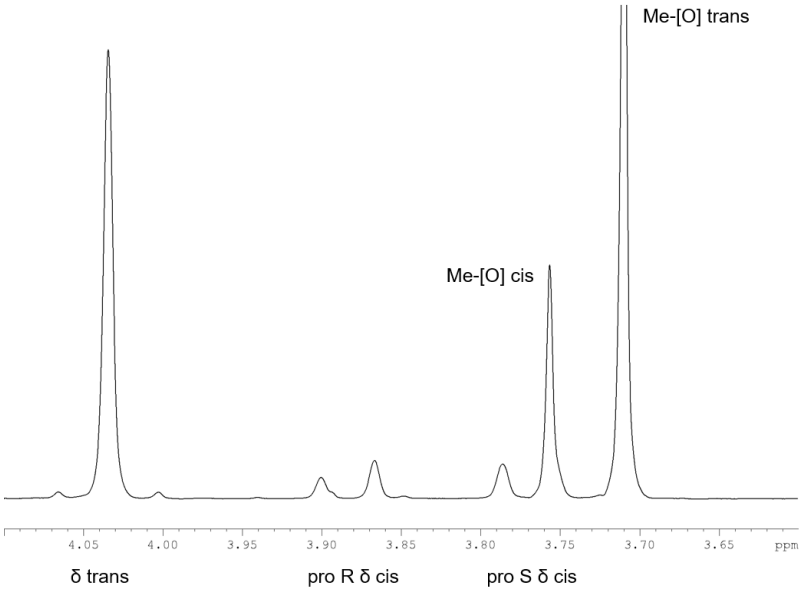


Figure 23 ^{19}F decoupled 1D ^{1}H spectrum of Ac-4,4-F_2Pro-OMe in D_2O. 298 K. 400 MHz. Note the very steep roof effect of the δ of the trans conformer.

Overlap can be most often resolved by using pure shift experiments [29] like PSYCHE [30] and PSYCHEDELIC [31] (*vide infra*). Strong coupling, however, remains an issue in these experiments, as is clear from Figures **17**, **22** and **29**. This will be elaborated in more detail in a later section (**2.3.1.2**).

The ^{1}H chemical shift values of all samples can be found in the supporting information (section **6.2**).

2.2.3. General assignment strategy of the ^{13}C spectra

Starting from the ^{1}H assignment, one can use the 2D ^{1}H-^{13}C HSQC to obtain the ^{13}C chemical shift values of all carbons, except the quaternary ones (see Figure **24**). To assign these quaternary carbons, the 2D ^{1}H-^{13}C HMBC was used. The ^{1}H signals of both the acetyl group and the methoxy group will show correlations to ^{13}C resonances at about 170 ppm, which is the expected position for carbonyl carbons of ester and amide groups. These correlations thus complete the ^{13}C chemical shift values of all compounds, which can be found in the supporting information (section **6.3**).

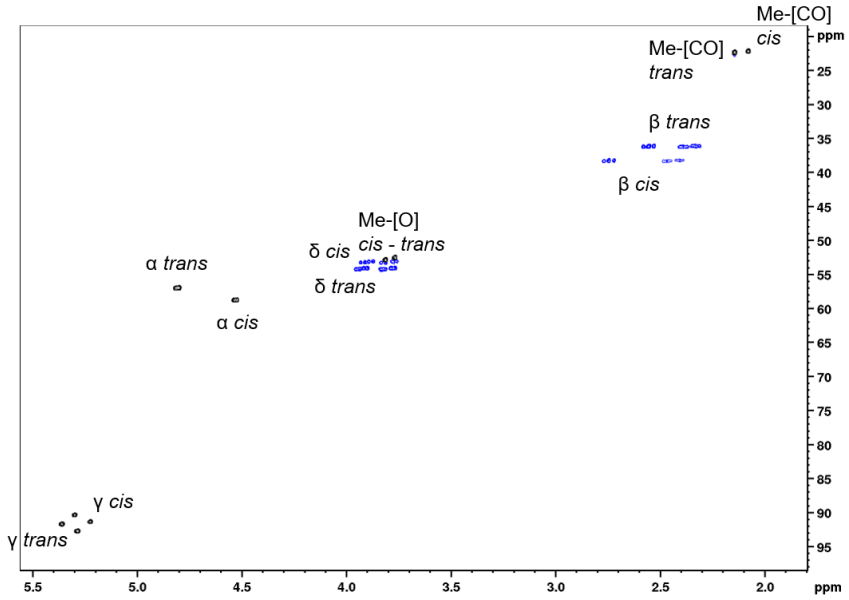


Figure 24 2D ^{1}H-^{13}C HSQC of Ac-(4S)-FPro-OMe in $CDCl_3$. 298K. 700 MHz.

2.2.4. General assignment strategy of the ^{19}F spectra

The assignment of the ^{19}F spectra of the monofluorinated samples is straightforward, only two peaks are found, a major and a minor one, which correspond to the *trans* and *cis* conformer respectively. The use of ^{1}H decoupled spectra was a prerequisite in some cases, as the many ^{1}H-^{19}F couplings led to significant overlap between both multiplets (see Figure **25**). In case of the difluorinated samples, four signals – two minor (*cis*) and two major (*trans*) doublets – are observed in the ^{1}H decoupled spectra. The doublets result from the $^{2}J_{FF}$ coupling and show distorted intensities (the "roof" effect) due to the onset of strong coupling conditions. In this case, a stereospecific assignment of both ^{19}F signals is required, which can be achieved using the 2D ^{1}H-^{19}F HOESY spectrum, as described in section **2.2.2**. ^{19}F chemical shift values can be found in the supporting information (section **6.4**).

2.3. Measuring coupling constants

The largest portion of experimental work was devoted to measuring coupling constants, as these are essential for the puckering analysis (see Chapter **3**). The way these coupling constants were measured will be discussed in this section, as well as the complications that arose. The theory behind scalar couplings was already discussed in section **2.1**.

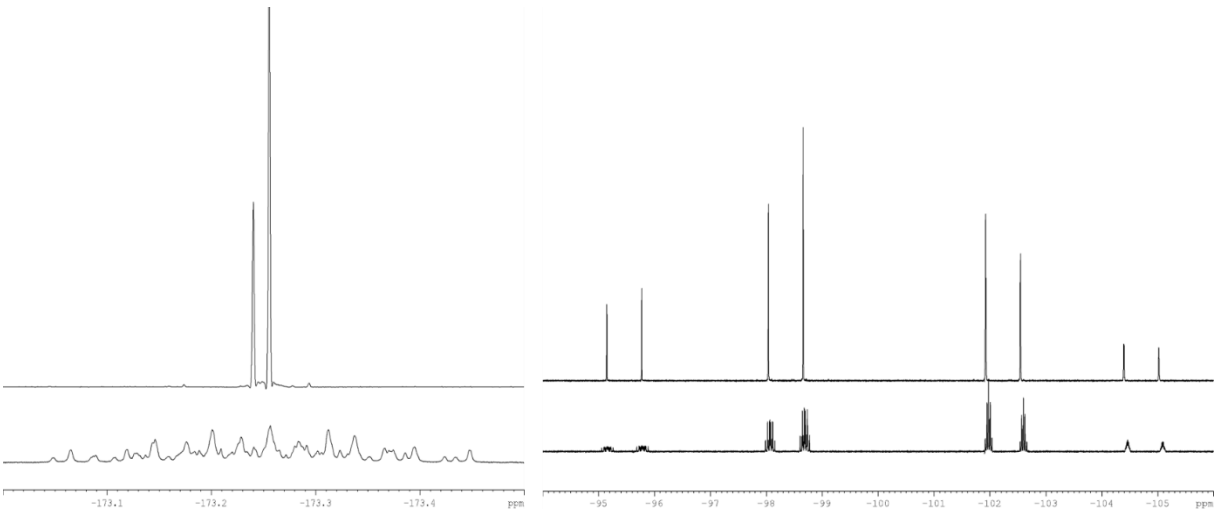


Figure 25 Left: 1D ^{19}F spectrum without (bottom) and with (top) ^{1}H decoupling of Ac-(4S)-FPro-OMe in D_2O. Right: 1D ^{19}F spectrum with (bottom) and without (top) ^{1}H decoupling of Ac-4,4-F_2Pro-OMe in D_2O. Both recorded at 298 K and on a 500 and 400 MHz spectrometer respectively.

2.3.1. Measuring ^{1}H-^{1}H couplings

2.3.1.1. PSYCHEDELIC

To measure ^{1}H-^{1}H couplings, the recently developed PSYCHEDELIC experiment was used [31]. PSYCHEDELIC stands for Pure Shift Yielded by CHirp Excitation to DELiver Individual Couplings. In this experiment, which is based on the PSYCHE pure shift experiment [30], a ^{1}H is selected, and a 2D J-resolved spectrum is generated that contains only the couplings to this ^{1}H as simple doublets (see Figure **26**). In the 2D J spectrum, the doublets are dispersed along a –45° pattern. This can be tilted by 45° to align the splitting fully parallel along the F_1 dimension, leaving only chemical shift information along F_2, and providing a resolution similar to the PSYCHE pure shift experiment. In principle, one can excite multiple ^{1}H's at the same time, but to retain the simple doublet splittings, and thus unambiguous interpretation, these ^{1}H's may not have mutual coupling partners amongst the ^{1}H's to which the couplings are measured. For example, one can always simultaneously select the same ^{1}H of both the *cis* and *trans* conformer safely, since these belong to different spin systems (see Figure **26**).

In order to actually measure the value of the scalar coupling, a slice along F_1 was extracted from the spectrum for each doublet and smoothed in order to more accurately measure the distance between the two peaks than in the 2D spectrum. This smoothing consists of increasing the digital resolution of the extracted slice by applying an inverse Fourier transform, back to the time domain, application of additional zero-filling, and finally Fourier transformation of the result to generate the smoothed trace. Further

Lorentz-to-Gauss resolution enhancement could also be applied prior to this step (see Figure **27**). Now, a more accurate estimate of the peak splitting and thus the coupling constant can be made.

The values of all measured coupling constants can be found in the supporting information (section **6.5**).

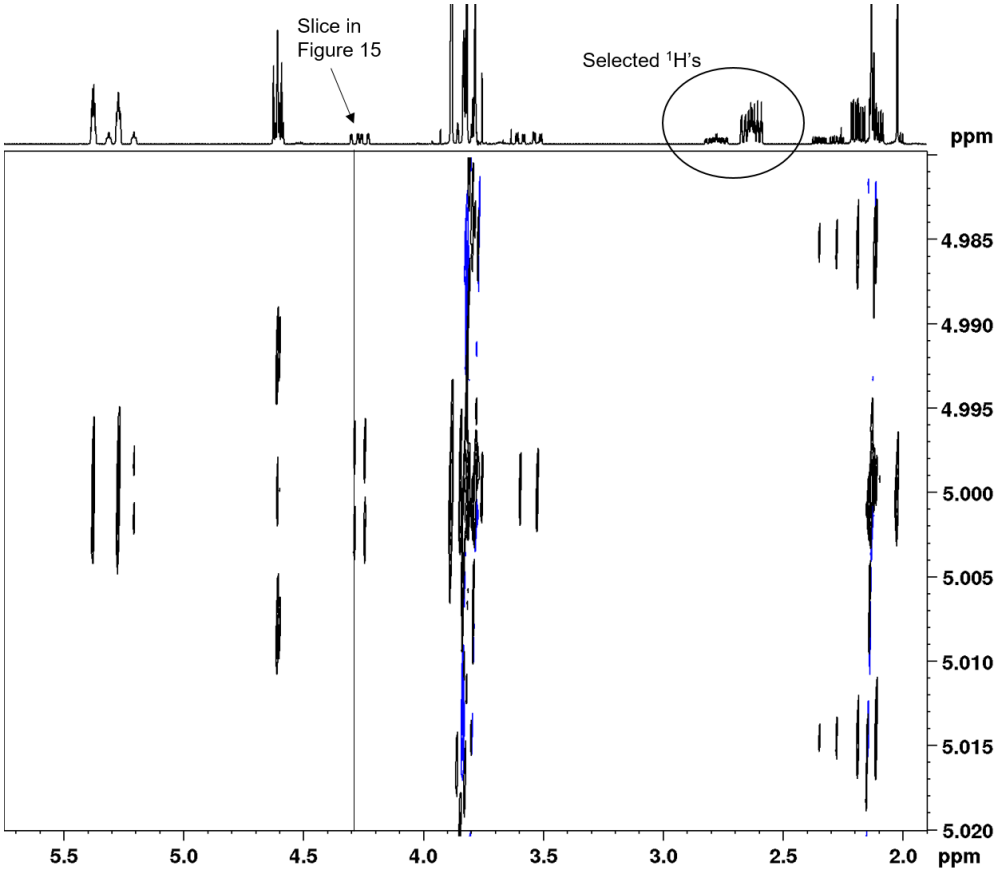


Figure 26 PSYCHEDELIC of Ac-(4R)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz. pro R β (major + minor) excitation.

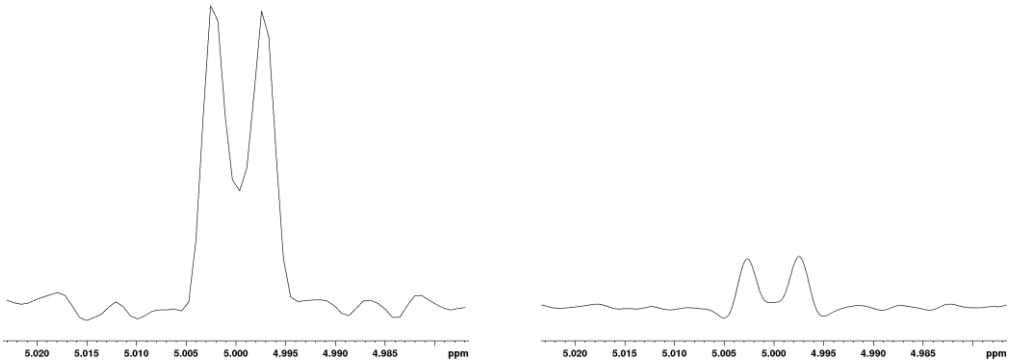


Figure 27 Left: Extracted 1D slice of Figure 15. Right: Same 1D slice after smoothing. There is an intensity decrease due to a Lorentz-to-Gauss window function that was additionally applied.

2.3.1.2. Encountered difficulties

A first difficulty that one can encounter is that the coupling is so small that the peak splitting observed along F_1 is close to or smaller than the linewidth of an individual signal (which here is typically about 1 Hz). In that case, the two peaks of the doublet cannot be resolved, as in Figure **27**, and possibly only one peak maximum is seen, such as in Figure **28**. Sometimes this can be overcome by applying further Lorentz-to-Gauss resolution enhancement. The latter not only transforms the Lorentzian lineshape to the more favourable Gaussian one, it can also trade sensitivity for resolution (see Figure **28**). However, sometimes even this did not help. In such cases, the only information that could be reported, was an upper limit for the coupling of 1 Hz.

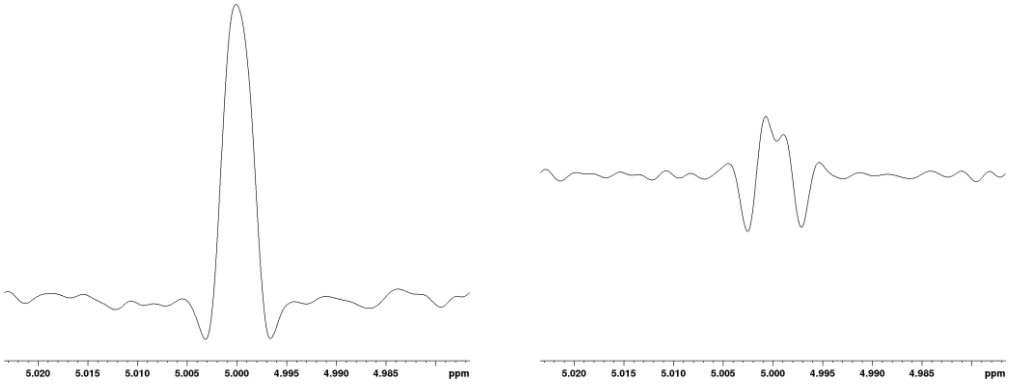


Figure 28 Ac-(4R)-FPro-OMe in CDCl3. 298 K. 500 MHz. γ excitation, pro S δ observation. Left: line broadening parameter set to -1.5. Right: line broadening parameter set to -5.

Another problem that was encountered is spectral overlap, which was already discussed in section **2.2.2.1.2**. When there is overlap between multiplets, the selective radio-frequency pulse used to select protons in PSYCHEDELIC cannot distinguish between both protons, selecting them both. As mentioned before, this is only an issue when these two protons have a mutual coupling partner, as then no longer simple doublets will be obtained along F_1.

A final, and more serious, problem arises when protons are very strongly coupled, *i.e.*, the coupling between them is of the same order of magnitude as their resonance frequency difference. When a proton is selected in PSYCHEDELIC that couples to both of these protons, strong coupling artefact peaks arise at the position of the strongly coupled spins, impeding straightforward spectral interpretation and coupling measurement (see Figure **29**). If the strongly coupled protons possess practically the same chemical shift, a peak splitting may be obtained that is assumed to be close to

the average of the two couplings. If, on the other hand, the two strongly coupled spins are selected in PSYCHEDELIC, the sum of their couplings to a third spin can be measured on their mutual coupling partner. There were four such very problematic cases. In three of them, either a sum or average could still be measured. These cases were: the γ protons of the *cis* conformer of Ac-Pro-OMe in $CDCl_3$, the δ protons of the *trans* conformer of Ac-(4*R*)-FPro-OMe in $CDCl_3$, and the β protons of the *trans* conformer of Ac-(4*S*)-FPro-OMe in D_2O. The case were no couplings could be measured was the γ protons of the *trans* conformer of Ac-Pro-OMe in D_2O, as these two very strongly coupled protons were additionally overlapping (and thus strongly coupled) with one of the two β protons. In general, accurately measuring couplings between and with strongly coupled spins still remains a methodological challenge.

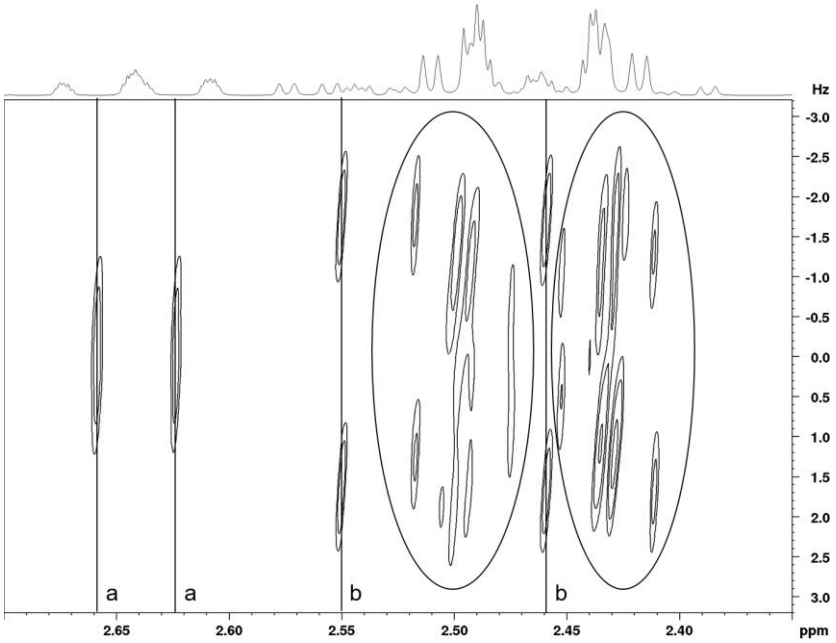


Figure 29 Ac-(4S)-FPro-OMe in D_2O. 298 K. 500 MHz. γ excitation, β region. a: pro S β cis. b: pro R β cis. Strong coupling artefacts can be observed (circled) for the trans conformer.

2.3.2. Measuring ^1H-^{19}F couplings

For the monofluorinated molecules, ^1H-^{19}F couplings can simply be measured from a 1D PSYCHE pure shift experiment [32]. In a pure shift experiment, ^1H-^1H splittings have been eliminated, meaning only the ^1H-^{19}F couplings, will be present as doublets and thus easily measurable (see Figure **30**). For Ac-4,4-F_2Pro-OMe, there are two fluorines present, meaning that in the 1D PSYCHE, the peaks of ^1H's that couple with ^{19}F will be split into a doublet of doublets. This is problematic, because even if one can then accurately extract these couplings from the spectrum, one in principle cannot

know which of the fluorine atoms is involved in the particular coupling. An alternative experiment is a heteronuclear ^{1}H-^{19}F PSYCHEDELIC experiment developed in our group (unpublished result). The experiment works similar as a homonuclear PSYCHEDELIC experiment, except that now a single fluorine is selected, so that ^{1}H-^{19}F splittings will be observed along F_1, while all ^{1}H-^{1}H and ^{1}H-^{19}F splittings are suppressed along F_2.

All ^{1}H-^{19}F coupling values can be found in the supporting information (section **6.6**).

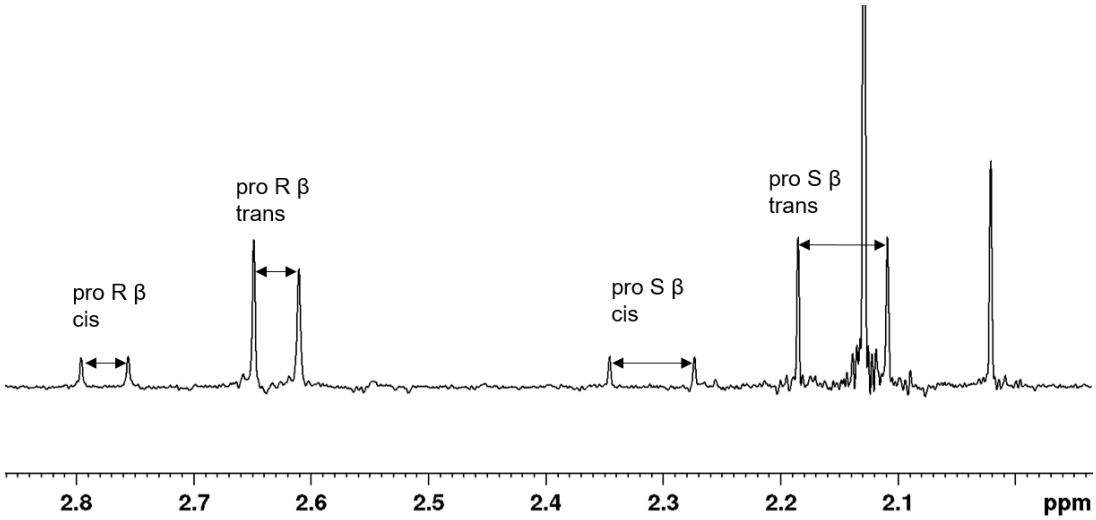


Figure 30 1D PSYCHE Ac-(4R)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

^{1}H-^{19}F couplings are generally larger than ^{1}H-^{1}H couplings, so there were never any issues with them being smaller than the linewidth. However, spectral overlap and strong coupling between protons can pose a problem. Sometimes it was not possible to measure a ^{1}H-^{19}F coupling from the 1D PSYCHE due to overlap. For example, for Ac-(4*S*)-FPro-OMe in $CDCl_3$, the pro-*R* δ doublets of both the *cis* and *trans* conformer overlap and also the methoxy peaks are interfering. Fortunately, in this case, it was possible to measure the ^{1}H-^{19}F couplings from the 2D ^{1}H-^{13}C HSQC, as these signals are resolved along the ^{13}C dimension, although this is somewhat less accurate. In some cases where couplings could not be measured in the 1D PSYCHE, the HSQC could also not provide a solution. For the *trans* conformer of Ac-4,4-F_2Pro-OMe in D_2O, the two δ ^{1}H's are very strongly coupled. Similarly as described above for homonuclear PSYCHEDELIC, it was in this case only possible to measure an average of the couplings.

2.4. NMR techniques for measuring relaxation and kinetics

In this section, a short description will be provided of the techniques used in Chapter **4** for the study of the *cis/trans* exchange process.

2.4.1. Inversion recovery

When the z-magnetization of a spin is perturbed from its equilibrium value (I_z^0) after an rf-pulse, it will relax back to its equilibrium state with a unique longitudinal relaxation time constant T_1. The inversion recovery experiment is used to measure the value of T_1 for each resonance in the spectrum. After a *p* pulse, the z-magnetization undergoing T_1 relaxation as a function of time can be described by the following equation.

$$I_z(t) = (1 - \cos(p))I_z^0 e^{-\frac{t}{T_1}}$$

In the inversion recovery experiment, a 180° pulse is applied, which will invert the z-magnetization (cos(*p*) = -1), which will then start to relax back to its equilibrium state according to the above equation. After a certain delay, a 90° pulse is applied and acquisition is performed. The intensity of the observed signal will depend on the duration of the delay in the same way as the above equation. By doing a series of such measurements with a variable delay, one can plot the obtained intensities as of function of the delay, and fit the above function to the data points. In that way, one can retrieve the value for T_1.

2.4.2. Selective inversion recovery

When two spins are in chemical exchange with one another, the selective perturbation of the z-magnetization of one of these spins will, in turn, influence the z-magnetization of the other. The extent of this influence will depend on the exchange rate k_{ex}, which is the sum of the forward (k_1) and the backward (k_{-1}) rate constants of the exchange reaction. If one would selectively invert the z-magnetization of one of the two spins, and track the intensity change of the other spin as a function of time as in a regular inversion recovery experiment, a value for k_{ex} can be assessed. Such inversion can be achieved by either applying a selective soft 180° pulse on one of the two exchanging spins, or by applying a $90°_x$ – delay – $90°_{-x}$ pulse sequence, in which the delay is the reciprocal value of two times the frequency difference of both resonances. In the latter experiment, the original orientation of the on-resonance peak's z-magnetization will be

preserved after this sequence, while the off-resonance peak's z-magnetization will have been inverted.

2.4.3. 2D EXSY

Another method for measuring k_{ex} is a 2D ^{1}H-^{1}H EXSY (EXchange SpectroscopY). In terms of pulse sequence, an EXSY is identical to a NOESY, but the cross-peaks of interest are in this case not caused by the nOe, but by chemical or conformational exchange between the two spins. These peaks will always possess the same sign as the diagonal peaks. The value for k_{ex} can be retrieved from the diagonal and cross-peak peak volumes after a certain EXSY mixing time.

A more theoretical explanation concerning selective inversion recovery and EXSY will be provided in Chapter **4**, as well as the reason why different kinds of experiments were used for the same purpose.

3. Puckering analysis

In this chapter, information about the puckering equilibrium of the studied molecules is derived from the measured couplings. First, the mathematical description of five-membered ring conformations will be addressed. Next, the relation between scalar couplings and torsion angles will be discussed. After that, data fitting procedures will be explained. Lastly, the results and conclusions for the studied molecules will be presented.

3.1. Mathematical description of five-ring conformations

The phenomenon of ring pucker – the deviation of planarity – was already discussed in Chapter **1**. The question is now, how to conveniently describe the ring pucker. Annotations like C$^{\gamma}$ *endo* and C$^{\gamma}$ *exo* are very limited given the amount of conformations that can potentially be described, and are also only relevant for proline related systems. It is clear that a more general way of describing conformations is needed. In general, in order to describe an *N*-membered ring conformation, *N*-3 parameters are needed, this means that two parameters will be needed for five-membered rings [33]. There are two formalisms that propose such parameters. The Cremer-Pople formalism, which is based on Cartesian coordinates [34], and the Altona-Sundaralingam formalism, which is based on endocyclic torsion angles [35]. Since it is more convenient to relate torsion angles to scalar couplings in a later step, the Altona-Sundaralingam formalism will here be preferred over the Cremer-Pople formalism [36].

The Altona-Sundaralingam formalism reduces the description of the ring pucker to two puckering parameters *P* and ν_{max} that are correlated to the endocyclic torsion angles ν_i (i = 0 – 4) in the following manner:

$$\nu_i = \nu_{max} \cos\left(P + \frac{4\pi}{5} i\right)$$

This is in fact a set of five equations, one for each endocyclic torsion angle. The inverse relationships to obtain *P* and ν_{max} from the torsion angles are:

$$\tan(P) = \frac{-\nu_1 + \nu_2 - \nu_3 + \nu_4}{2\nu_0 \left(\sin\left(\frac{\pi}{5}\right) + \sin\left(\frac{2\pi}{5}\right)\right)}$$

$$\nu_{max} = \frac{\nu_0}{\cos(P)}$$

As is evident from the above formula, *P* will be dependent on the numbering scheme. Throughout this thesis, the numbering scheme as depicted in Figure **31**, was used [36].

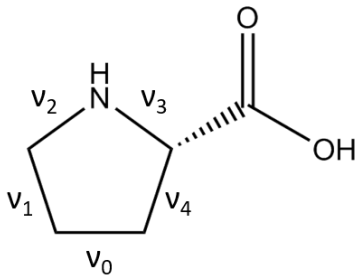


Figure 31 Used numbering scheme for the endocyclic torsion angles.

The two puckering parameters proposed by Altona and Sundaralingam can be interpreted in the following manner. *P* denotes the pseudorotation phase angle (0° – 360°), and indicates which atom(s) is (are) out of plane. When *P* is an odd multiple of 18°, the conformation will be a true envelope, while even multiples of 18° correspond to pure twist conformations. Of special interest for this investigation are the *P* values for the C^{γ} *exo* and C^{γ} *endo* conformations, which are 18° and 198° respectively for the numbering scheme used here. ν_{max} denotes the puckering amplitude (≥ 0°), and indicates how much the atom(s) is (are) tilted out of plane. A puckering amplitude of 0° would thus indicate that the ring is flat [36].

Because of the periodic nature of the pseudorotation phase, polar plots are very convenient for depicting five-membered ring conformations (see Figure **32**). *P* will be used as the angular coordinate and ν_{max} will be used as the radial coordinate. Such polar plots will be called pseudorotation plots or pseudorotation wheels [36].

3.2. The relationship between torsion angles and scalar couplings

As mentioned before, scalar couplings will be influenced by torsion angles, which is the reason why conformational analysis can be performed by measuring coupling constants. In 1959, Martin Karplus proposed a simple mathematical relationship between $^3J_{HH}$ coupling constants and torsion angles (φ), based on the comparison of experimental coupling constants and theoretical valence bond calculations [37]. Expressions relating scalar couplings and torsion angles are called Karplus relations, named after its pioneer.

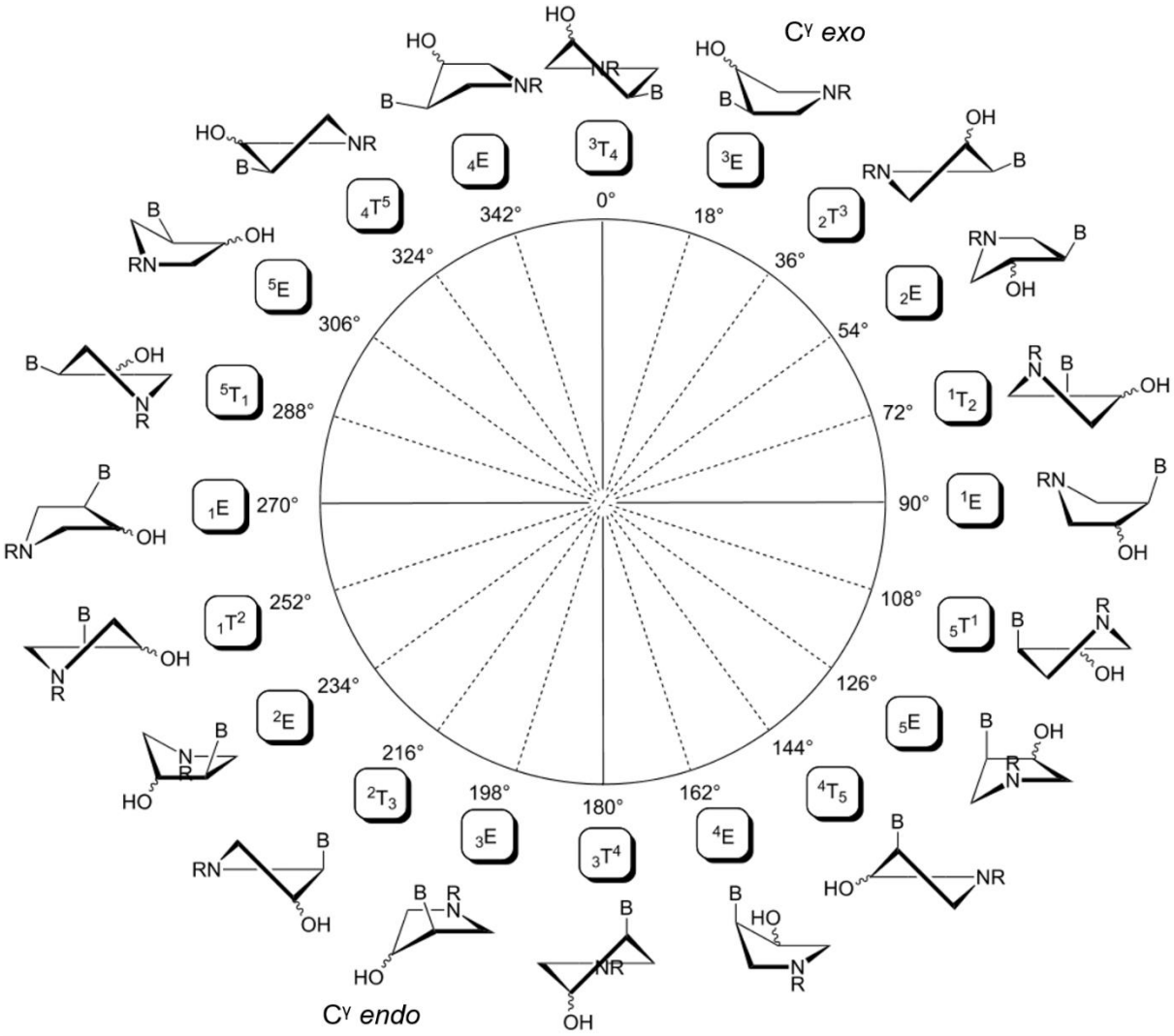


Figure 32 Pseudorotation plot with indicated positions of the 10 possible envelope and 10 possible twist conformations. This figure was taken from [38], with some small adaptations. With the used numbering scheme, the carbon carrying the B substituent is the β carbon, and the carbon with the OH substituent is the γ carbon.

$$^{3}J = A\cos^{2}(\varphi) + B\cos(\varphi) + C$$

The values for the *A*, *B* and *C* constants are derived empirically for specific circumstances, for example, for certain substituents. Over the years, a lot of refinements have been made with respect to this simple equation, mainly concerning the effects of substituents, which have the largest impact on the scalar coupling next to the torsion angle [36]. The substituent effects have been parameterized by the so-called group electronegativities λ_i. These are used in the Diez-Donders equation, which is a generalized Karplus equation for J_{HCCH} couplings and which was used throughout the course of this thesis [39, 40].

$$J_{HCCH} = C_0 + C_1\cos(\varphi) + C_2\cos(2\varphi) + C_3\cos(3\varphi) + S_2\sin(2\varphi)$$

$$C_0 = 6.97 - 0.58\sum_i \lambda_i - 0.24(\lambda_1\lambda_2 + \lambda_3\lambda_4)$$

$$C_1 = -1.06$$

$$C_2 = 6.55 - 0.82\sum_i \lambda_i + 0.20(\lambda_1\lambda_4 + \lambda_2\lambda_3)$$

$$C_3 = -0.54$$

$$S_2 = 0.68\sum_i \xi_i\lambda_i^2$$

Values for λ for various substituent groups have been deduced by Diez *et al.* ξ is 1 for odd subscript values and -1 for even subscript values. As is evident from the equations, C_0 and C_2, and thus J_{HCCH}, are dependent on the substituent numbering scheme, which can be seen in Figure **33**.

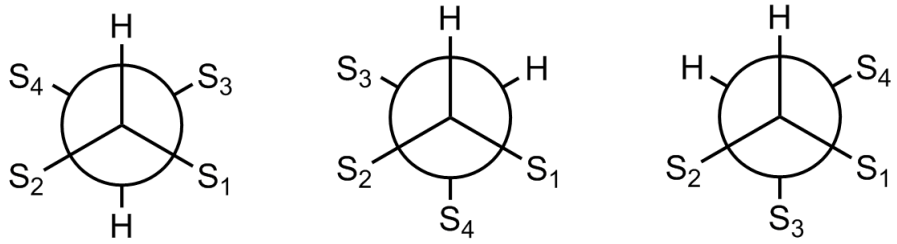


Figure 33 Substituent numbering scheme used in the Diez-Donders equation. Figure taken from [36].

Marshall *et al.* studied the non-equivalence of *exo-exo* and *endo-endo* couplings in norbornanes and found that non-bonded interactions can also have an influence on scalar couplings [41]. This so-called Barfield effect can be explained from Figure **34**. The reduced distance between H_1/H_2 and X will reduce the $^3J_{H1H2}$ scalar coupling, however it will not affect the $^3J_{H3H4}$ scalar coupling, nor any of the transoid couplings. It has been proposed to correct for the Barfield effect by adding a correction term to the Karplus relation.

$$\Delta J = T_b\cos^2(P - P_b)$$

T_b is dependent on the nature of atom X, and is 1 Hz for nitrogen and 2 Hz for carbon. P_b is the pseudorotation phase for which the Barfield effect will be the most pronounced. In case of the numbering scheme used, this will be 234° for α-β couplings, 90° or 270° for β-γ couplings (above or below the ring respectively), or 306° or 126° for γ-δ couplings (above or below the ring respectively). Note that the Barfield

correction only needs to be applied for cisoid couplings and when $|P - P_b| \leq 90°$, and that it will always lead to a reduction of the scalar coupling [36].

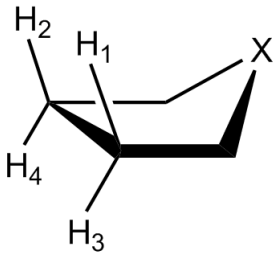


Figure 34 Illustration of the Barfield effect in a five-membered ring. Figure taken from [36].

To relate ^{1}H-^{19}F couplings to H-C-C-F torsion angles, a Karplus relation that is also parameterized with group electronegativities was used [42]. However, this relation also needs bond angles (a_{FCC} and a_{HCC}) as parameters and obtaining sensible values for those angles can be difficult as will be seen later.

$$J_{HCCF} = 40.61\cos^2(\phi) - 4.22\cos(\phi) + 5.88 + \sum_i \lambda_i[-1.27 - 6.20\cos^2(\xi_i\phi + 0.20\lambda_i)] - 3.72\left(\frac{a_{FCC} + a_{HCC}}{2} - 110\right)\cos^2(\phi)$$

Karplus relations deliver exocyclic torsion angles. However, the Altona-Sundaralingam formalism, is built around endocyclic torsion angles. Luckily, the conversion between exocyclic and endocyclic torsion angles is straightforward.

$$\nu_{exo} = A\nu_{endo} + B$$

For an ideal tetrahedral geometry, *A* will be 1 and *B* will be 0° for cisoid couplings or ±120° for transoid couplings. When deviating from the ideal tetrahedral geometry, small corrections have to be applied to *A* and *B*. However, in this thesis, ideal tetrahedral geometry was always assumed [36].

3.3. Fitting conformations

Two fitting programs were used, a graphical user interface (GUI), written by P.M.S. Hendrickx [43], and an in-house MatLab script, written by D. Sinnaeve. Initially, the GUI was used. However, due to some shortcomings of the GUI, the new MatLab script was written. The fitting procedure in this script follows the same principle as the one used in the GUI, with the advantages that the known coding of the algorithm provides full control, ^{1}H-^{19}F couplings can be used (using the Karplus relation of Thibaudeau *et*

al. [42]), and the Barfield correction can optionally be applied. A comparison between the GUI and the MatLab script is made in the supporting information (section **6.7.1**).

3.3.1. MatLab script

To derive one conformation from a set of experimental scalar couplings, the fitting program works in the following way. Starting from initial guesses for *P* and v_{max}, endocyclic torsion angles are calculated via the Altona-Sundaralingam formalism. The obtained endocyclic torsion angles are then converted to exocyclic torsion angles between the protons for which the coupling constants were measured. These exocyclic torsion angles are then fed into the Diez-Donders equations, together with the correct λ parameters, to yield the estimated scalar coupling constants. These are subsequently compared with the experimental coupling constants by calculating the root-mean-squared difference (RMSD) as goodness-of-fit.

$$RMSD = \sqrt{\frac{1}{N}\sum_{i}^{N}(J_i^{calc} - J_i^{expt})^2}$$

An optimization procedure is used to update the initial *P* and v_{max} parameters in order to minimize the RMSD. These updated parameters are compared to the original parameters and convergence is checked. Iteration of the previously described steps is performed until convergence is reached [36].

An equilibrium between two conformations can also be fitted, but then there are some differences. Five parameters now need to be fitted, so initial values have to be entered for P_A, P_B, $v_{max,A}$, $v_{max,B}$ and x_A. The latter is the fraction of conformer A, from which the fraction of conformer B follows, since $x_A + x_B = 1$. All the previously described steps up until the calculation of the coupling constants are now performed for two conformers separately. Then, an extra step is added in which weighted average coupling constants are calculated, using the following formula.

$$J_i^{calc} = x_A J_i^A + x_B J_i^B$$

Now these calculated scalar coupling constants are used to calculate the RMSD. The optimization procedure remains the same [36, 43].

The MatLab script also allows estimating the impact of the precision of the coupling measurement on the fitting. This is done using a Monte Carlo analysis [44]. In this procedure, within 200 (one conformer) or 500 (two conformers) iterations, Gaussian noise with a spread equal to a precision of 0.1 Hz (see supporting information, section **6.7.2**) is added to the experimental coupling values in a semi-random fashion. These modified coupling constants are then used to fit the conformation(s), so in fact 200 or 500 conformations will be fitted. This multitude of fittings, will be used for two purposes. First, they allow the calculation of 95% uncertainties of the fitted parameters. Second, the results from the 200/500 different one/two conformations can be plotted in a pseudorotation plot, providing a visual indication of the uncertainty on the fitted parameters due to the measurement precision.

3.3.2. Approximations and limitations

It is important to realize that, even when the group electronegativities in the generalized Diez-Donders equations are known, a Karplus equation will always be an approximate description of the relation between torsion angle and scalar coupling, and it is thus limited in its accuracy. For the fluoroproline systems, further complications arise that will negatively affect the accuracy of the Karplus relations. First, not all group electronegativity values required, have been reported in literature. This means that for some substituents, approximations for this value had to be made (see supporting information for details), giving the Karplus relation an even more approximate nature.

Furthermore, bond angles (a_{FCC} and a_{HCC}) are needed as input for the ^{1}H-^{19}F Karplus relation. Fluorination can indeed significantly perturb the ideal tetrahedral bond angles, and this is not easy to estimate. A solution would be resorting to crystal structures, however, only for Ac-(4*R*)-FPro-OMe has this been reported in literature [45]. For the other fluorinated analogues results from quantum chemical calculations had to be used. These were kindly provided by dr. Ilya Kuprov (Southampton University). Using the Gaussian program, he performed geometry optimizations for the four molecules with DFT, using the M06 functional on the cc-pVDZ level of theory. From this, sensible values for the required bond angles could be obtained, which are listed in the supporting information.

3.4. Results and discussion

Because the results obtained via both the GUI and the MatLab script (when ^{1}H-^{19}F couplings are not used) are very similar, only the results from the MatLab script will be discussed here. All fit results (pseudorotation plots, fitted parameters, 95% uncertainties, RMSD values and residuals) can be found in the supporting information (sections **6.7.4** and **6.7.5**). The input parameters used can also all be found in the supporting information. For obtaining initial *P* and v_{max} values, a preliminary scan of the parameter space was performed, so the actual fitting can start close the RMSD minimum.

3.4.1. Ac-Pro-OMe

For Ac-Pro-OMe, no fitting could be performed for the *trans* conformer in D_2O as only two couplings could be measured, due to strong coupling artefacts (see Chapter **2**). The *cis* conformer in $CDCl_3$ is also plagued by strong coupling artefacts, but here averages of couplings could be measured, so fitting could be performed. However, results for this particular case have to be interpreted with caution.

When trying to fit one conformation, a single well-defined conformation appears to be obtained only for the *cis* conformer in $CDCl_3$. The *P* value is 175.7°, v_{max} is 17.3° and the RMSD is only 0.6 Hz, which seems to indicate a rather good result. The fitted conformation is close to a C^{γ} *endo* conformation, which is the predominant conformer for Ac-Pro-OMe according to literature [1]. However, it is reported that Ac-Pro-OMe does not exclusively adopt a C^{γ} *endo* conformation, but rather a 7:3 ratio of C^{γ} *endo* and C^{γ} *exo* [1]. In that respect, it is unexpected that fitting one conformation to the data yields such a low RMSD. In contrast, in case of the *trans* conformer in $CDCl_3$ and the *cis* conformer in D_2O, fitting one conformation to the data does not yield a reliable result, as the uncertainties on the *P* value are of the same order of magnitude as the *P* value and the uncertainties on v_{max} are even five to six orders of magnitude larger than v_{max}. The RMSD values are also quite high for both cases (2.6 Hz and 3.2 Hz respectively). These observations all indicate that assuming one conformation for Ac-Pro-OMe in these cases cannot adequately explain the observed scalar couplings, and that indeed a second conformer should be assumed. The fact that a single conformation could adequately fit the data of the *cis* conformer in $CDCl_3$, could be caused by the high proportion of average couplings used. This significantly reduces

the restrictiveness on the system imposed by the experimental data, which might make it easier to fit one conformation. It would indeed be very surprising if such a significant difference in ring pucker equilibrium between the *cis* and *trans* form in chloroform would be reflected by only such limited differences in coupling constants between the two forms (see Table **7** supporting info).

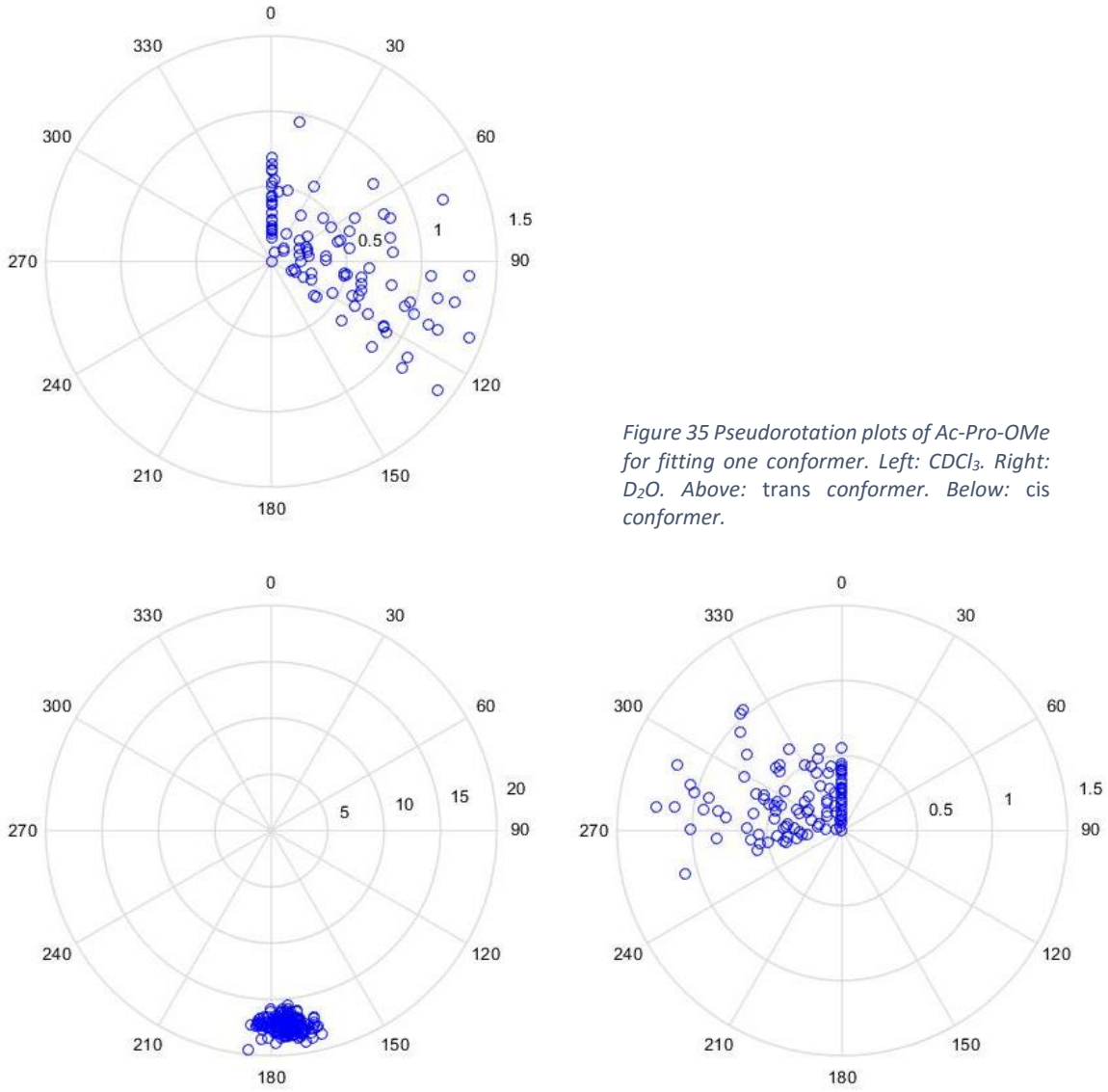


Figure 35 Pseudorotation plots of Ac-Pro-OMe for fitting one conformer. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

When assuming two conformers, the results shown in Figure **36** are obtained. Again, the fitting for the *cis* conformer in $CDCl_3$ proves to be less reliable, as it leads to rather ill-defined conformations. For the other two cases, the RMSD values have decreased compared to fitting one conformation (to 0.5 Hz and 0.8 Hz). The fitted values for *trans* Ac-Pro-OMe in $CDCl_3$ were P = 187.9°, v_{max} = 31.5° for the major pucker, and P = 17.1°, v_{max} = 49.7° for the minor pucker, with a fraction of the minor pucker of 29.5%.

The fitted values for *cis* Ac-Pro-OMe in D_2O were P = 182.8°, v_{max} = 36.4° for the major pucker, and P = 13.4°, v_{max} = 43.7° for the minor pucker, with a fraction of the minor pucker of 16.9%. From Figure **36**, one can clearly see that the uncertainty on the v_{max} value of the minor pucker is quite large in all cases, which is not very surprising. Since this conformer is only present as a small fraction, the weighted average couplings calculated in the script will only vary slightly, even for larger variations of pseudorotation parameters of the minor fraction. These obtained values very much confirm what has been already found in literature (*vide supra*), as a C^{γ} *endo* and C^{γ} *exo* conformer correspond to P values of 198° and 18° respectively. It can also be noted that the C^{γ} *exo* fraction is larger for the *trans* conformer than for the *cis* conformer, which can be expected based on the coupled equilibria described in Chapter **1**.

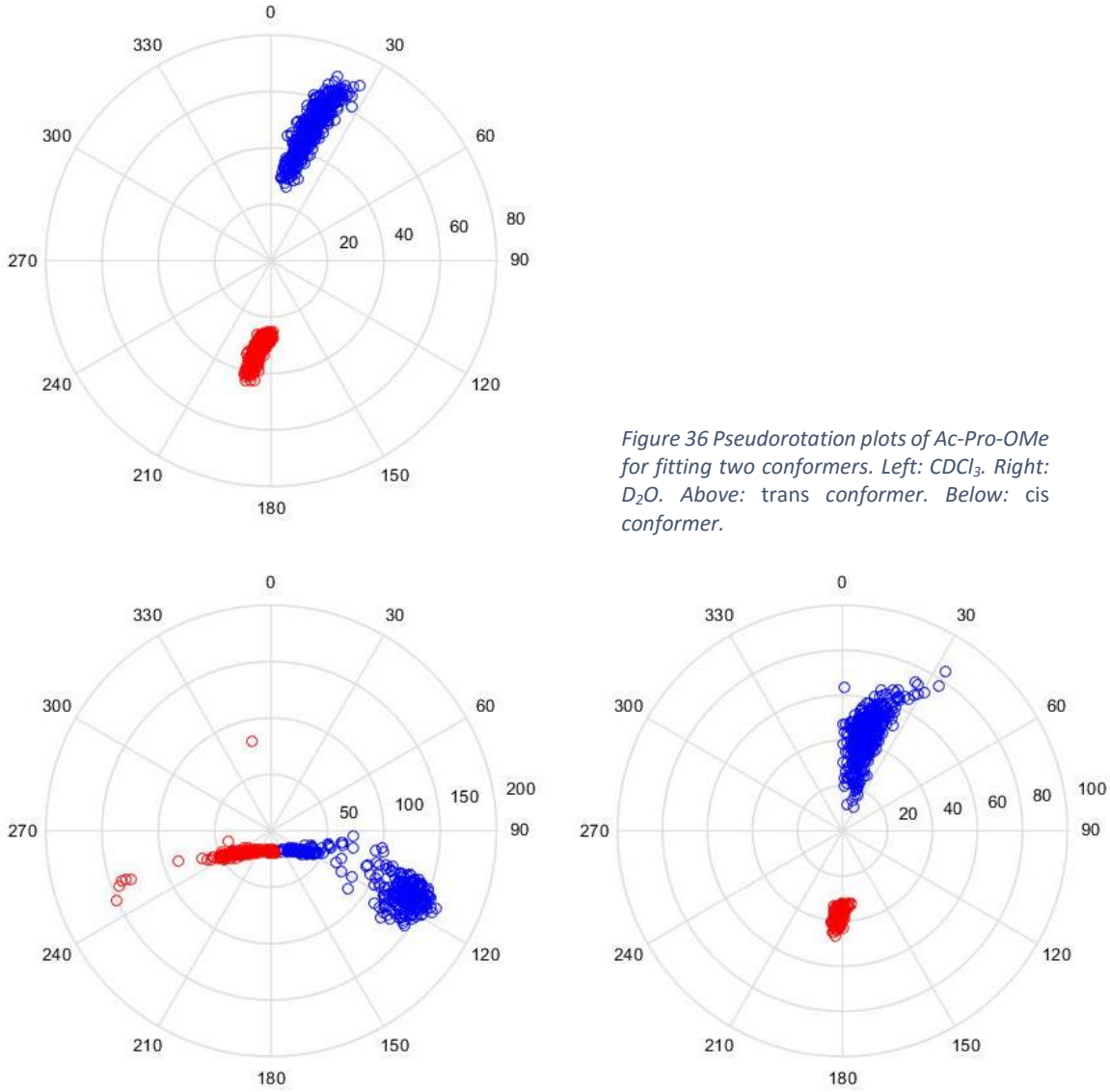
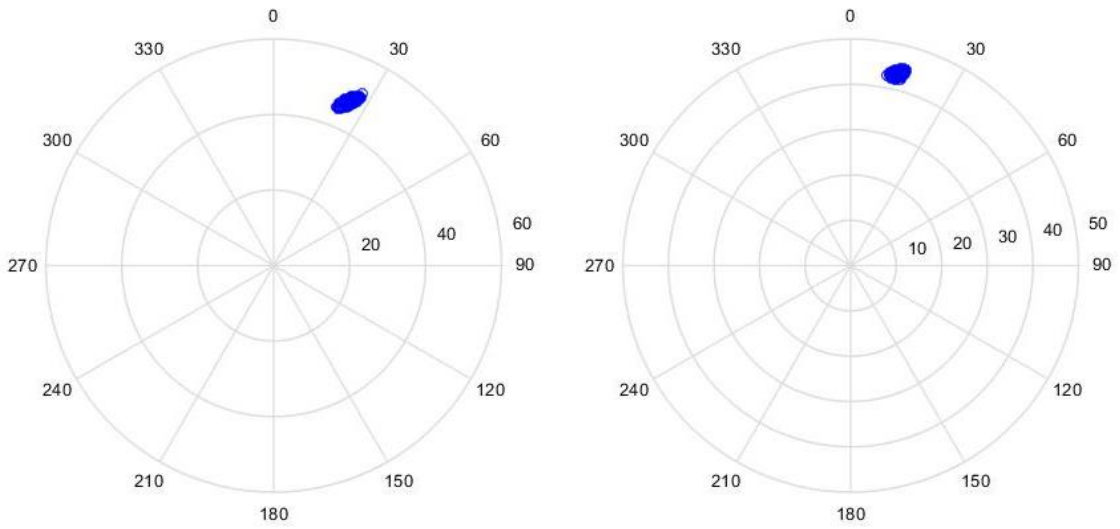


Figure 36 Pseudorotation plots of Ac-Pro-OMe for fitting two conformers. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

*3.4.2. Ac-(4*R*)-FPro-OMe*

Fitting was performed with a maximum of six ^{1}H-^{1}H couplings and four ^{1}H-^{19}F couplings. However, this maximum number of couplings could not be reached for the *trans* conformer in $CDCl_3$, as two ^{1}H-^{1}H couplings were measured as an average and two ^{1}H-^{19}F couplings could not be measured.

When fitting one conformer without using the ^{1}H-^{19}F couplings, the pseudorotation plots in Figure **37** are generated. From these plots, it can be seen that rather well-defined puckers are obtained. In all cases, similar ring puckers are found with pseudorotation phases ranging from 13.5° to 25.0° and puckering amplitudes ranging from 40.7° to 48.0°. RMSD values all fall around 0.4 Hz, except for the *trans* conformer in $CDCl_3$, where 0.94 Hz is obtained. This might be related to the fact that not all individual couplings were known in this case (*vide supra*). Nevertheless, these are quite low RMSD values, which indicates a high goodness-of-fit. The low uncertainty values on the fitted parameters also indicate that the puckering equilibrium is skewed to one side, because if a second conformer was present with reasonable percentages, the uncertainties are typically larger, as was seen for Ac-Pro-OMe. According to literature [1], the ring pucker of Ac-(4*R*)-FPro-OMe is indeed expected to show a large preference for the C^{γ} *exo* conformer, which has a pseudorotation phase of 18°. The results from the fitting thus confirm the presence of this conformer, and its predominance. It can be noted that the v_{max} values are slightly higher than for Ac-Pro-OMe.



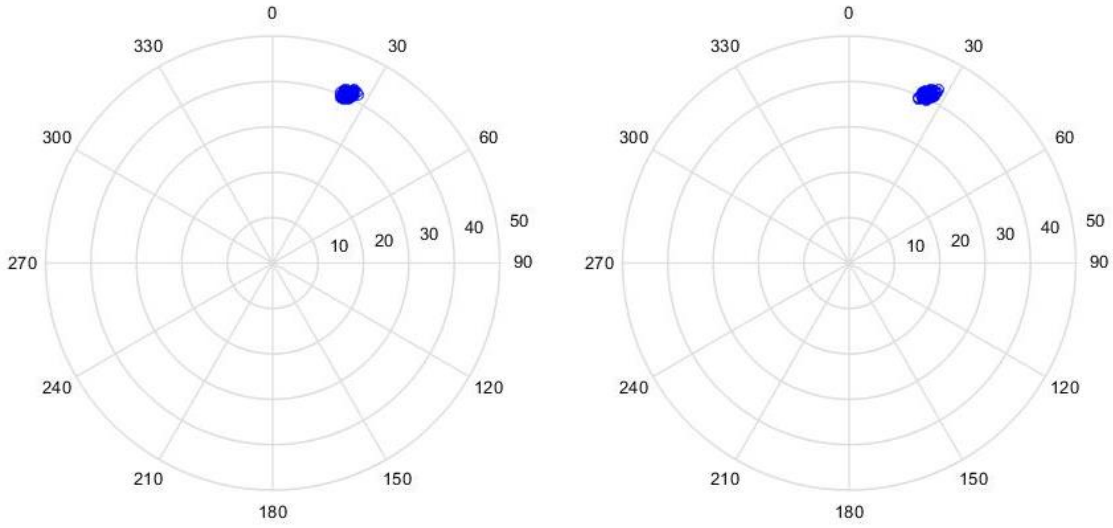


Figure 37 Pseudorotation plots of Ac-(4R)-FPro-OMe for fitting one conformer, without use of 1H-19F couplings. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

If one includes ^{1}H-^{19}F couplings in the conformational analysis, one sees that roughly the same results are obtained, which would be expected. However, the RMSD values are systematically higher (for all but one, more than 1 Hz), which may point out that the ^{1}H-^{19}F Karplus relation is not as accurate as those for ^{1}H-^{1}H, since the highest residuals were often found for ^{1}H-^{19}F couplings (all values are in the supporting information, as well as the pseudorotation plots, sections **6.7.4** and **6.7.5**).

In order to fit two conformations, ^{1}H-^{19}F couplings are needed as input as six ^{1}H-^{1}H couplings are insufficient to fit five parameters (see supporting information, section **6.7.1**). For the *trans* conformers in both solvents, the uncertainty on the fitted fraction of the minor pucker is so large (much larger than the fitted value of the fraction), that it is impossible to say that there are two conformations. For the *cis* conformers in both solvents, the uncertainty of the fraction is smaller, however the fitted v_{max} values for the minor conformer are so unreasonably high (104.0° in $CDCl_3$ and 147.2° in D_2O) that these would not be realistic conformations (some endocyclic torsion angles would be over 100°). For these cases, the major conformer coincides rather well with the C^{γ} *exo* conformation, and the uncertainties on its *P* and v_{max} values are very small. This also indicates that it is not possible to prove the presence of a second conformer, because the large uncertainties of the minor pucker imply that this pucker does not contribute much to the eventual coupling constants. It is reasonable to think that there probably will exist a puckering equilibrium with a small fraction of C^{γ} *endo*. However, with the limited number of couplings and the approximations used in the Karplus

relations, it does not appear possible to detect this. If we want to be able to confirm this hypothesis, more accurate Karplus relations or additional experimental data will be needed. Pseudorotation plots and all fitted parameters are in the supporting information (section **6.7.4** and **6.7.5**).

3.4.3. Ac-(4S)-FPro-OMe

The *trans* conformer in D_2O is expected to yield less reliable results, as two pairs of ^{1}H-^{1}H couplings had to be measured as averages, and two ^{1}H-^{19}F couplings could not be measured. For the *cis* conformer in D_2O, also two ^{1}H-^{19}F couplings could not be measured.

When looking at Figure **38**, which are the plots for fitting one conformation without use of the ^{1}H-^{19}F couplings, one can see that three of the four fittings yield approximately the same conformation. Only the *trans* conformer in D_2O, yields another conformation, but this is less well-restrained due to use of average coupling values. Discarding, this result, one can see that the other results have values for *P* ranging from 199.3° to 206.5° and values for v_{max} ranging from 37.1° to 41.6°. The uncertainties due to measurement errors are a bit higher than for Ac-(4*R*)-FPro-OMe. Nonetheless, also for this molecule it can be deduced that there will be a large preference for one conformer. According to literature [1], this should be the C^{γ} *endo* conformation, which has a *P* value of 198°. This is indeed in good agreement with the fitted results. When including ^{1}H-^{19}F couplings (again discarding the results for the *trans* conformer in D_2O), the *P* values shift to lower values for both cases in $CDCl_3$ (179.4° for *trans* and 175.8° for *cis*), which is more a twist conformation ($_{\gamma}T^{\beta}$, which means C^{γ} is below the plane and C^{β} above), but nevertheless still close to 198°. Just like for Ac-(4*R*)-FPro-OMe, the RMSD values are higher upon inclusion of ^{1}H-^{19}F couplings (see supporting info).

When fitting two conformations (necessitating the use of ^{1}H-^{19}F couplings, *vide supra*), the story for Ac-(4*S*)-FPro-OMe is very similar as for Ac-(4*R*)-FPro-OMe, although now for all four cases, the uncertainties on the fitted fractions were much higher than the values of the fitted fractions themselves. This again indicates that the presence of a second conformer cannot be proven with the currently used approximations. A similar conclusion as for Ac-(4*R*)-FPro-OMe can thus be drawn, but now with the C^{γ} *exo* conformer as the minor fraction. Again, the pseudorotation plots can be found in the supporting information (section **6.7.5**).

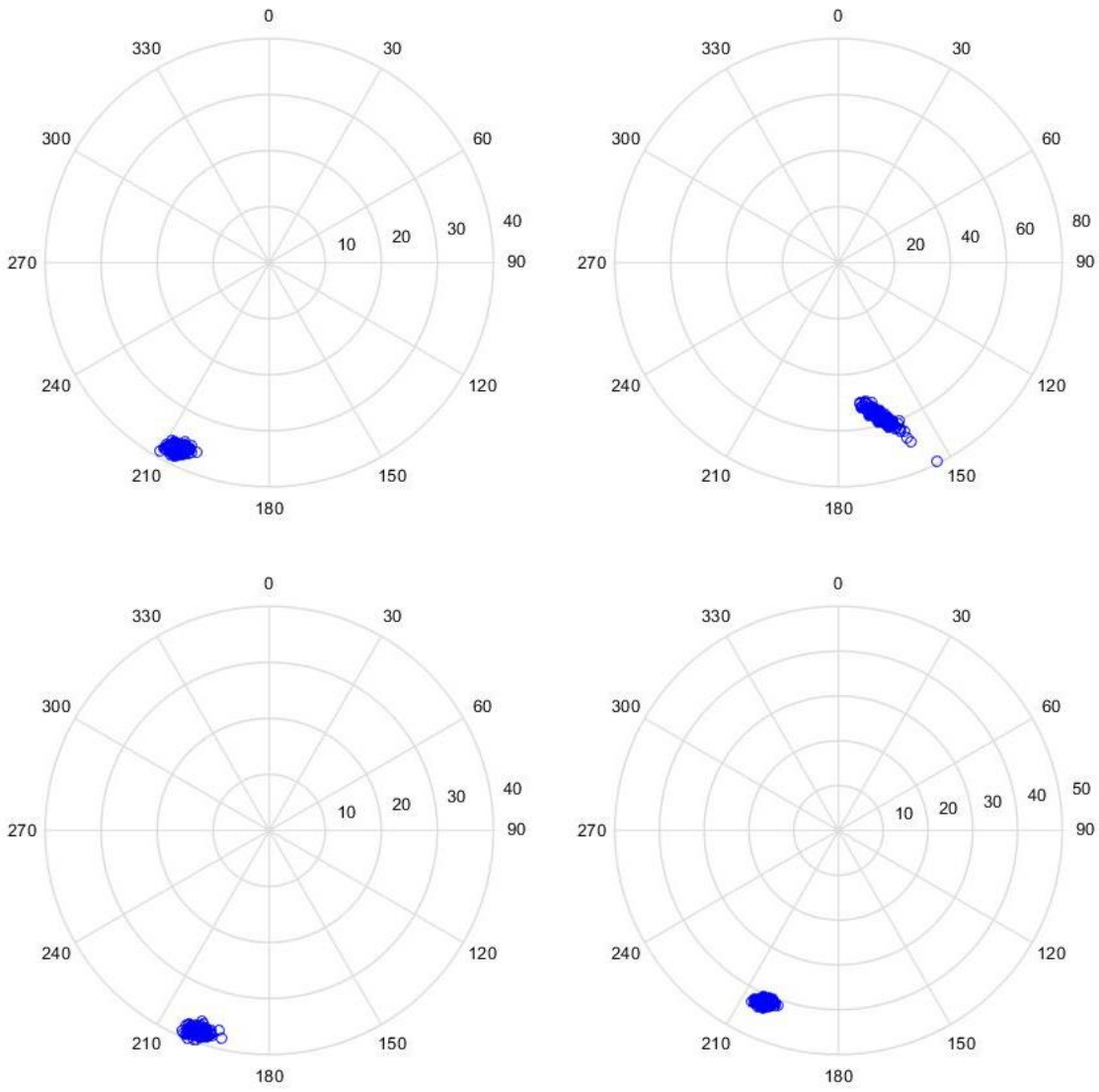


Figure 38 Pseudorotation plots of Ac-(4S)-FPro-OMe for fitting one conformer without use of ^{1}H-^{19}F couplings. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

3.4.4. Ac-4,4-F_2Pro-OMe

No fitting could be performed for Ac-4,4-F_2Pro-OMe without using ^{1}H-^{19}F couplings as the number of vicinal ^{1}H-^{1}H couplings (two) is insufficient. For the *cis* conformer in $CDCl_3$, one ^{1}H-^{19}F coupling could not be measured and for the *trans* conformer in D_2O, two pairs of ^{1}H-^{19}F couplings were measured as averages.

When fitting one conformation, except for the *trans* conformer in $CDCl_3$, well-defined conformations could be fitted, which for the two *cis* conformers are close to a C^{γ} *endo* conformation (P = 167.7°; ν_{max} = 21.3°, in $CDCl_3$ and P = 188.1°; ν_{max} = 29.4°, in D_2O). For the *trans* conformer in D_2O, a rather unexpected conformation is fitted (P value is 133.9°). However, the reliability of this result is reduced because of the featured

average couplings in the dataset. According to literature [4], Ac-4,4-F_2Pro-OMe should be close to Ac-Pro-OMe in terms of conformational behaviour, with a predominant C^{γ} *endo* conformation, but also a considerable fraction of C^{γ} *exo*. Just like for Ac-Pro-OMe, high RMSD values are obtained (even for both *cis* conformers, 4.1 Hz in $CDCl_3$ and 2.7 Hz in D_2O), which indeed suggests that the assumption of one conformer is inadequate.

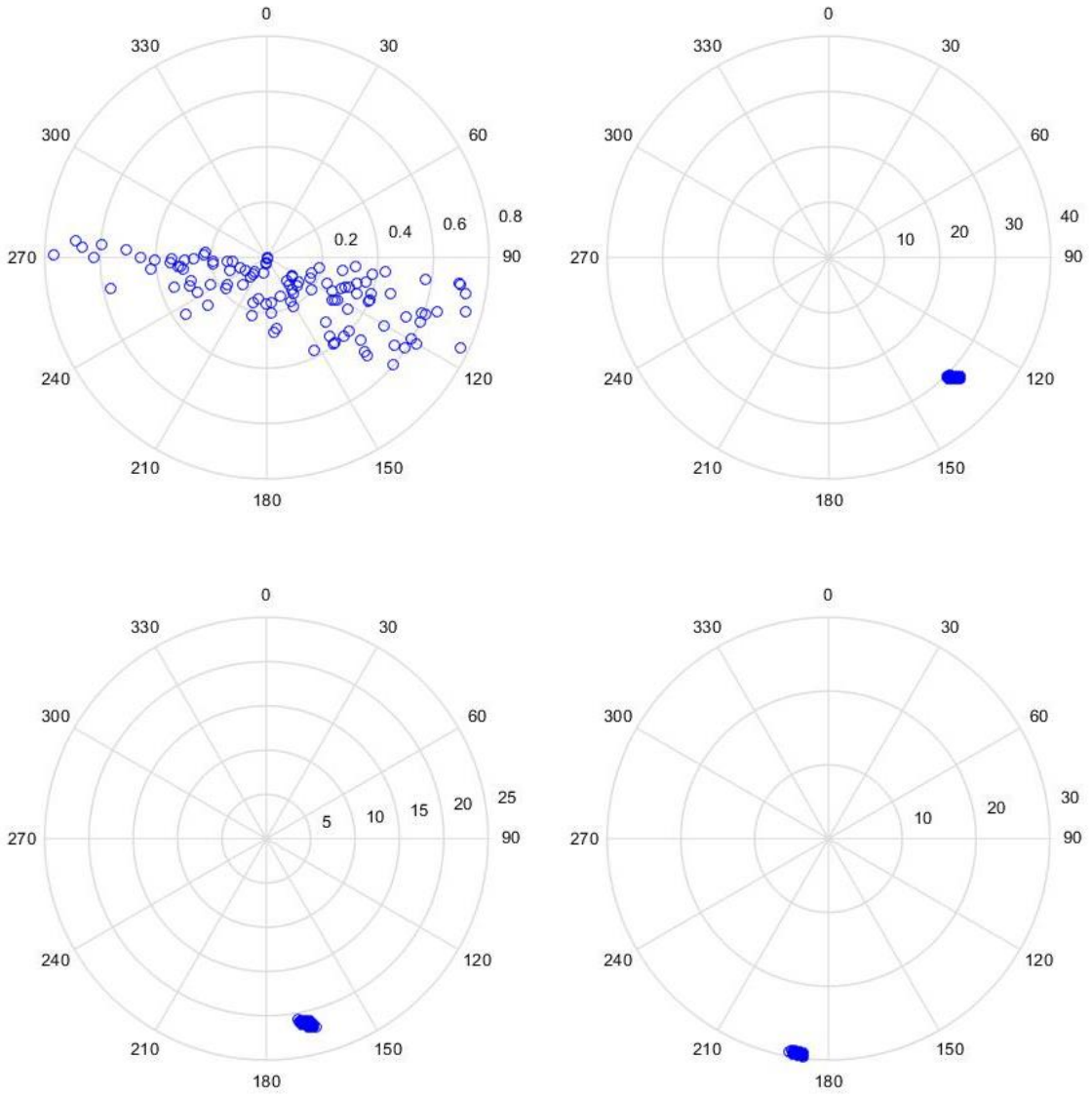


Figure 39 Pseudorotation plots of Ac-4,4-F_2Pro-OMe for fitting one conformation, using ^1H-^{19}F couplings. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

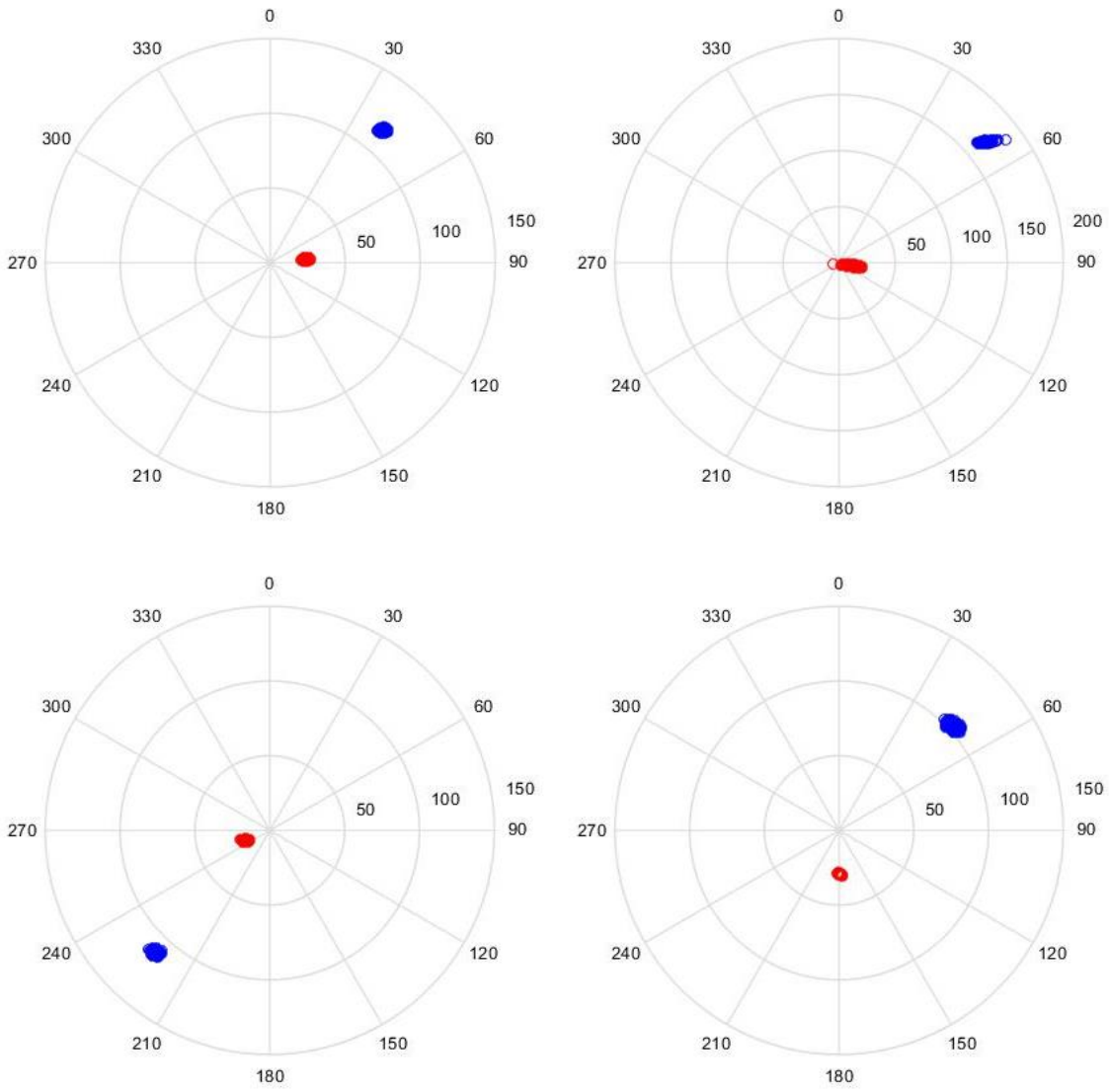


Figure 40 Pseudorotation plots of Ac-4,4-F_2Pro-OMe for fitting two conformations, using ^{1}H-^{19}F couplings.. Left: $CDCl_3$. Right: D_2O. Above: trans *conformer. Below:* cis *conformer.*

When fitting two conformations (Figure **40**), two rather well-defined conformations can always be fitted for Ac-4,4-F_2Pro-OMe. However, the fitted conformations are inconsistent for the four different cases, and are also not close to either a C$^\gamma$ *endo* or C$^\gamma$ *exo* conformation. Furthermore, the v_{max} values of the minor conformers are always unreasonably high.

The observations in literature concerning the ring pucker of Ac-4,4-F_2Pro-OMe (*vide supra*) are based on the assumption that the two opposing stereoelectronic effects of the fluorine atoms will completely cancel each other out. Renner *et al.* [4], drew this conclusion by comparing the patterns of ^{1}H-^{1}H and ^{1}H-^{19}F coupling constants of different proline analogues, without attempting an extensive conformational analysis as presented here. They reasoned that comparable patterns should yield comparable

conformations and from this they concluded that Ac-4,4-F_2Pro-OMe should be predominantly C$^{\gamma}$ *endo*. It is clear that the possibility of other ring puckers than C$^{\gamma}$ *endo* or C$^{\gamma}$ *exo* was never considered. It is thus not impossible that different conformations are significantly populated, and that this is reflected in our conformational analysis. However, 8 out of 10 used couplings were ^{1}H-^{19}F couplings, for which the ^{1}H-^{19}F Karplus relation needs to be used, which is a relation of which the accuracy for the fluoroproline analogues is not fully established, considering also the approximations used for group electronegativities and bond angles. It could be that other parameterizations than the ones used now are needed. Therefore, it can be concluded that further research into refining the Karplus relations will be needed to draw a confident conclusion concerning Ac-4,4-F_2Pro-OMe.

If it is assumed, that only the C$^{\gamma}$ *endo* or C$^{\gamma}$ *exo* conformations are possible (with the values obtained from Ac-(4*S*)-FPro-OMe, P = 205.2°, v_{max} = 36.8°; and Ac-(4*R*)-FPro-OMe, P = 26.8°, v_{max} = 50.0°, respectively), the experimental couplings of the Ac-4,4-F_2Pro-OMe samples are used to fit the fraction based on the fact that the lowest RMSD value gives the best fit. Except for the *trans* conformer in D_2O, where insufficient individual couplings could be measured, such an analysis was performed. Consistently it was found that the C$^{\gamma}$ *exo* conformer is more predominant than the C$^{\gamma}$ *endo* conformer. This contradicts the literature assumption that Ac-4,4-F_2Pro-OMe behaves similarly as Ac-Pro-OMe, so still it can be concluded that further research is required. However, care has to be taken not to attach too much importance to these results, as even the minimal RMSD value is quite high, 2.8 Hz (plots can be found in the supporting information, section **6.7.6**).

3.4.5. Fitting with the Barfield correction

In all previous results, no Barfield corrections were taken into account. The analyses were repeated on those samples for which (nearly) all individual ^{1}H-^{1}H couplings could be measured. For all cases except two, the Barfield correction gave similar to equal results, but with higher or equal RMSD values. For the *cis* conformer of Ac-4,4-F_2Pro-OMe and Ac-(4*R*)-FPro-OMe both in $CDCl_3$, completely different puckers were obtained. However, in those cases the RMSD values were a lot higher than before applying the Barfield correction. It can be concluded that the proposed Barfield correction does not provide an added value for the investigated samples. All results can be found in the supporting information (sections **6.7.4** and **6.7.5**).

4. Thermodynamics and kinetics of the *cis/trans* isomerism

The phenomenon of *cis/trans* isomerism was already described in Chapter **1**, and in Chapter **2** it was explained that this kind of isomerism also takes place in the investigated molecules. In this chapter, both the thermodynamics and kinetics of the *cis/trans* isomerism will be discussed.

4.1. Investigation of the thermodynamics

4.1.1. Theory

The central property in chemical thermodynamics is the equilibrium constant *K*. In case of the investigated equilibrium (Figure **41**) it can be described as follows.

$$K_{trans/cis} = \frac{[trans]}{[cis]}$$

In which [*trans*] and [*cis*] are the equilibrium concentrations of the *trans* and *cis* conformer respectively. The equilibrium constant can be related to the standard Gibbs free energy of the reaction $\Delta G°_{trans/cis}$ by the following formula.

$$K_{trans/cis} = e^{-\frac{\Delta G^0_{trans/cis}}{RT}}$$

In order to retrieve the reaction enthalpy $\Delta H°_{trans/cis}$ and the reaction entropy $\Delta S°_{trans/cis}$, this equation has to be rearranged.

$$\Delta H^0_{trans/cis} - T\Delta S^0_{trans/cis} = -RT\ln(K_{trans/cis})$$

If $-RT\ln(K)$ is plotted versus temperature, the intercept will be equal to the reaction enthalpy and the negative of the slope will be equal to the reaction entropy (see supporting information section **6.8.2**).

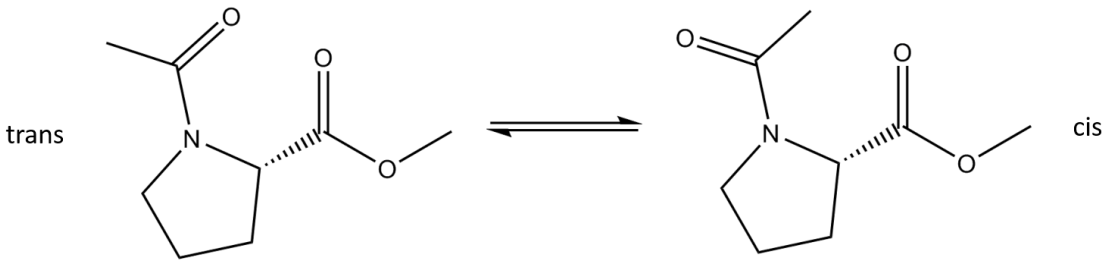


Figure 41 The cis/trans isomerism for Ac-Pro-OMe.

$K_{trans/cis}$ can be easily measured from quantitative 1D spectra, since the intensities of the peaks should reflect the concentrations of both forms. So for a certain resonance *i*, $K_{trans/cis}$ can also be written as follows.

$$K_{trans/cis} = \frac{I_i^{trans}}{I_i^{cis}}$$

Resonances fit to provide intensities by integration, should be resolved from other resonances. Typically, either one of the methyl resonance pairs in the 1D ^{1}H spectrum, or the ^{19}F resonances in case of the monofluorinated compounds were used, except for Ac-(4*R*)-FPro-OMe, where a resolved β ^{1}H multiplet had to be used. Spectra were measured at six different temperatures in the case of $CDCl_3$, and three in D_2O. The spectra of the $CDCl_3$ samples were measured non-quantitatively, however, this is justifiable, as T_1 relaxation for a certain resonance will be approximately the same for the *cis* as for the *trans* conformer (see supporting information, section **6.8.1**).

4.1.2. Results and discussion

The results for $K_{trans/cis}$ are presented in Tables **2** and **3**. Three trends can be seen. Firstly, $K_{trans/cis}$ drops with increasing temperature. According to Le Châtelier's principle, when temperature is increased and the equilibrium shifts in the opposite direction, it implies that the reaction is exothermic. This confirms that the *trans* conformer is more energetically stable than the *cis* conformer, as already discussed in Chapter **1**. Secondly, Ac-(4*R*)-FPro-OMe has the highest values for $K_{trans/cis}$ while Ac-(4*S*)-FPro-OMe has the lowest values, while the values for Ac-Pro-OMe and Ac-4,4-F_2Pro-OMe fall in-between. It thus appears that the *cis/trans* equilibrium follows a similar trend as the ring pucker. As discussed in Chapter **1**, it is known that a stabilizing $n \rightarrow \pi^*$ interaction is present in the *trans* form when a C$^{\gamma}$ *exo* pucker is adopted, providing an explanation for this correlation. As shown in Chapter **3**, Ac-(4*R*)-FPro-OMe has a large preference for the C$^{\gamma}$ *exo* pucker, which shifts the equilibrium more towards the *trans* conformer. Ac-(4S)-FPro-OMe, on the other hand, has a large preference for the C$^{\gamma}$ *endo* pucker, which will shift the *cis/trans* equilibrium in the opposite direction. However, $K_{trans/cis}$ is still larger than 1, so the trans conformer is still more prevalent. The puckering equilibrium of Ac-4,4-F_2Pro-OMe is not really biased towards one form, which explains why its $K_{trans/cis}$ values are closer to those of Ac-Pro-OMe. Lastly, the $K_{trans/cis}$ values in $CDCl_3$ are always lower than those in D_2O. A

possible explanation for this was provided by Siebler *et al.* [46]. They theorized that the increased charge separation induced by the $n \rightarrow \pi^*$ interaction between the adjacent carbonyl groups is better solvated and therefore more stabilized in D_2O. Since these interactions only occur in the *trans* form, this can be a reason why the population of the *trans* form is higher in D_2O than in $CDCl_3$.

Table 2 $K_{trans/cis}$ values for the $CDCl_3$ samples at different temperatures.

T(K)	Ac-Pro-OMe	Ac-(4R)-FPro-OMe	Ac-(4S)-FPro-OMe	Ac-4,4-F_2Pro-OMe
283	4.36	4.12	1.80	3.29
288	4.22	4.48	1.80	3.18
293	4.03	4.32	1.74	3.10
298	3.89	4.20	1.70	3.03
303	3.76	4.00	1.68	2.95
308	3.67	3.89	1.65	2.92

Table 3 $K_{trans/cis}$ values for the D_2O samples at different temperatures.

T(K)	Ac-Pro-OMe	Ac-(4R)-FPro-OMe	Ac-(4S)-FPro-OMe	Ac-4,4-F_2Pro-OMe
298	5.04	7.23	2.50	3.51
308	4.79	6.44	2.41	3.33
333	4.03	5.19	2.18	3.02

As described above, $\Delta H°_{trans/cis}$ and $\Delta S°_{trans/cis}$ values can be obtained from these data by linear regression (Table **4**). The negative values for $\Delta H°_{trans/cis}$ point out that the reaction is exothermic, and the absolute values more or less show the same trends between different molecules and different solvents as could be seen for $K_{trans/cis}$. The values for Ac-4,4-F_2Pro-OMe in D_2O are not very reliable, since the regression is not good ($R^2 = 0.89$). The values for $\Delta S°_{trans/cis}$ are all negative, which indicates that the *trans* conformer is more ordered than the *cis* conformer. A possible rationalization can be found when considering the $n \rightarrow \pi^*$ interactions (see Chapter **1**). These interactions can only occur in the *trans* conformer, and will limit the atomic motion, making the molecules more rigid and ordered.

Table 4 $\Delta H°_{trans/cis}$ (in kJ/mol) and $\Delta S°_{trans/cis}$ (in J/mol K) values for the investigated molecules in both solvents.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	$CDCl_3$	D_2O	$CDCl_3$	D_2O	$CDCl_3$	D_2O	$CDCl_3$	D_2O
$\Delta H°_{trans/cis}$	-5.1 ± 0.4	-5.4 ± 2.9	-5.3 ± 0.6	-7.7 ± 2.0	-2.5 ± 0.2	-3.2 ± 0.3	-3.8 ± 0.3	-3.5 ± 0.9
$\Delta S°_{trans/cis}$	-5.9 ± 1.3	-4.5 ± 9.2	-5.9 ± 2.1	-9.6 ± 6.5	-4.0 ± 0.8	-3.2 ± 1.0	-3.5 ± 1.0	-1.3 ± 2.9

4.2. Investigation of the kinetics

4.2.1. Theory

Consider a system whereby a spin is in equilibrium between two identities *I* and *S* with an equilibrium constant *K*, a forward reaction rate constant k_1 and a backward reaction rate constant k_{-1}.

$$I \leftrightharpoons S$$

When an equal relaxation rate constant R_1 is assumed in both identities – which is justifiable, see the results of the inversion recovery experiments in the supporting information, section **6.8.1** – the deviation of the magnetizations of the spins will evolve according to the following equations.

$$\begin{cases} \dfrac{d\Delta I(t)}{dt} = -(R_1 + k_1)\Delta I(t) + k_{-1}\Delta S(t) \\ \dfrac{d\Delta S(t)}{dt} = -(R_1 + k_{-1})\Delta S(t) + k_1\Delta I(t) \end{cases}$$

This can also be written in matrix representation:

$$\frac{d}{dt}\begin{pmatrix}\Delta I(t)\\ \Delta S(t)\end{pmatrix} = \begin{pmatrix}-(R_1 + k_1) & k_{-1}\\ k_1 & -(R_1 + k_{-1})\end{pmatrix}\begin{pmatrix}\Delta I(t)\\ \Delta S(t)\end{pmatrix}$$

Which abbreviated is:

$$\frac{d}{dt}\Delta\mathbf{M}(t) = [\mathbf{R} + \mathbf{K}]\Delta\mathbf{M}(t) \tag{4.1}$$

Since [**R** + **K**] is independent of time, the general solution to this differential equation is as follows.

$$\Delta\mathbf{M}(t) = e^{[\mathbf{R}+\mathbf{K}]t}\Delta\mathbf{M}(0)$$

Applying the matrix exponential to [**R** + **K**]*t* yields:

$$e^{[\mathbf{R}+\mathbf{K}]t} = \begin{pmatrix} \dfrac{k_1}{k_1+k_{-1}}e^{-(R_1+k_1+k_{-1})t} + \dfrac{k_{-1}}{k_1+k_{-1}}e^{-R_1 t} & -\dfrac{k_{-1}}{k_1+k_{-1}}e^{-(R_1+k_1+k_{-1})t} + \dfrac{k_{-1}}{k_1+k_{-1}}e^{-R_1 t} \\ -\dfrac{k_1}{k_1+k_{-1}}e^{-(R_1+k_1+k_{-1})t} + \dfrac{k_1}{k_1+k_{-1}}e^{-R_1 t} & \dfrac{k_{-1}}{k_1+k_{-1}}e^{-(R_1+k_1+k_{-1})t} + \dfrac{k_1}{k_1+k_{-1}}e^{-R_1 t} \end{pmatrix}$$

The following relation exists between the equilibrium constant, the equilibrium populations *p* of *I* and *S*, and the kinetic constants:

$$K = \frac{p_S}{p_I} = \frac{k_1}{k_{-1}}$$

Thus, it can be written that:

$$p_S = p_I \frac{k_1}{k_{-1}} \text{ or } p_I k_1 = p_S k_{-1}$$

And from the definition of $k_{ex} = k_1 + k_{-1}$, and with $p_I + p_S$ normalized to 1, it follows that:

$$\begin{cases} p_I k_{ex} = p_I k_1 + p_I k_{-1} = p_S k_{-1} + p_I k_{-1} = k_{-1} \\ p_S k_{ex} = p_S k_1 + p_S k_{-1} = p_S k_1 + p_I k_1 = k_1 \end{cases}$$

With these relations the exponential of [**R** + **K**]*t* can be rewritten:

$$e^{[\mathbf{R}+\mathbf{K}]t} = \begin{pmatrix} (p_I + p_S e^{-k_{ex}t}) e^{-R_1 t} & p_I (1 - e^{-k_{ex}t}) e^{-R_1 t} \\ p_S (1 - e^{-k_{ex}t}) e^{-R_1 t} & (p_S + p_I e^{-k_{ex}t}) e^{-R_1 t} \end{pmatrix} \quad (4.2)$$

These equations, (4.1) and (4.2), form the basis to derive the outcome of the selective inversion recovery and the EXSY experiments.

4.2.1.1. Selective inversion recovery equation

This experiment was introduced in Chapter **2**. When doing a selective inversion recovery, where only the resonance from *S* is inverted while that of *I* remains unaffected, the deviations in z-magnetizations immediately after the 180° pulse (at time 0) will be the following.

$$\begin{cases} \Delta I(0) = 0 \\ \Delta S(0) = -2S^0 = -2p_S M^0 = -2p_I M^0 \frac{k_1}{k_{-1}} \end{cases}$$

With M^0 the total equilibrium z-magnetization of spins *I* and *S* (with $I^0 = p_I M^0$ and $S^0 = p_S M^0$). When plugging these values in equation (4.1) (with (4.2) as the exponential of [**R** + **K**]*t*), a description of the evolution of the perturbation of the z-magnetization of spin *I* as a function of time is obtained.

$$\Delta I(t) = -2p_I M^0 \left(p_I e^{-R_1 t} - p_I e^{-(R_1 + k_{ex})t}\right) \frac{k_1}{k_{-1}}$$

This result reveals that, after selective inversion of *S*, the z-magnetisation of *I* will first decease with an initial rate of k_{ex}, then reach a minimum, and finally return to its

equilibrium value due to T_1- relaxation (see Figure **42**). Applying a 90° pulse at time *t* will convert the z-magnetization into detectable signal, the integral of which will abide the above expression. Measurement of spectra at different values of *t*, thus allows extraction of k_{ex} through curve fitting. Since, in practice, the selective inversion is imperfect, and the magnetization evolves before detection, a slightly modified equation is used to perform the data fitting (in which $I(t)$ is used instead of $\Delta I(t)$).

$$I(t) = I^0\left(1 - \left(ae^{-\frac{t}{T_1}} - be^{-t\left(\frac{1}{T_1}+k_{ex}\right)}\right)\right)$$

In which k_{ex}, I^0, *a* and *b* are the parameters that are being fitted, and in which T_1 (= $1/R_1$) is required as input. T_1 constants for all samples were measured with inversion recovery experiments (see Chapter **2**) and fitted values can be found in the supporting information (section **6.8.1**).

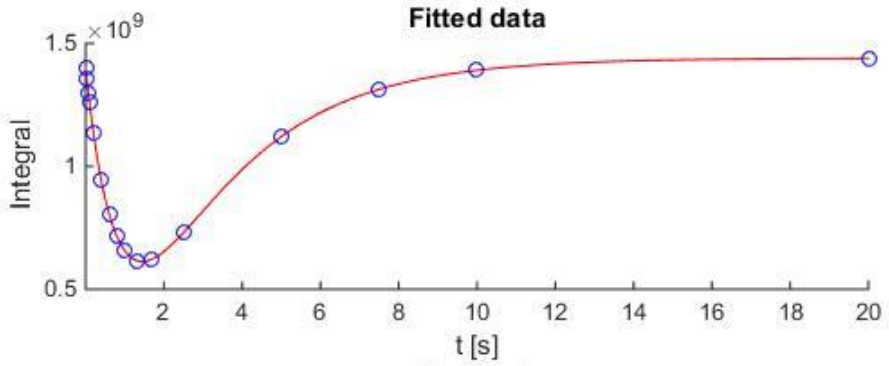


Figure 42 Selective inversion recovery of Ac-(4R)-FPro-OMe in $CDCl_3$ at 308 K.

4.2.1.2. EXSY equations

The Δ**M**(0) matrix, at the start of the mixing time and after the t_1-evolution, is given by:

$$\Delta\mathbf{M}(0) = \begin{pmatrix} -p_I M^0 \cos\omega_I t_1 \\ -p_S M^0 \cos\omega_S t_1 \end{pmatrix} \quad (4.3)$$

With ω the Larmor frequency. This means that for spin *I*, the peak intensity will evolve as (plugging in (4.2) and (4.3) in (4.1)):

$$-M^0 \cos\omega_I t_1 \left[p_I(p_I + p_S e^{-k_{ex}t})e^{-R_1 t}\right] - M^0 \cos\omega_S t_1 \left[p_I p_S(1 - e^{-k_{ex}t})e^{-R_1 t}\right]$$

The cos $\omega_I t_1$ part represents the diagonal peak and the cos $\omega_S t_1$ part represents the cross-peak. A similar expression can be found for spin *S*:

$$-M^0 \cos \omega_S t_1 \left[p_S(p_S + p_I e^{-k_{ex}t})e^{-R_1 t}\right] - M^0 \cos \omega_I t_1 \left[p_I p_S(1 - e^{-k_{ex}t})e^{-R_1 t}\right]$$

In this case, the cos $\omega_S t_1$ part represents the diagonal peak and the cos $\omega_I t_1$ part represents the cross-peak.

To find k_{ex}, the ratio of the integrals of the diagonal peak and the cross-peak at the same F_2 frequency, has to be taken:

$$r_I = \frac{I_I^{cross}}{I_I^{diag}} = \frac{p_S(1 - e^{-k_{ex}t})}{p_I + p_S e^{-k_{ex}t}}$$

An analogous expression can be found for spin *S*. With some rearrangement, the following equations for k_{ex} can be obtained:

$$k_{ex} = -\frac{\ln\left(\frac{1 - \frac{p_I}{p_S} r_I}{1 + r_I}\right)}{t_m}$$

$$k_{ex} = -\frac{\ln\left(\frac{1 - \frac{p_S}{p_I} r_S}{1 + r_S}\right)}{t_m}$$

With t_m the EXSY mixing time. Via these formulae, values for k_{ex} could be determined from 2D EXSYs. These were measured at various mixing times to check the consistency of the results, and which also provided an estimation of error for k_{ex}. The signals used to determine k_{ex}, both for the selective inversion recovery and the EXSY, were the ^{19}F signals in case of the monofluorinated samples, or the (resolved) methyl ^{1}H signals for the other samples.

4.2.1.3. Arrhenius and Eyring equations

When the values for k_{ex} are determined for different temperatures, the activation energy of the reaction (E_a) and the pre-exponential factor (A) can be determined by using the Arrhenius equation.

$$k = Ae^{-\frac{E_a}{RT}}$$

The k in this equation is not k_{ex}, but rather k_1 or k_{-1}, so for both the forward and backward reaction, E_a and A can be determined (k_1 and k_{-1} can be derived from k_{ex} if K is known, *vide supra*). In order to obtain E_a and A, the Arrhenius equation is linearized:

$$\ln(k) = \ln(A) - \frac{E_a}{RT}$$

The values of ln(*k*) are plotted as a function of 1/*RT* and linear regression is performed. The slope will be equal to -E_a and the intercept will be equal to ln(*A*). The resulting graphs can be found in the supporting information (section **6.8.3**).

The Eyring equation is similar to the Arrhenius equation and it relates k_1 or k_{-1} to the enthalpy of activation $\Delta H^{\ddagger}$ and the entropy of activation $\Delta S^{\ddagger}$.

$$k = \frac{k_B T}{h} e^{\frac{\Delta S^{\ddagger}}{R}} e^{-\frac{\Delta H^{\ddagger}}{RT}}$$

The linear form of this equation is as follows:

$$R \ln\left(\frac{hk}{k_B T}\right) = -\Delta H^{\ddagger} \frac{1}{T} + \Delta S^{\ddagger}$$

With k_B the Boltzmann constant and *h* Planck's constant. The values of $R \ln\left(\frac{hk}{k_B T}\right)$ are plotted versus 1/*T*. Linear regression will yield -$\Delta H^{\ddagger}$ as the slope and $\Delta S^{\ddagger}$ as the intercept. The resulting graphs can be found in the supporting information (section **6.8.4**).

4.2.2. Results and discussion

In first instance, only the selective inversion recovery was used to determine the k_{ex} values, as it was expected that this strategy would allow a faster sampling of various mixing times. However, this approach had some issues. First, a selective pulse (Iburp) was used for excitation. Such a pulse is quite long, meaning that relaxation and exchange will already take place during this time, before the start of acquisition. When k_{ex} is small compared to the inverse of the pulse duration, this may result in a situation where the sampling of the curve starts at a position that is too late for reliably fitting k_{ex} values. Especially at low temperatures this was the case, or when the signals of *I* and *S* were too close to each other, necessitating a long duration of the selective pulse. An alternative approach was then used for selective inversion, using a $90°_{x}$ — Δ — $90°_{-x}$ sequence, with Δ a delay equal to half of the inverse of the frequency difference between *I* and *S*, and the rf-frequency set to the resonance frequency of spin *I* (see Chapter **2**). This approach, which only works for singlets, is generally shorter than a selective pulse, making it easier to fit the function to the data. However, there were still

cases where k_{ex} turned out to be too small (*i.e.* in D_2O at 308 K, or most Ac-Pro-OMe samples), so the fitting is still problematic.

Finally, we resorted to 2D EXSY to determine k_{ex}. For Ac-Pro-OMe in $CDCl_3$ at 308 K, where the selective inversion recovery experiments worked out, it was validated that the same result was obtained with EXSY. For this reason, all other cases where the selective inversion recovery experiments were already measured and yielded reliable fitting, no EXSY experiments (which takes ca. 4.5 h per sample) were measured additionally. When the selective inversion recovery was used, 95% errors (ε_{95}) were calculated using a Monte Carlo analysis [44], similar to the one described in Chapter **3**. When the EXSY was used, 95% errors were calculated based on the standard deviation (σ) of all cross-peak volumes per temperature.

$$\varepsilon_{95} = \frac{\sigma}{\sqrt{n}} 1.96$$

The number of integrated cross -peaks was typically six (n = 6), since one EXSY spectrum delivers two cross-peaks on either side of the diagonal, and three different mixing times were measured.

The obtained values for k_{ex} for each compound in D_2O and $CDCl_3$, and at different temperatures, can be found in Tables **5** and **6**. Three trends can be seen in the results. Firstly, k_{ex} increases with increasing temperature. This is logical, since increasing the temperature, increases the amplitude of atomic displacements, facilitating the rotation around the amide bond. Secondly, k_{ex} increases with increasing amount of fluorine atoms in the molecule. As already discussed in Chapter **1**, the fluorine atoms withdraw electron density from the amide bond, decreasing its double bond character and thus making rotation around the bond easier. Lastly, a clear solvent effect is visible. Exchange in D_2O is strikingly slower than in $CDCl_3$, and in D_2O a difference in k_{ex} can be observed between the monofluorinated molecules, while they have approximately the same k_{ex} values in $CDCl_3$, so the solvent effect is not uniform. There is a change in charge separation in the amide bond during the rotation. This causes the carbonyl oxygen atom to have a larger partial negative charge in a planar amide (equilibrium state) than in an orthogonal amide (transition state). Therefore, the equilibrium state is more polar than the transition state. This explains why rotation over the amide bond is slower in D_2O, which shows better solvation for polar compounds than $CDCl_3$ [47].

Table 5 k_{ex} values in s^{-1} for the $CDCl_3$ samples at different temperatures. Determined via EXSY: all temperatures for Ac-Pro-OMe and 303 K for the other samples. All others determined via selective inversion recovery.

T(K)	Ac-Pro-OMe	Ac-(4R)-FPro-OMe	Ac-(4S)-FPro-OMe	Ac-4,4-F_2Pro-OMe
298	0.0973 ± 0.0073	0.2479 ± 0.0087	0.2058 ± 0.0205	0.7372 ± 0.0465
303	0.1849 ± 0.0181	0.4239 ± 0.0348	0.4327 ± 0.0105	1.4768 ± 0.0113
308	0.3295 ± 0.0389	0.7309 ± 0.0098	0.7544 ± 0.0125	2.5056 ± 0.0658

Table 6 k_{ex} values in s^{-1} for the D_2O samples at different temperatures. Determined via EXSY: all temperatures for Ac-Pro-OMe and 308 K for the other samples. All others determined via selective inversion recovery.

T(K)	Ac-Pro-OMe	Ac-(4R)-FPro-OMe	Ac-(4S)-FPro-OMe	Ac-4,4-F_2Pro-OMe
308	0.0354 ± 0.0080	0.1069 ± 0.0105	0.0415 ± 0.0063	0.1481 ± 0.0094
333	0.3693 ± 0.0086	0.8526 ± 0.0441	0.5200 ± 0.0269	1.7008 ± 0.0402

The result for Ac-Pro-OMe at 308 K can be compared with the result obtained by Eberhardt *et al*. [48]. They found a value for k_{ex} of 0.024 ± 0.17 s^{-1}. Given the large uncertainty on their result, both results are comparable.

Obtained values for E_a; A; $\Delta H^{\ddagger}$; and $\Delta S^{\ddagger}$ can be found in Table **7**. These values were only obtained for $CDCl_3$ samples, since in D_2O, only two data points were collected due to time restrictions, which is insufficient to perform linear regression. The first observation that can be made is that the activation energy and activation enthalpy for the backward reaction (from *trans* to *cis*) are always higher than for the forward reaction. This is logical since the *trans* conformer is more stable than the *cis* conformer, so the activation energy or activation enthalpy for the backward reaction will be the sum of the activation energy of the forward reaction and the energy difference between the *cis* and the *trans* conformer. Thus, according to the *K* values determined in the previous section, the difference in E_a between the forward and the backward reaction should be the smallest for Ac-(4*S*)-FPro-OMe and the largest for Ac-(4*R*)-FPro-OMe with the other two molecules in between. The difference is not the largest for Ac-(4*R*)-FPro-OMe, but keep in mind that these values are only determined by a three point linear regression, so the error is still quite large (see Table **7**). It can be noticed that $\Delta H^{\ddagger}$ is always about 2.5 kJ/mol less than E_a, a trend that is also visible in the results of Thomas *et al.* [49]. $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ were also determined by Renner *et al.* [4], and when comparing the results, it can be noticed that they are not in agreement with each other, and sometimes even opposing. However, it needs to be noticed that the samples studied by Renner *et al.* were dissolved in D_2O, while our studied samples were dissolved in $CDCl_3$, and solvent can surely have an effect on the difference in stability of the *cis/trans* forms (*vide supra*). It also needs to be emphasized that the results obtained here are only preliminary and have high uncertainties due to the limited number of data points. However, it is striking that the values determined here are

following a trend that correlates with the puckering equilibrium, as Ac-(4*R*)-FPro-OMe has the highest activation enthalpy and the lowest activation entropy, while Ac-(4*S*)-FPro-OMe has the lowest activation enthalpy and the highest activation entropy, and both the activation enthalpy and activation entropy of Ac-Pro-OMe and Ac-4,4-F_2Pro-OMe have intermediate values. It could be that $n \rightarrow \pi^*$ interactions also play a role in the transition state, and thus stabilize and order it, just as was proposed before when discussing the thermodynamic data (*vide supra*). But still, one has to be cautious in interpreting the results, and more data points need to be measured to reduce the uncertainties of these results.

Table 7 E_a values in kJ/mol; A values in 10^{15} s^{-1}; $\Delta H^{\ddagger}$ values in kJ/mol; and $\Delta S^{\ddagger}$ values in J/mol K for the $CDCl_3$ samples, for both the forward and backward reaction.

	Ac-Pro-OMe		Ac-(4*R*)-FPro-OMe		Ac-(4*S*)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	k_1	k_{-1}	k_1	k_{-1}	k_1	k_{-1}	k_1	k_{-1}
E_a	92.5 ± 11.6	95.5 ± 14.8	82.0 ± 0.55	84.5 ± 14.4	98.9 ± 41.7	99.7 ± 53.5	92.6 ± 41.6	96.0 ± 37.9
A	1.26 ± 5.83	1.11 ± 6.49	0.05 ± 0.10	0.03 ± 0.18	28.6 ± 473	23.7 ± 504	9.51 ± 157	12.5 ± 188
$\Delta H^{\ddagger}$	89.9 ± 11.7	93.0 ± 14.8	79.5 ± 0.54	82.0 ± 14.3	96.4 ± 41.7	97.2 ± 53.5	90.0 ± 41.6	93.5 ± 37.9
$\Delta S^{\ddagger}$	35.7 ± 38.7	34.7 ± 49.0	8.41 ± 18.0	4.96 ± 47.3	61.7 ± 138	60.1 ± 177	52.5 ± 137	54.8 ± 125

5. General conclusion

5.1. Concluding remarks

The goal of this thesis was to evaluate the potential of the recently developed NMR methodology (PSYCHEDELIC [31]) for measuring ^{1}H-^{1}H and ^{1}H-^{19}F couplings in available fluorinated proline variants, and to use these couplings in combination with generalized Karplus equations [39, 42] and the Altona-Sundaralingam formalism [35] to reinvestigate their ring pucker. Although the pucker of the studied compounds has previously been established [1], an extensive analysis of all $^{3}J_{HH}$ and $^{3}J_{HF}$ couplings using the aforementioned conformational analysis strategy has not yet been performed. Also the thermodynamics and kinetics of the *cis/trans* equilibrium of the amide bond of these (fluoro)proline analogues were investigated. The explored methodology, and identification of potential pitfalls, will also serve to swiftly engage into the study of novel fluorinated proline compounds whose properties are fully unknown, such as the 3,4-difluorinated prolines.

Generally, PSYCHEDELIC provides a very straightforward way for obtaining couplings, as the 2D J-resolved spectrum that is generated is easy to interpret. However, one major problem still plagues unhindered interpretation of PSYCHEDELIC spectra and pure shift spectra in general, and that is when protons are strongly coupled. How to accurately extract individual couplings from proton signals that are involved in strong coupling, still remains a challenge for these types of experiments. This caused that not all individual couplings could be measured for the fluoroprolines. It is clear that future developments in this research area will be about resolving strong coupling artefacts.

Puckering analysis of Ac-Pro-OMe, Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe, all confirmed the results from puckering analyses conducted in the past. Ac-Pro-OMe is predominantly C^{γ} *endo*, but with a considerable fraction of C^{γ} *exo* (ca. 15% for the *cis* form, and 30% for the *trans* form, which is logical since *trans* prefers the *exo* pucker). Ac-(4*R*)-FPro-OMe and Ac-(4*S*)-FPro-OMe both show a strong preference for a single pucker, being C^{γ} *exo* and C^{γ} *endo* respectively. The presence of a potential second, minimally populated pucker could not be detected using the current state of the methodology. The puckering analysis of Ac-4,4-F_2Pro-OMe confirmed the assumption made in literature [4] that there will be no strong preference for one pucker but rather

an equilibrium between multiple puckers. However, the assumption that Ac-4,4-F_2Pro-OMe will have a similar puckering behaviour as Ac-Pro-OMe could not be confirmed using the available ^{1}H-^{19}F Karplus curves. Although, based on comparison of the only two vicinal ^{1}H-^{1}H couplings alone, it is indeed not unreasonable to assume this. Thus, it is clear that there are still some challenges to be overcome, especially when the use of ^{1}H-^{19}F couplings is necessary.

The pucker analysis was supplemented with a thermodynamic and kinetic study of the *cis/trans* equilibrium involving the amide bond. This delivered results that are in agreement with observations in literature, and that could be rationalized. The *trans* conformer is always more prevalent than the *cis* [1], and a correlation between the puckering preference and the *cis/trans* equilibrium was demonstrated [5]. The *trans* conformer was found to be more predominant in D_2O than in $CDCl_3$ [46]. The exchange rate between the *cis* and *trans* forms increases with increasing number of fluorine atoms [1], and exchange always occurs faster in $CDCl_3$ than in D_2O [47].

5.2. Future prospects

In terms of puckering analysis, a proper validation of the used ^{1}H-^{19}F Karplus relation in the context of fluoroprolines is in order. Another issue is the lack of correct group electronegativities and bond angles, which further affect the accuracy of the used Karplus relations. Extensive computational studies of the investigated molecules can possibly offer a combined solution for these problems.

In terms of kinetic studies, future prospects include extending the temperature dependent data to increase the precision of the obtained Arrhenius or Eyring parameters, and to confirm or disprove the rationalizations that were made concerning the observed trends between the different molecules. Again, computational studies may be able to point out which interactions are determining for these parameters.

In the future, it will be the intention to use these methodologies to perform puckering analyses and thermodynamic and kinetic studies on the less well-studied or fully unknown fluoroproline analogues, such as the 3-fluoroprolines and the 3,4-difluoroprolines, and the NMe_2 derivatives instead of the OMe derivatives, as these more closely resemble proline residues incorporated in a peptide sequence [46]. Then, these results can be used to create a library of possible fluoroproline analogues and their respective conformational preferences. Eventually, the purpose is to incorporate

fluoroproline in proline rich regions of intrinsically disordered proteins to modulate their properties or to facilitate their study using NMR. As already indicated in Chapter **1**, the potential of ^{19}F NMR as a site specific structural analysis technique is anticipated to be considerable, because of the absence of amide protons in PRR motifs. Gaining knowledge of dynamics and structure of PRRs can possibly lead to a better understanding of the protein-protein interactions that lead to molecular recognition, signal transduction and gene regulation.

6. Supporting information

6.1. NMR experiments

6.1.1. The basics of NMR

NMR is a structural analysis technique based on nuclear magnetism, which results from the fact that a nucleus has an intrinsic magnetic moment, related to its nuclear spin. When a spin ½ nucleus is placed in an external magnetic field (B_0), the magnetic moment interacts with the magnetic field (Zeeman effect), leading to a splitting of energy levels according to the spin state (labelled α and β). The nuclear spins will also precess around B_0 with an angular frequency $\omega_0 = -\gamma B_0$ (called the Larmor frequency), with γ the gyromagnetic ratio of the nucleus. The ensemble of spins in an NMR sample leads to a net macroscopic nuclear magnetization, aligned and proportional with the external magnetic field. By means of radio frequency (RF) pulses, this magnetization can be tilted orthogonal to the magnetic field, where it will precess with a frequency v_0 ($= \omega_0/2\pi$). This, in turn, induces an oscillating current with the same frequency v_0 in a detection solenoid. Fourier transform of this oscillating signal from the time domain to the frequency domain yields an NMR spectrum. Since the value of γ differs for each isotope, the measured frequencies will be nucleus dependent, hence the different types of NMR (^{1}H NMR, ^{13}C NMR, ^{19}F NMR,…). The exact magnetic field experienced by a nucleus also depends on the chemical environment of the nucleus, leading to a distinctive frequency for each nucleus in a molecule. Because this frequency is proportional to the magnitude of the external B_0-field, the δ-value, a B_0 independent measure for the chemical shift is usually employed, reported in parts per million (ppm) rather than Hertz (Hz), against a reference [50].

6.1.2. 2D ^{1}H-^{1}H COSY

In a COSY (COrrelation SpectroscopY) spectrum, which is a homonuclear 2D spectrum (both dimensions represent the same type of nucleus, ^{1}H in our case), cross-peaks can be seen between two ^{1}H's between which there exists a scalar coupling [51]. In most cases, a cross-peak indicates that the two ^{1}H's are positioned geminal or vicinal to each other. However, when longer-range couplings are present, cross-peaks can be observed between ^{1}H's separated by more than three bonds.

6.1.3. 2D ^{1}H-^{13}C HSQC

An HSQC (Heteronuclear Single Quantum Correlation) spectrum allows one to see correlations between ^{1}H and ^{13}C resonances that are coupled through a one-bond $^{1}J_{CH}$ coupling. This can facilitate the assignment of crowded parts of the ^{1}H spectra, where there is lot of overlap of peaks, as such peaks may be resolved along the ^{13}C dimension. This type of spectrum also offers a straightforward way to detect geminal ^{1}H signals, since these should correlate with the same ^{13}C chemical shift. In this research, multiplicity edited HSQC (m.e. HSQC) spectra were used, which means that the sign of the cross-peak encodes the ^{13}C multiplicity. Positive signs indicate CH- or CH_3-groups, while negative signs correspond to CH_2-groups [51].

6.1.4. 2D ^{1}H-^{13}C HMBC

In an HMBC (heteronuclear multiple bond correlation) spectrum, one retrieves correlations between coupled ^{1}H and ^{13}C spins that are separated by more than one bond. In practice, ^{1}H and ^{13}C spins separated by two or three bonds possess a coupling strong enough to lead to a cross-peak, although longer-range couplings may sometimes be present. The HMBC spectrum is particularly useful to assign quaternary carbons, as these cannot be assigned from an HSQC.

6.1.5. 2D ^{1}H-^{1}H NOESY

In a NOESY (Nuclear Overhauser Enhancement SpectroscopY) spectrum, the correlations that are visible, result from nuclear magnetic dipole-dipole interactions between the spins. These interactions thus occur through-space, in contrast with the previously described 2D experiments, where the correlations resulted from through-bond scalar couplings. Under the right conditions, the intensity of an nOe cross-peak is inversely proportional to the sixth power of the distance between the ^{1}H's involved. This implies that only ^{1}H's that are spatially close, will exhibit a detectable nOe cross-peak. The strength of an nOe cross-peak depends on the NOESY mixing time, which is the time during which the nOe has to build up [51].

6.1.6. 2D ^{1}H-^{19}F HOESY

A HOESY (Heteronuclear Overhauser Enhancement SpectroscopY) spectrum is essentially a heteronuclear version of a NOESY spectrum. In this project, ^{1}H-^{19}F HOESY spectra were used in which the direct signal acquisition took place on ^{1}H, in

other words, the ^{1}H dimension was the direct dimension (F_2), while the ^{19}F dimension was the indirect dimension (F_1).

6.1.7. Experimental

Samples were prepared by dissolving 2.05 mg of Ac-Pro-OMe, 2.27 mg of Ac-(4*R*)-FPro-OMe, 2.27 mg of Ac-(4*S*)-FPro-OMe, and 2.49 mg of Ac-4,4-F_2Pro-OMe into ca. 600 µl of $CDCl_3$ (Eurisotop, 99.8%) or D_2O (Eurisotop, 99.9%).

All NMR measurements were performed on either one of four different spectrometers: a Bruker Avance II spectrometer equipped with either a 5mm ^{1}H-^{13}C-^{31}P TXI-Z probe or a 5mm ^{1}H-^{13}C-^{19}F TXO-Z probe, operating at ^{1}H, ^{13}C and ^{19}F frequencies of 500.13 MHz, 125.76 MHz and 470.59 MHz respectively; a Bruker Avance III spectrometer equipped with a 5mm BBI-Z probe, operating at ^{1}H and ^{13}C frequency of 500.13 MHz and 125.76 MHz respectively; a Bruker Avance II spectrometer equipped with a 5mm ^{1}H-^{13}C-^{31}P TXI-Z probe, operating at ^{1}H and ^{13}C frequency of 700.13 MHz and 176.05 MHz respectively; or a Bruker Avance II spectrometer equipped with a 5mm ^{1}H-BB-^{19}F TBI-Z probe, operating at ^{1}H and ^{19}F frequency of 400.13 MHz and 376.50 MHz respectively. All measurements were performed at 298.0 K unless mentioned otherwise.

For 1D ^{19}F spectra (with or without ^{1}H decoupling), a double spin echo of a few microseconds long and a 16-step phase cycle was applied to suppress the ^{19}F background signal from the probe [52]. 2D spectra measured for spectral assignment included ^{1}H-^{1}H gCOSY, ^{1}H-^{1}H NOESY using zero-quantum suppression [53] during the 500 ms mixing time, ^{1}H-^{13}C multiplicity edited gHSQC using adiabatic ^{13}C pulses, a ^{1}H-^{13}C gHMBC optimized for couplings of 7 Hz, and a ^{1}H-^{19}F HOESY with a mixing time of 500 ms. Spectral windows were optimized to ensure detection of all correlations in every dimension. Typically, total time domain sampled in the F_2 and F_1 dimensions were 2048 and 512 respectively. Prior to Fourier transform, all spectra were multiplied with a square cosine bell window function (square sine bell in the case of the COSY and HMBC) and zero filled until a 2048×2048 real data matrix was obtained in the frequency dimension.

Pure shift spectra were recorded using the PSYCHE [30] method, available from the University of Manchester NMR methodology group website (http://nmr.chemistry.manchester.ac.uk/), and applying an interferogram acquisition

scheme [29]. The length of the time domain data chunks had a duration matched to the dwell time in the t_2 time domain dimension to avoid discontinuities between chunks. The duration of the data chunks was chosen to be between 50 ms and 100 ms. Typically, 32 data chunks or more were acquired. The pure shift data was reconstructed using a macro available from the University of Manchester NMR methodology group website. The RF-field strength of the amplitude-modulated double chirp pulse (PSYCHE element) was typically set to achieve about a 20° flip angle. PSYCHEDELIC [31] spectra were set up similarly as the PSYCHE pure shift spectra. The spectral width of the 2DJ indirect dimension was set to ca. 20 Hz, and fine-tuned to be exactly in a power of two ratio with the direct dimension spectral width in order to facilitate 45° tilting. Resolution enhancing Lorenz-to-Gauss window functions were applied, width line broadening parameters (lb) and Gaussian broadening (gb) parameters in TopSpin were standardly set to –1.5 and 0.3 respectively along both dimensions of the 2DJ spectrum, or as a function of the required resolution to measure couplings. For the ^{1}H selective pulses, 180° RSNOB [54] or RE-BURP [55] pulses were used, with a duration chosen as a function of the spectral region to be excited. For ^{1}H-^{19}F couplings, an in-house pulse sequence was used that work similarly as PSYCHEDELIC, but applies ^{19}F selective pulses.

For ^{19}F T_1 measurements, inversion recovery experiments with power gated ^{1}H decoupling were used, with an acquisition time of 3.5 s, and an interscan delay of 12.5 s. Typically, 12 points were sampled, with recovery delays ranging from 10 µs to 18 s seconds. For the selective inversion recovery experiments, 16 data points were sampled up to a 20 s recovery delay, and a 16 s interscan delay was used. For ^{1}H T_1 measurements, regular inversion recovery experiments with similar delays as for ^{19}F were used. For 2D ^{1}H-^{1}H EXSY experiments, a similar experimental set up and spectrum processing of the 2D NOESY spectra was used. An interscan delay of 15 s was used. The spectral width was reduced in order to limit the number of time domain points needed in t_1 to achieve sufficient resolution in order to resolve the methyl cross-peaks from the diagonal peaks. A similar set up was applied for the ^{19}F-^{19}F EXSY experiments, except that no zero-quantum suppression was applied.

6.2. ^{1}H chemical shift tables

All spectra were recorded at 298 K at a 700 MHz (Ac-Pro-OMe and Ac-(4S)-FPro-OMe in $CDCl_3$) or 500 MHz (rest) spectrometer. The reported chemical shift values are in ppm. Which label refers to which ^{1}H can be seen in the figure below.

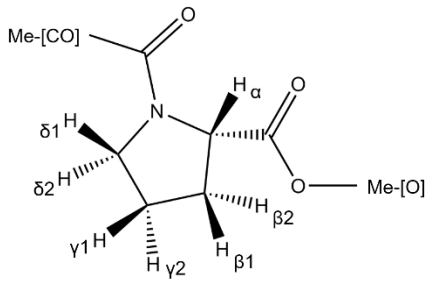


Figure 35 ^{1}H labels.

This type of labelling scheme is more convenient than working with pro R and pro S as it allows a more straightforward way of comparing the results. Suffix 1 indicates that the hydrogen atom is pointing up, while suffix 2 indicates that the hydrogen atom is pointing down. In most of the cases, the pro R hydrogen atom will be the one pointing up, and the pro S hydrogen atom will be the one pointing down. For the β position in Ac-Pro-OMe however, the opposite is true. In case of the difluorinated analogue, no γ chemical shifts can be reported as they are both substituted by fluorine. For the monofluorinated analogues, only one of the γ hydrogen atoms is replaced by fluorine, a chemical shift value for the one that is not substituted will be reported.

6.2.1. $CDCl_3$ samples

Table 8 ^{1}H chemical shift values for all $CDCl_3$ samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
Me-[CO]	2.12	2.01	2.13	2.02	2.15	2.08	2.12	2.03
Me-[O]	3.76	3.80	3.78	3.82	3.77	3.82	3.79	3.84
α	4.52	4.41	4.61	4.60	4.81	4.53	4.74	4.62
β1	2.21	2.31	2.63	2.78	2.37	2.44	2.72	2.84
β2	2.02	2.19	2.15	2.31	2.56	2.75	2.51	2.72
γ1	2.02	1.95	/	/	5.33	5.27	/	/
γ2	2.10	1.95	5.33	5.26	/	/	/	/
δ1	3.53	3.59	3.84	4.27	3.80	3.80	3.91	4.05
δ2	3.68	3.66	3.84	3.56	3.93	3.91	4.00	3.84

6.2.2. D_2O samples

Table 9 ^1H chemical shift values for all D_2O samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
Me-[CO]	2.04	1.93	2.07	1.96	2.08	2.00	2.05	1.97
Me-[O]	3.68	3.73	3.70	3.75	3.71	3.74	3.71	3.76
α	4.37	4.63	4.51	4.84	4.67	4.84	4.69	4.97
β1	2.22	2.28	2.64	2.77	2.47	2.51	2.81	2.90
β2	1.94	2.14	2.15	2.36	2.47	2.64	2.56	2.75
γ1	1.94	1.91	/	/	5.36	5.32	/	/
γ2	1.94	1.78	5.37	5.30	/	/	/	/
δ1	3.57/3.59	3.40	3.95	4.05	3.83	3.67	4.03	3.88
δ2	3.57/3.59	3.47	3.83	3.46	3.92	3.71	4.03	3.77

It was not possible to assess pro R and pro S for the δ hydrogens of the trans conformer of Ac-Pro-OMe.

6.3. ^{13}C chemical shift tables

All spectra were recorded at 298 K at a 700 MHz (Ac-Pro-OMe and Ac-(4S)-FPro-OMe in $CDCl_3$) or 500 MHz (rest) spectrometer. The reported chemical shift values are in ppm. Which label refers to which ^{13}C can be seen in the figure below.

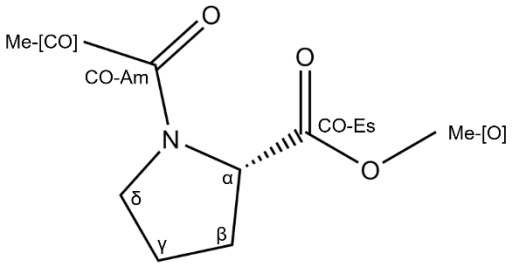


Figure 36 ^{13}C labels.

6.3.1. $CDCl_3$ samples

Table 10 ^{13}C chemical shift values for all $CDCl_3$ samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
Me-[CO]	22.27	22.23	22.25	21.69	22.24	22.08	22.05	21.00
Me-[O]	52.22	52.57	52.41	52.81	52.47	52.81	52.70	53.14
α	58.46	60.17	57.18	58.22	56.92	58.70	56.21	58.03
β	29.41	31.46	36.08	38.24	36.08	38.20	37.53	39.45
γ	24.78	22.80	91.87	90.21	92.15	90.78	126.22	*
δ	47.72	46.28	54.17	52.70	54.07	53.11	54.36	52.90
CO-Am	169.35	169.50	169.33	169.86	169.33	169.77	169.24	169.64
CO-Es	172.83	172.64	172.43	172.28	171.13	171.30	170.68	170.42

*γ ^{13}C chemical shift of the Ac-4,4-F_2Pro-OMe *cis* conformer could not be determined, correlations were not visible in the HMBC.

6.3.2. D_2O samples

Table 11 ^{13}C chemical shift values for all D_2O samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
Me-[CO]	21.16	21.16	21.25	20.74	21.23	21.12	21.12	20.12
Me-[O]	52.87	53.18	53.02	53.38	53.05	53.34	53.28	53.63
α	59.03	60.66	57.65	58.46	57.37	59.09	56.61	58.49
β	29.15	30.56	35.55	37.04	35.49	37.15	36.42	38.04
γ	24.20	22.29	92.75	91.32	93.37	92.24	126.13	126.02
δ	48.38	46.58	54.46	52.75	54.52	53.16	53.83	52.55
CO-Am	172.94	173.28	173.20	173.93	173.36	173.84	173.25	173.81
CO-Es	175.00	174.59	174.23	173.93	173.75	173.96	172.68	172.60

6.4. ^{19}F chemical shift tables

All spectra were recorded at 298 K at a 400 MHz (Ac-4,4-F_2Pro-OMe in D2O) or 500 MHz (rest) spectrometer. The reported chemical shift values are in ppm. Which label refers to which ^{19}F can be seen in the figure below. For Ac-(4R)-FPro-OMe, there is no fluorine atom at the γ2 position, while for Ac-(4S)-FPro-OMe, there is no fluorine atom at the γ1 position. In case of Ac-4,4-F_2Pro-OMe, the fluorine at the γ1 position is pro R, and the fluorine at the γ2 position is pro S.

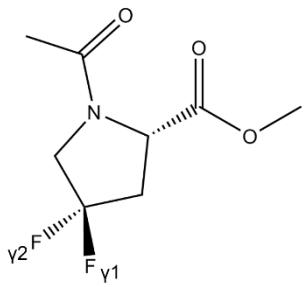


Figure 37 ^{19}F labels.

6.4.1. $CDCl_3$ samples

Table 12 ^{19}F chemical shift values for all $CDCl_3$ samples.

	Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis
γ1	-176.96	-177.87	/	/	-98.19	-101.43
γ2	/	/	-172.85	-173.12	-99.11	-96.74

6.4.2. D_2O samples

Table 13 ^{19}F chemical shift values for all D_2O samples.

	Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis
γ1	-177.85	-177.79	/	/	-102.23	-104.71
γ2	/	/	-173.26	-173.25	-98.35	-95.46

6.5. ^1H-^1H coupling tables

All spectra were recorded at 298 K at a 700 MHz (Ac-(4R)-FPro-OMe in $CDCl_3$) or 500 MHz (rest) spectrometer. The reported coupling constants are in Hz. The same labels as in part 1 are used.

6.5.1. $CDCl_3$ samples

Table 14 Measured ^1H-^1H couplings for all $CDCl_3$ samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
α - β1	8.7	8.6	8.0	8.3	9.9	9.5	9.3	9.4
α - β2	4.0	2.8	9.1	8.1	1.6°	0.8°°	5.4	3.2
β1 - γ1	7.6	9.0†	/	/	4.1	3.8	/	/
β1 - γ2	8.6	9.0†	1.7°	1.7	/	/	/	/
β2 - γ1	*	5.3†	/	/	1.2°	0.9°°	/	/
β2 - γ2	6.5	5.3†	4.3	4.4	/	/	/	/
γ1 - δ1	7.0	8.1†	/	/	4.4	4.3	/	/
γ1 - δ2	4.8	6.2†	/	/	<1	<1	/	/
γ2 - δ1	7.3	8.1†	2.2‡	1.0°	/	/	/	/
γ2 - δ2	7.8	6.2†	2.2‡	3.6	/	/	/	/

*: coupling could not be measured, †: measured coupling is an average of two couplings, ‡: measured coupling is a sum of two couplings, °: coupling was measured with Lorentzian line broadening set to -3, °°: coupling was measured with Lorentzian line broadening set to -5.

6.5.2. D_2O samples

Table 15 Measured ^1H-^1H couplings for all D_2O samples.

	Ac-Pro-OMe		Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis	trans	cis
α - β1	8.8	8.9	7.8	8.7	2.9†	9.7	9.8	9.8
α - β2	4.7	2.6	10.1	8.2	2.9†	0.8°°	4.6	2.1
β1 - γ1	*	7.1	/	/	1.6†	3.4	/	/
β1 - γ2	*	11.1	1.3°	1.5°	/	/	/	/
β2 - γ1	*	3.5	/	/	1.6†	1.1°°	/	/
β2 - γ2	*	6.9	3.8	4.3	/	/	/	/
γ1 - δ1	*	7.6	/	/	3.7	3.8	/	/
γ1 - δ2	*	3.6	/	/	<1	<1	/	/
γ2 - δ1	*	9.2	<1	0.8°°	/	/	/	/
γ2 - δ2	*	8.6	3.2	3.4	/	/	/	/

*: coupling could not be measured, †: measured coupling is an average of two couplings, ‡: measured coupling is a sum of two couplings, °: coupling was measured with Lorentzian line broadening set to -3, °°: coupling was measured with Lorentzian line broadening set to -5.

6.6. ^1H-^{19}F coupling tables

All spectra were recorded at 298 K at a 400 MHz (Ac-4,4-F_2Pro-OMe in D_2O) or 500 MHz (rest) spectrometer. The reported coupling constants are in Hz. The same labels as indicated in the figure below are used.

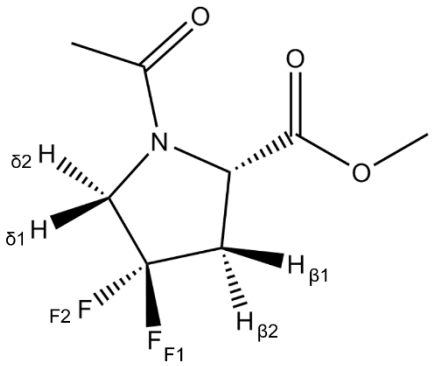


Figure 38 Labels used for reporting ^1H-^{19}F couplings.

6.6.1. $CDCl_3$ samples

Table 16 Measured ^1H-^{19}F couplings for all $CDCl_3$ samples.

	Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis
β1 - F1	19.1	20.1	/	/	13.8	11.8
β1 - F2	/	/	40.5	42.2	14.2	20.0
β2 - F1	38.1	36.2	/	/	13.7	7.4
β2 - F2	/	/	18.8	15.9	13.7	*
δ1 - F1	*	21.3	/	/	11.5	12.2
δ1 - F2	/	/	33.6	37.5°	10.9	15.6
δ2 - F1	*	35.9	/	/	13.3	9.9
δ2 - F2	/	/	25.6	27.2	13.1	13.5

*: coupling could not be measured, †: measured coupling is an average of two couplings, °: coupling was measured in the HSQC.

6.6.2. D_2O samples

Table 17 Measured ^1H-^{19}F couplings for all D_2O samples.

	Ac-(4R)-FPro-OMe		Ac-(4S)-FPro-OMe		Ac-4,4-F_2Pro-OMe	
	trans	cis	trans	cis	trans	cis
β1 - F1	18.5	21.1	/	/	11.8	10.3
β1 - F2	/	/	*	46.1°	19.6	21.5
β2 - F1	41.9	39.2	/	/	10.6	4.2
β2 - F2	/	/	*	16.7	13.6	14.7
δ1 - F1	21.9	21.0	/	/	10.9†	10.5
δ1 - F2	/	/	38.7	*	14.5†	19.8
δ2 - F1	38.2	37.2	/	/	10.9†	5.9
δ2 - F2	/	/	25.0	*	14.5†	16.9

*: coupling could not be measured, †: measured coupling is an average of two couplings, °: coupling was measured in the HSQC.

6.7. Puckering analysis

6.7.1. Advantages and disadvantages of the GUI

The advantages of the GUI are that it is user friendly, and data input is limited and straightforward, certainly for the group electronegativities for which an editor can be used. In the editor, one can graphically assign a group electronegativity value to every possible substituent, and for a certain coupling, the program automatically picks the needed λ's and also automatically assigns to the correct positions in the Diez-Donders equation. The program also allows the use of sums of couplings as input, which is useful when no individual couplings could be measured but average couplings could be obtained.

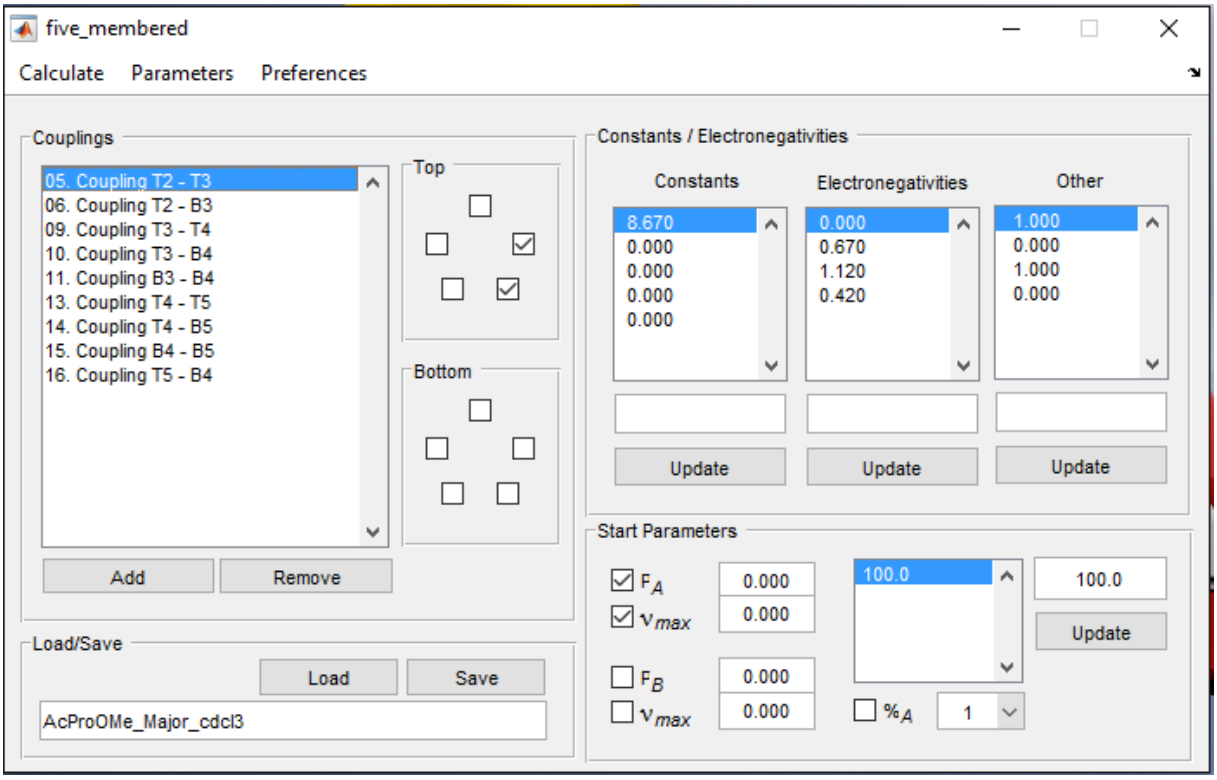


Figure 39 Input window of the graphical user interface.

However, there are some shortcomings. The program works as a black box, with limited information on what happens on the background. It also only works with ^{1}H-^{1}H couplings, which is problematic for Ac-4,4-F_2Pro-OMe, for which only two ^{1}H-^{1}H couplings can be used, which makes fitting of even one conformation impossible. For example, for the monofluorinated molecules, a maximum of six experimental ^{1}H-^{1}H couplings are available, which means two conformations can in principle be fitted, though the results may be unreliable (using six experimental parameters to fit five parameters). It is thus clear that ^{1}H-^{19}F couplings are essential for these compounds. Furthermore, the above-described Barfield correction is not implemented in the GUI. Finally, the generated pseudorotation plots for two conformations turned out not straightforward to interpret, as were the uncertainties of the fitted parameters. In order to overcome these problems, we resorted to writing a new fitting script that better suited our needs.

For one case, a comparison of the generated plots by both the GUI and the MATLAB script will be made (see Figure **6**). In the GUI, a plot of the RMSD of every point in conformational space is made, while in the MATLAB script every fitted conformation during the Monte Carlo analysis is plotted. This yields two different plots, in case of the GUI, a shallow RMSD well indicates a lower goodness-of-fit, while in case of the

MATLAB script, a broad spread of points indicates a higher uncertainty on the fitted value due to measurement errors.

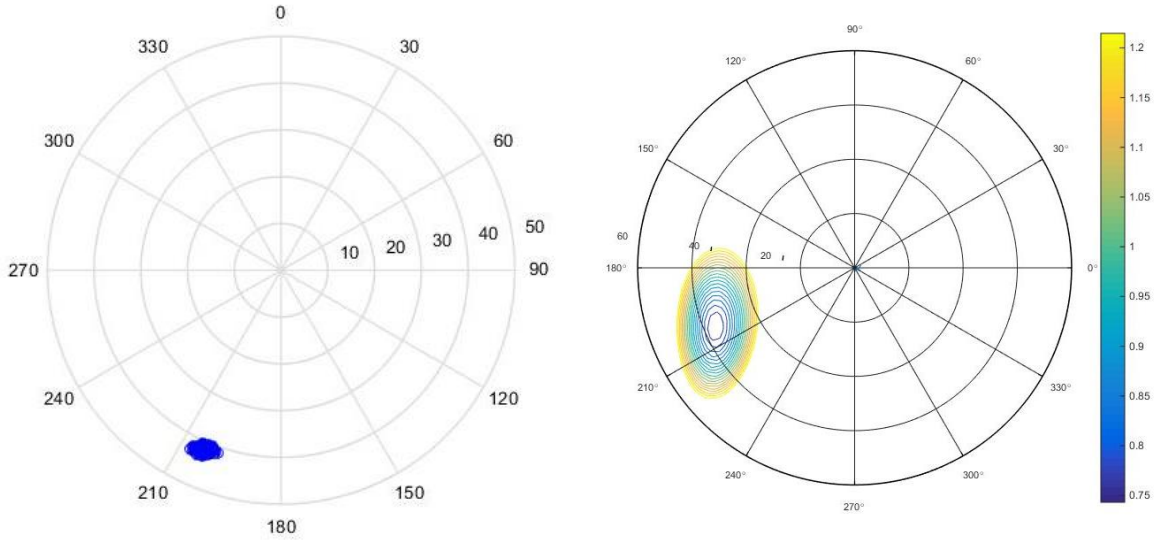


Figure 40 Pseudorotation plots for the cis *conformer of Ac-(4S)-FPro-OMe in D_2O. Left: From the MATLAB script. Right: From the GUI. Note that the polar coordinate is arranged differently.*

6.7.2. Precision

Most ^{1}H-^{1}H couplings could be measured multiple times. One can select ^{1}H A and observe ^{1}H B, but also the other way around. Furthermore, for fluorinated compounds, the β, γ and δ protons are also splitted along F_2 due to ^{1}H-^{19}F coupling, meaning that the same coupling can be measured twice in the same PSYCHEDELIC spectrum. Since all these splittings, though measured on different occasions, should yield the same value, we are able to make an estimate of the precision of the measurement. We found the spread in measured values to be in the order of magnitude of 0.1 Hz.

6.7.3. Input parameters

For the couplings, the same labels are used as in sections **6.4** and **6.5**, only now F is replaced by γ for the ^{1}H-^{19}F couplings. The λ value of H is 0, that of COOR is 0.42, that of NR_2 is 1.12, and that of F is 1.37. The λ values of CHFR and CF_2R were approximated by that of CH_2F (0.65), those of $CH_2CH_2NR_2$ and $CH_2CH(NR_2)COOR$ were approximated by that of CH_2CH_2OR (0.67), that of $CH(NR_2)COOR$ was approximated by that of CH_2COOR (0.68), and that of CH_2NR_2 was approximated by that of CH_2OR (0.68). All used λ values can be found in Table **11**. The used values for B, T_b and P_b (if applicable) can be found in Table **12**. For parameter A, a value of 1 was always used. In Table **13**, the used bond angles in the ^{1}H-^{19}F Karplus relation can

be found. Those for Ac-(4*R*)-FPro-OMe were obtained via the available crystal structure, the values for the other molecules were obtained via computational calculations.

Table 18 All used group electronegativities, both for the Diez-Donders and the H-C-C-F Karplus equation.

	Ac-Pro-OMe				Ac-(4*R*)-FPro-OMe				Ac-(4*S*)-FPro-OMe				Ac-4,4-F_2Pro-OMe			
	λ_1	λ_2	λ_3	λ_4	λ_1	λ_2	λ_3	λ_4	λ_1	λ_2	λ_3	λ_4	λ_1	λ_2	λ_3	λ_4
α-β1	0	0.67	1.12	0.42	0	0.65	1.12	0.42	0	0.65	1.12	0.42	0	0.65	1.12	0.42
α-β2	0.67	0	1.12	0.42	0.65	0	1.12	0.42	0.65	0	1.12	0.42	0.65	0	1.12	0.42
β1-γ1	0	0.68	0.67	0	0.68	0	0	0.68	0.68	0	1.37	0.68	1.37	0.68	0.68	0
β1-γ2	0.68	0	0.67	0	0.68	0	0.68	1.37	0.68	0	0.68	0	0.68	1.37	0.68	0
β2-γ1	0	0.68	0	0.67	0	0.68	0	0.68	0	0.68	1.37	0.68	1.37	0.68	0	0.68
β2-γ2	0.68	0	0	0.67	0	0.68	0.68	1.37	0	0.68	0.68	0	0.68	1.37	0	0.68
γ1-δ1	0.67	0	0	1.12	0.67	0	0	1.12	0.67	1.37	0	1.12	0.67	1.37	0	1.12
γ1-δ2	0.67	0	1.12	0	0.67	0	1.12	0	0.67	1.37	1.12	0	0.67	1.37	1.12	0
γ2-δ1	0	0.67	0	1.12	1.37	0.67	0	1.12	0	0.67	0	1.12	1.37	0.67	0	1.12
γ2-δ2	0	0.67	1.12	0	1.37	0.67	1.12	0	0	0.67	1.12	0	1.37	0.67	1.12	0

Table 19 Values for B (deg), T_b (Hz) and P_b (deg) parameters. T_b and P_b are not applicable (n.a.) for transoid couplings (B ≠ 0).

	B	T_b	P_b
α-β1	0	2	234
α-β2	-120	n.a.	n.a.
β1-γ1	0	1	90
β1-γ2	-120	n.a.	n.a.
β2-γ1	120	n.a.	n.a.
β2-γ2	0	1	270
γ1-δ1	0	2	306
γ1-δ2	-120	n.a.	n.a.
γ2-δ1	120	n.a.	n.a.
γ2-δ2	0	2	126

Table 20 Used a_{FCC} and a_{HCC} angles (in degrees) as input for the ^{1}H-^{19}F Karplus relation. For Ac-(4S)-FPro-OMe and Ac-4,4-F_2Pro-OMe, these were calculated for trans *and* cis *separately.*

	Ac-(4*R*)-FPro-OMe		*trans* Ac-(4*S*)-FPro-OMe		*cis* Ac-(4*S*)-FPro-OMe		*trans* Ac-4,4-F_2Pro-OMe		*cis* Ac-4,4-F_2Pro-OMe	
	a_{FCC}	a_{HCC}	a_{FCC}	a_{HCC}	a_{FCC}	a_{HCC}	a_{FCC}	a_{HCC}	a_{FCC}	a_{HCC}
β1 - F1	108.03	110.2	/	/	/	/	112.52	108.52	112.42	108.38
β1 - F2	/	/	108.04	109.68	108.47	109.7	109.27	108.52	109.89	108.38
β2 - F1	108.03	110.3	/	/	/	/	112.52	110.92	112.42	110.94
β2 - F2	/	/	108.04	111.07	108.47	110.89	109.27	110.92	109.89	110.94
δ1 - F1	107.72	110.5	/	/	/	/	112.37	110.4	111.83	109.57
δ1 - F2	/	/	109.59	111.95	109.57	111.19	110.42	110.4	110.39	109.57
δ2 - F1	107.72	109.3	/	/	/	/	112.37	110.77	111.83	109.48
δ2 - F2	/	/	109.59	110.45	109.57	109.13	110.42	110.77	110.39	109.48

6.7.4. Output parameters

All fitted parameters and RMSD values from the puckering analysis can be found in Tables **14**, **15**, **16** and **17** (one molecule per table). The residuals (the difference between the calculated and experimental couplings) can be found in Tables **18**, **19**, **20** and **21** (one molecule per table). The residuals are defined as follows:

$$J^{expt} - J^{calc}$$

Table 21 Puckering fit output for Ac-Pro-OMe.

		$CDCl_3$		D_2O	
		trans	cis	trans	cis
1 conformer	P	76.4° ± 64.3°	175.7° ± 7.9°	/	300.7° ± 53.6°
	v_{max}	0.0° ± 0.7°	17.3° ± 1.3°	/	0.0° ± 0.6°
	RMSD	2.56 Hz	0.63 Hz	/	3.18 Hz
2 conformers	P_1	17.1° ± 7.1°	111.5° ± 19.8°	/	13.4° ± 10.2°
	$v_{max,1}$	49.7° ± 17.5°	143.2° ± 99.2°	/	43.7° ± 20.5°
	P_2	187.9° ± 5.9°	185.0° ± 52.0°	/	182.8° ± 3.7°
	$v_{max,2}$	31.5° ± 7.5°	17.9° ± 33.7°	/	36.4° ± 4.4°
	x_1	29.5% ± 6.2%	7.3% ± 50.5%	/	16.9% ± 4.8%
	RMSD	0.54 Hz	0.51 Hz	/	0.81 Hz
2 conformers with Barfield	P_1	19.3° ± 5.0°	/	/	17.7° ± 10.7°
	$v_{max,1}$	47.7° ± 14.0°	/	/	39.0° ± 20.3°
	P_2	183.9° ± 5.0°	/	/	181.2° ± 3.6°
	$v_{max,2}$	31.4° ± 6.3°	/	/	36.5° ± 4.8°
	x_1	29.6% ± 5.8%	/	/	17.6% ± 6.3%
	RMSD	0.73 Hz	/	/	1.00 Hz

Table 22 Puckering fit output for Ac-(4R)-FPro-OMe.

		$CDCl_3$		D_2O	
		trans	cis	trans	cis
1 conformer without ^1H-^{19}F	P	24.9° ± 2.1°	24.1° ± 1.7°	13.5° ± 1.7°	25.0° ± 1.7°
	v_{max}	48.0° ± 2.0°	40.7° ± 1.1°	43.5° ± 0.9°	41.1° ± 1.2°
	RMSD	0.94 Hz	0.37 Hz	0.44 Hz	0.45 Hz
1 conformer with ^1H-^{19}F	P	24.0° ± 1.8°	13.1° ± 0.7°	3.1° ± 0.7°	14.5° ± 0.8°
	v_{max}	45.9° ± 0.8°	38.6° ± 0.2°	44.5° ± 0.3°	40.3° ± 0.2°
	RMSD	1.44 Hz	1.69 Hz	0.68 Hz	0.98 Hz
2 conformers with ^1H-^{19}F	P_1	108.2° ± 39.8°	219.8° ± 3.2°	16.2° ± 63.3°	54.4° ± 2.7°
	$v_{max,1}$	184.7° ± 77.7°	104.0° ± 3.4°	44.4° ± 98.7°	147.2° ± 7.3°
	P_2	23.9° ± 4.1°	8.0° ± 1.2°	59.0° ± 46.4°	22.0° ± 1.2°
	$v_{max,2}$	47.1° ± 8.9°	37.5° ± 0.5°	150.0° ± 92.0°	41.2° ± 0.5°
	x_1	9.2% ± 21.0%	15.3% ± 1.0%	96.0% ± 85.9%	8.0% ± 0.6%
	RMSD	0.83 Hz	0.49 Hz	0.52 Hz	0.41 Hz
1 conformer without ^1H-^{19}F with Barfield	P	24.9° ± 2.2°	205.7° ± 0.5°	13.5° ± 1.8°	25.0° ± 1.7°
	v_{max}	48.0° ± 2.1°	142.5° ± 0.8°	43.5° ± 1.0°	41.1° ± 1.2°
	RMSD	0.94 Hz	3.67 Hz	0.44 Hz	0.45 Hz

Table 23 Puckering fit output for Ac-(4S)-FPro-OMe.

		$CDCl_3$		D_2O	
		trans	cis	trans	cis
1 conformer without ^1H-^{19}F	P	206.5° ± 2.9°	199.3° ± 3.0°	165.2° ± 5.4°	203.2° ± 2.6°
	v_{max}	37.1° ± 1.1°	38.0° ± 1.1°	56.3° ± 6.8°	41.6° ± 1.2°
	RMSD	0.74 Hz	0.68 Hz	1.02 Hz	0.76 Hz
1 conformer with ^1H-^{19}F	P	179.4° ± 1.0°	175.8° ± 0.8°	166.8° ± 9.8°	198.7° ± 2.1°
	v_{max}	40.7° ± 0.3°	45.9° ± 0.3°	56.4° ± 15.7°	50.7° ± 0.7°
	RMSD	1.90 Hz	2.46 Hz	3.56 Hz	1.51 Hz
2 conformers with ^1H-^{19}F	P_1	135.3° ± 4.6°	175.8° ± 64.8°	158.5° ± 31.7°	190.6° ± 49.0°
	$v_{max,1}$	150.0° ± 8.5°	45.9° ± 74.8°	67.6° ± 64.4°	48.3° ± 54.5°
	P_2	180.6° ± 11.9°	175.8° ± 50.5°	166.8° ± 72.5°	198.7° ± 52.7°
	$v_{max,2}$	40.7° ± 73.7°	45.9° ± 59.1°	56.4° ± 25.2°	50.7° ± 67.6°
	x_1	7.5% ± 7.7%	49.6% ± 83.1%	0.0% ± 78.5%	0.0% ± 87.3%
	RMSD	1.81 Hz	2.46 Hz	3.56 Hz	1.51 Hz
1 conformer without ^1H-^{19}F with Barfield	P	187.2° ± 4.4°	185.4° ± 4.6°	/	193.7° ± 4.1°
	vmax	31.7° ± 1.1°	34.3° ± 1.1°	/	38.1° ± 1.0°
	RMSD	1.24 Hz	1.11 Hz	/	1.25 Hz

Table 24 Puckering fit output for Ac-4,4-F2Pro-OMe.

		$CDCl_3$		D_2O	
		trans	cis	trans	cis
1 conformer	P	144.6° ± 98.4°	167.7° ± 1.7°	133.9° ± 1.3°	188.1° ± 1.3°
	v_{max}	0.0° ± 0.4°	21.3° ± 0.4°	31.3° ± 0.8°	29.4° ± 0.2°
	RMSD	5.03 Hz	4.06 Hz	4.63 Hz	2.73 Hz
2 conformers	P_1	40.4° ± 0.9°	223.1° ± 1.3°	50.5° ± 1.1°	47.4° ± 2.3°
	$v_{max,1}$	116.2° ± 1.8°	111.8° ± 2.0°	169.3° ± 5.1°	103.4° ± 2.5°
	P_2	84.3° ± 2.2°	247.5° ± 3.5°	102.8° ± 15.2°	179.1° ± 1.8°
	$v_{max,2}$	23.9° ± 1.9°	17.8° ± 1.9°	13.2° ± 5.8°	29.2° ± 0.6°
	x_1	40.4% ± 0.6%	41.0% ± 0.7%	53.8% ± 1.4%	19.0% ± 0.6%
	RMSD	1.12 Hz	0.76 Hz	0.47 Hz	1.29 Hz
2 conformers with Barfield	P_1	40.4° ± 0.9°	38.1° ± 3.4°	/	49.2° ± 2.4°
	$v_{max,1}$	116.2° ± 1.9°	72.4° ± 10.1°	/	104.7° ± 2.7°
	P_2	84.3° ± 2.3°	162.2° ± 5.7°	/	175.7° ± 1.8°
	$v_{max,2}$	23.9° ± 1.9°	30.9° ± 1.9°	/	29.7° ± 0.7°
	x_1	40.4% ± 0.6%	28.4% ± 1.1%	/	19.3% ± 0.6%
	RMSD	1.12 Hz	1.08 Hz	/	1.36 Hz

Table 25 Residuals of the puckering analysis of Ac-Pro-OMe, only for fitting two conformers, without Barfield correction.

	$CDCl_3$		D_2O
	trans	*cis*	*cis*
α-β1	1.0	0.3	1.4
α-β2	0.3	0.3	0.3
β1-γ1	0.4	0.8†	0.1
β1-γ2	0.7	0.8†	1.4
β2-γ1	*	0.2†	0.7
β2-γ2	-0.7	0.2†	-0.1
γ1-δ1	0.0	0.5†	0.6
γ1-δ2	0.5	0.7†	0.9
γ2-δ1	-0.1	0.5†	0.4
γ2-δ2	0.5	0.7†	0.9

*: coupling was not measured and could therefore not be used. †: residual of a sum of couplings.

Table 26 Residuals of the puckering analysis of Ac-(4R)-FPro-OMe, only for fitting one conformer, without Barfield correction

	without ^{1}H-^{19}F couplings				with ^{1}H-^{19}F couplings			
	$CDCl_3$		D_2O		$CDCl_3$		D_2O	
	trans	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*
α-β1	0.5	0.5	0.9	0.9	0.5	1.0	1.6	1.5
α-β2	0.5	0.1	0.6	0.3	0.6	-0.9	-0.4	-0.8
β1-γ	0.1	0.7	-0.1	0.5	0.3	0.6	-0.3	0.3
β2-γ	0.6	-0.1	0.0	-0.2	0.4	0.0	0.2	0.0
γ-δ1	-1.9†	0.0	0.0	-0.3	-1.9†	0.2	0.1	-0.1
γ-δ2	-1.9†	0.3	-0.1	0.1	-1.9†	-0.2	-0.4	-0.2
β1-F	/	/	/	/	-1.6	-3.0	-0.7	-1.2
β2-F	/	/	/	/	-2.7	-2.9	0.2	-0.5
δ1-F	/	/	/	/	*	-1.6	-0.1	-1.0
δ2-F	/	/	/	/	*	-2.4	-0.9	-2.0

*: coupling was not measured and could therefore not be used. †: residual of a sum of couplings.

Table 27 Residuals of the puckering analysis of Ac-(4S)-FPro-OMe, only for fitting one conformer, without Barfield correction.

	without ^{1}H-^{19}F couplings				with ^{1}H-^{19}F couplings			
	$CDCl_3$		D_2O		$CDCl_3$		D_2O	
	trans	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*
α-β1	1.4	1.3	-0.4†	1.4	2.7	3.0	-0.4†	2.1
α-β2	0.0	-0.4	-0.4†	-0.4	0.9	0.0	-0.4†	0.0
β1-γ	-0.9	-0.8	-2.0†	-0.9	0.1	0.4	-2.0†	0.2
β2-γ	0.4	0.0	-2.0†	0.0	-0.1	-0.9	-2.0†	-0.9
γ-δ1	0.7	0.6	-0.1	0.6	0.2	0.4	0.0	1.4
γ-δ2	0.1	0.2	0.2	-0.1	0.3	0.2	0.1	-0.7
β1-F	/	/	/	/	-0.9	-0.2	*	3.2
β2-F	/	/	/	/	-1.6	-1.5	*	0.5
δ1-F	/	/	/	/	3.3	5.1	6.9	*
δ2-F	/	/	/	/	3.6	4.6	4.9	*

*: coupling was not measured and could therefore not be used. †: residual of a sum of couplings.

Table 28 Residuals of the puckering analysis of Ac-4,4-F_2Pro-OMe, only for fitting two conformers, without Barfield correction.

	$CDCl_3$		D_2O	
	trans	*cis*	*trans*	*cis*
α-β1	1.2	0.8	1.1	1.7
α-β2	0.3	-0.2	0.2	0.0
β1-F1	0.7	-0.6	0.3	-2.0
β1-F2	0.6	-0.5	0.3	-1.3
β2-F1	0.3	-0.3	0.0	-1.9
β2-F2	0.0	*	-0.3	-0.1
F1-δ1	1.1	0.9	-0.2†	0.6
F1-δ2	2.2	1.1	-0.2†	1.6
F2-δ1	1.8	1.3	-0.3†	1.1
F2-δ2	0.8	0.3	-0.3†	0.8

*: coupling was not measured and could therefore not be used. †: residual of a sum of couplings.

6.7.5. Pseudorotation plots

In the following figures, the ordering will always be as follows unless stated otherwise. Left: $CDCl_3$ samples. Right: D_2O samples. Top: *trans* conformers. Bottom: *cis* conformers.

6.7.5.1. Pseudorotation plots generated by the GUI

<u>Ac-Pro-OMe</u>

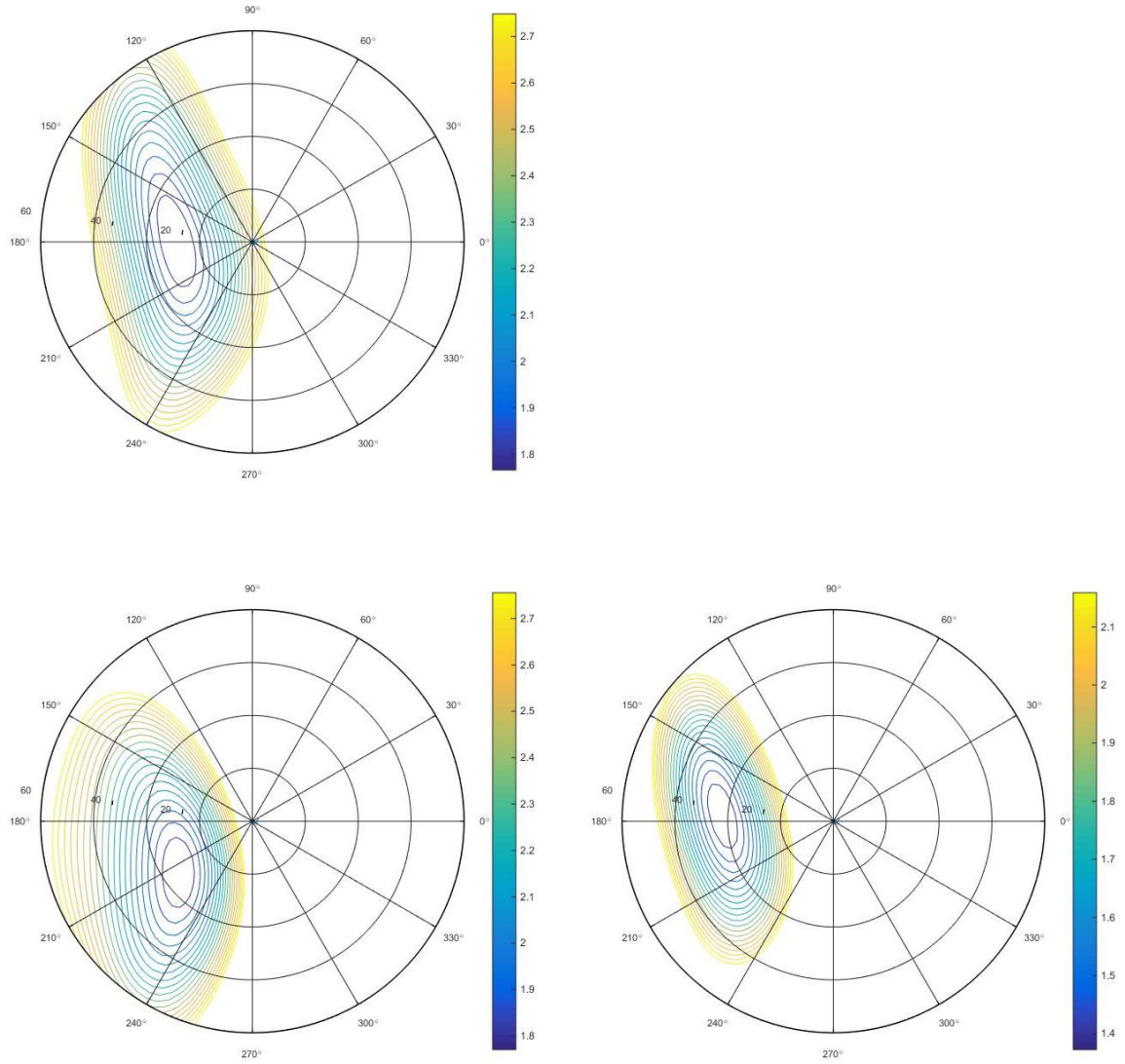


Figure 41 Pseudorotations plots from the GUI for Ac-Pro-OMe.

Ac-(4*R*)-FPro-OMe

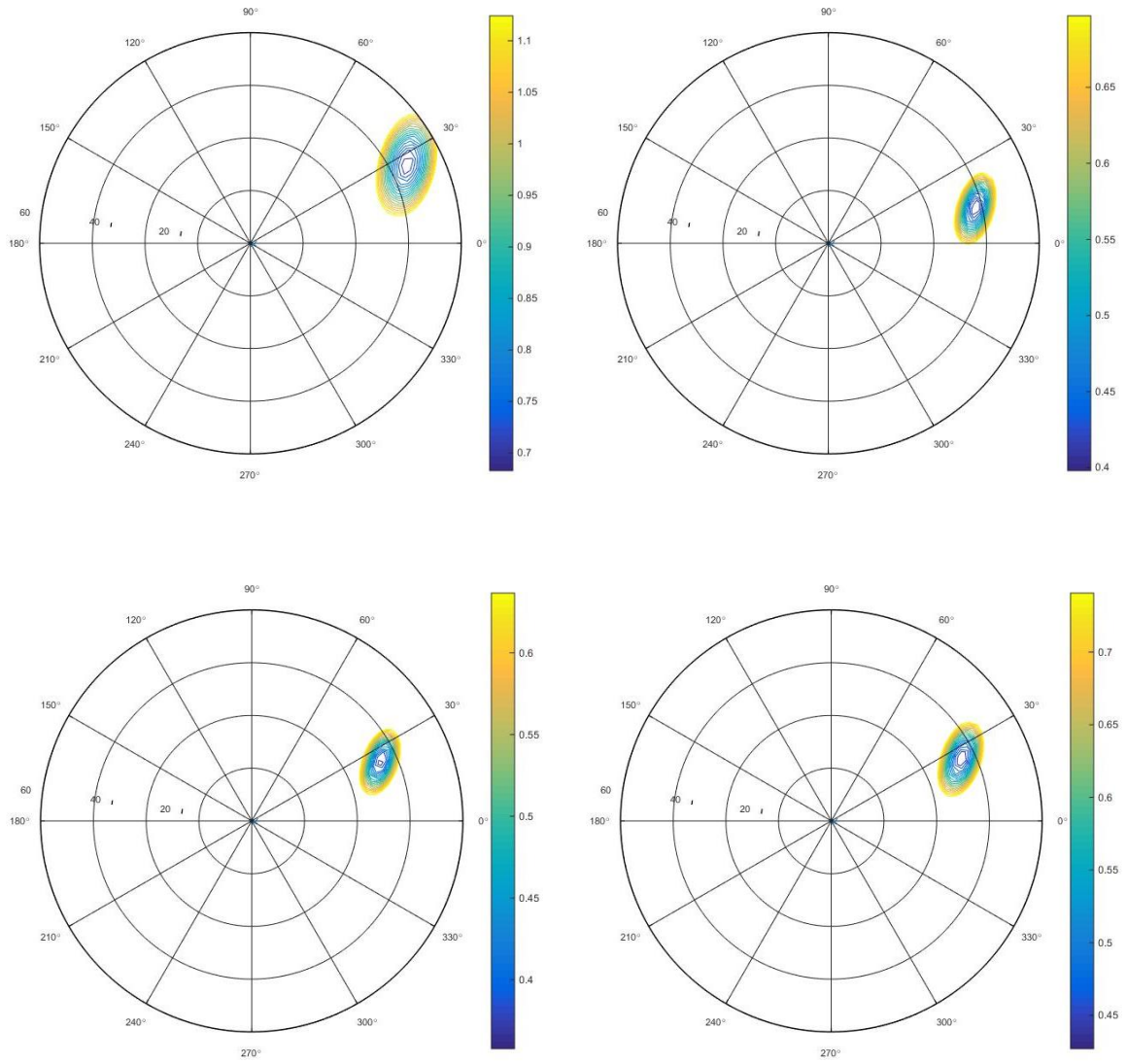


Figure 42 Pseudorotation plots from the GUI for Ac-(4R)-FPro-OMe.

Ac-(4*S*)-FPro-OMe

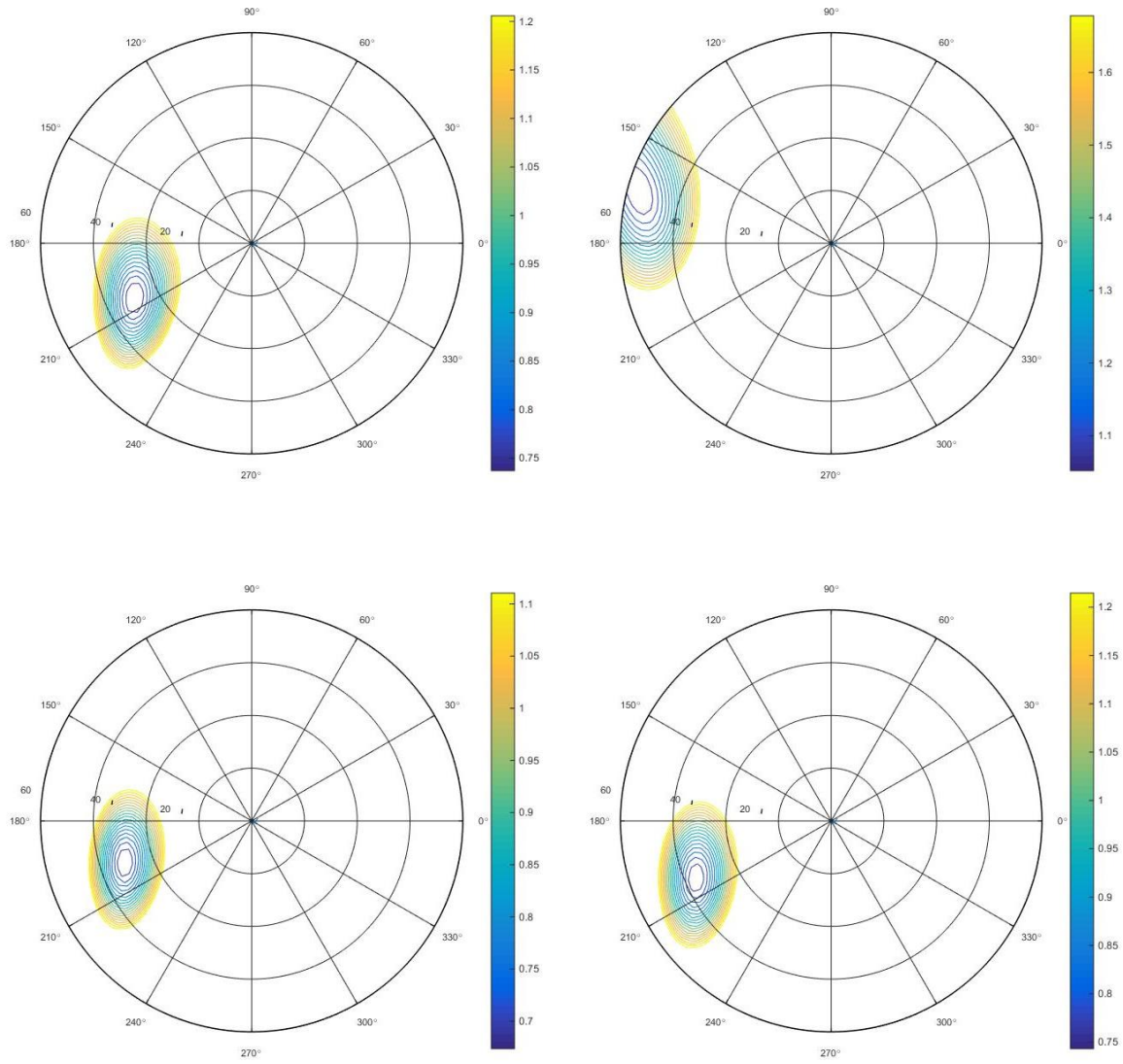


Figure 43 Pseudorotation plots from the GUI for Ac-(4S)-FPro-OMe.

6.7.5.2. Pseudorotation plots generated by the MatLab script

Warning: not all pseudorotation plots depicted below have the same scale!

Ac-(4*R*)-FPro-OMe with ^{1}H-^{19}F couplings

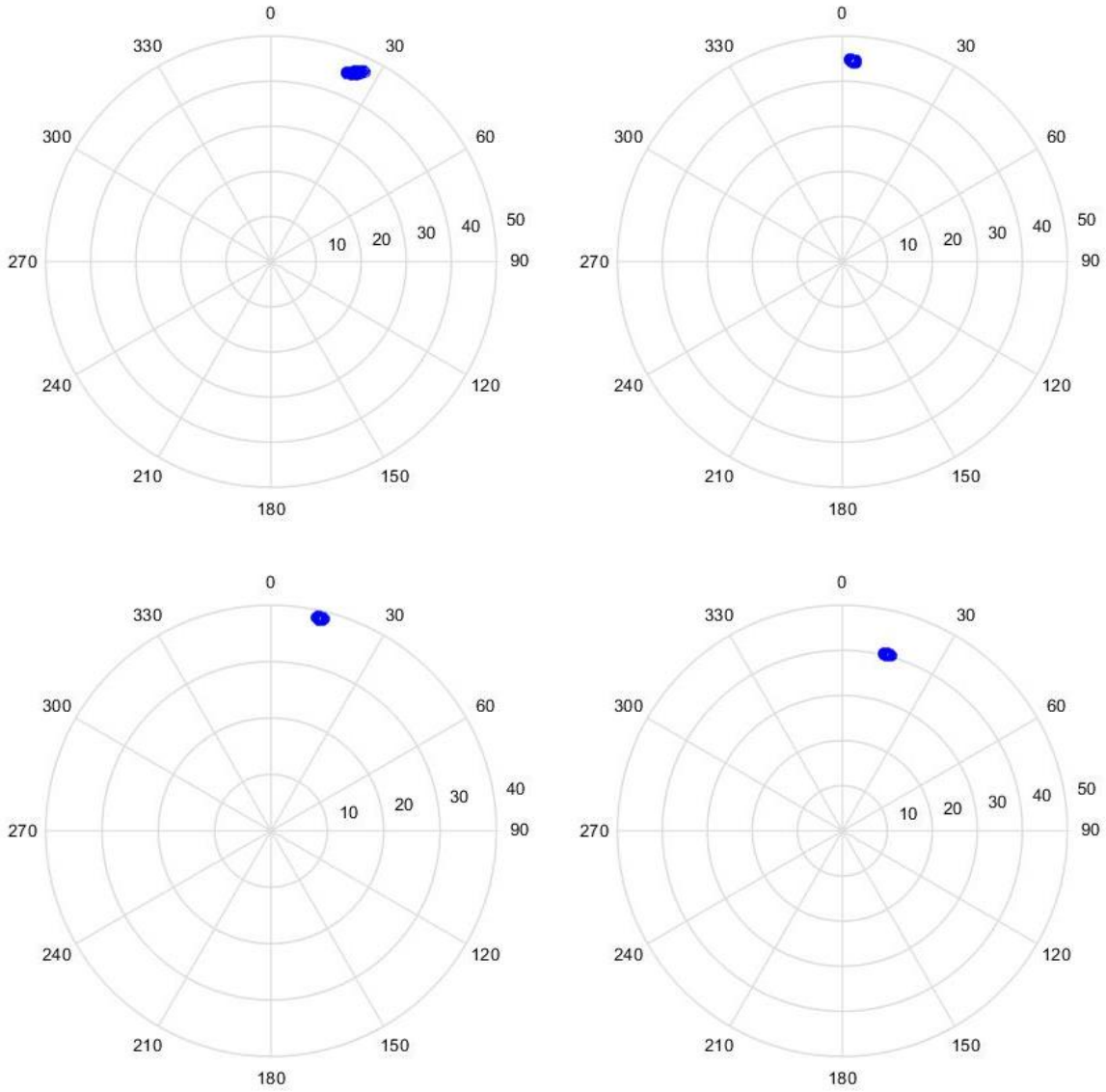


Figure 44 Pseudorotation plots from the MatLab script for Ac-(4R)-FPro-OMe, with use of ^{1}H-^{19}F couplings. For fitting one conformer.

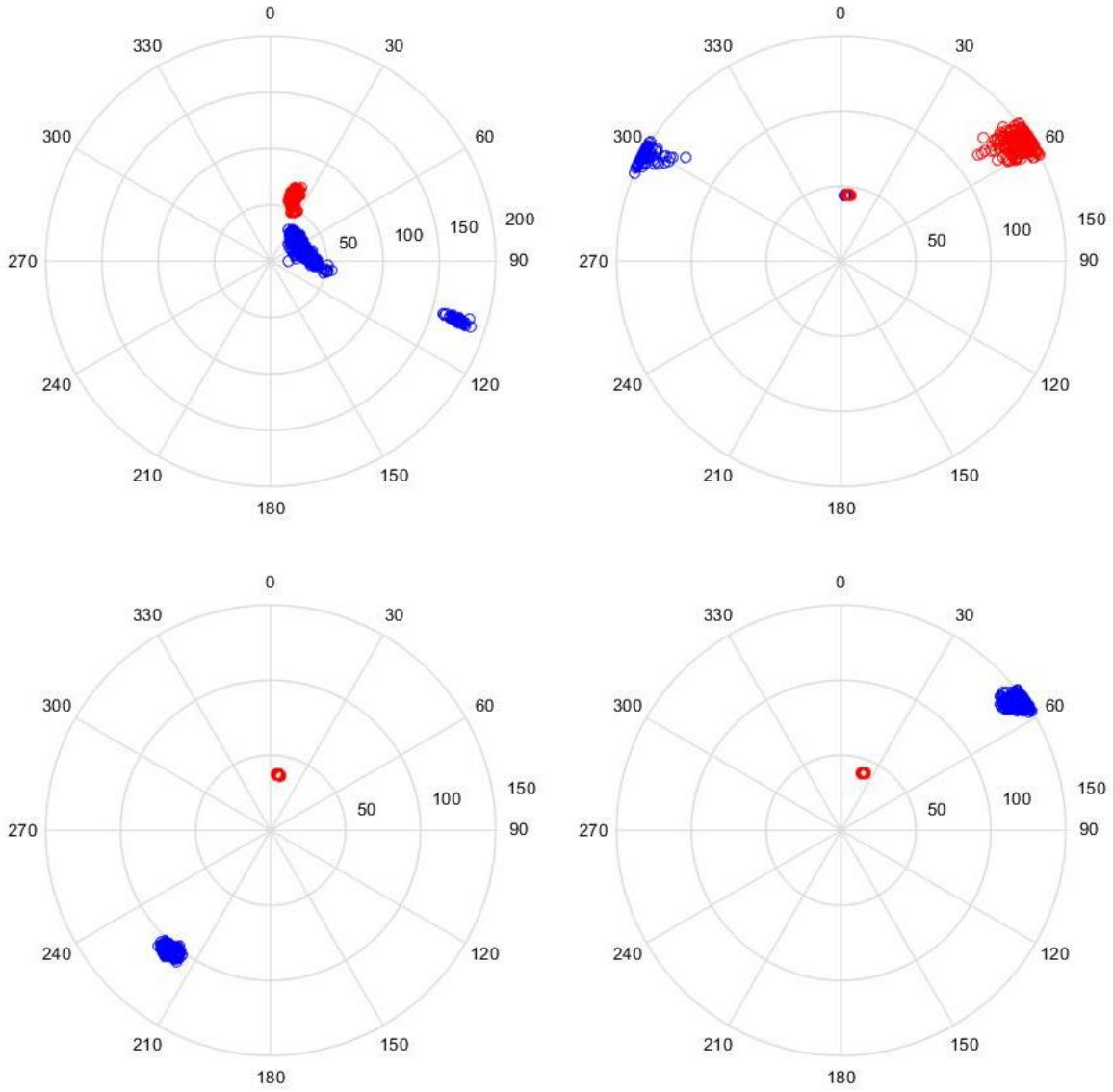


Figure 45 Pseudorotation plots from the MatLab script for Ac-(4R)-FPro-OMe, with use of ^{1}H-^{19}F couplings. For fitting two conformers.

Ac-(4*S*)-FPro-OMe with ^{1}H-^{19}F couplings

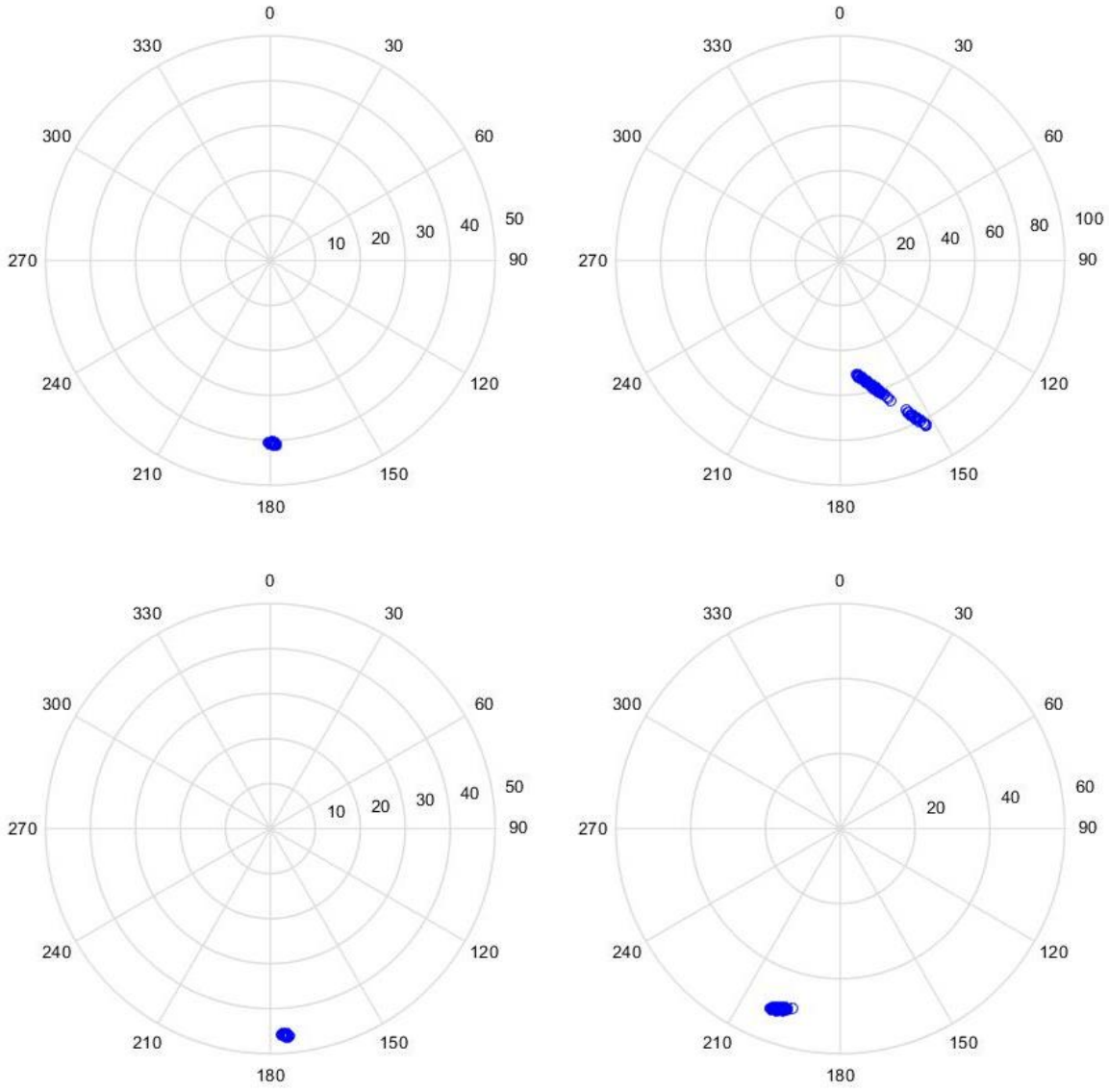


Figure 46 Pseudorotation plots from the MatLab script for Ac-(4S)-FPro-OMe, with use of ^{1}H-^{19}F couplings. For fitting one conformer.

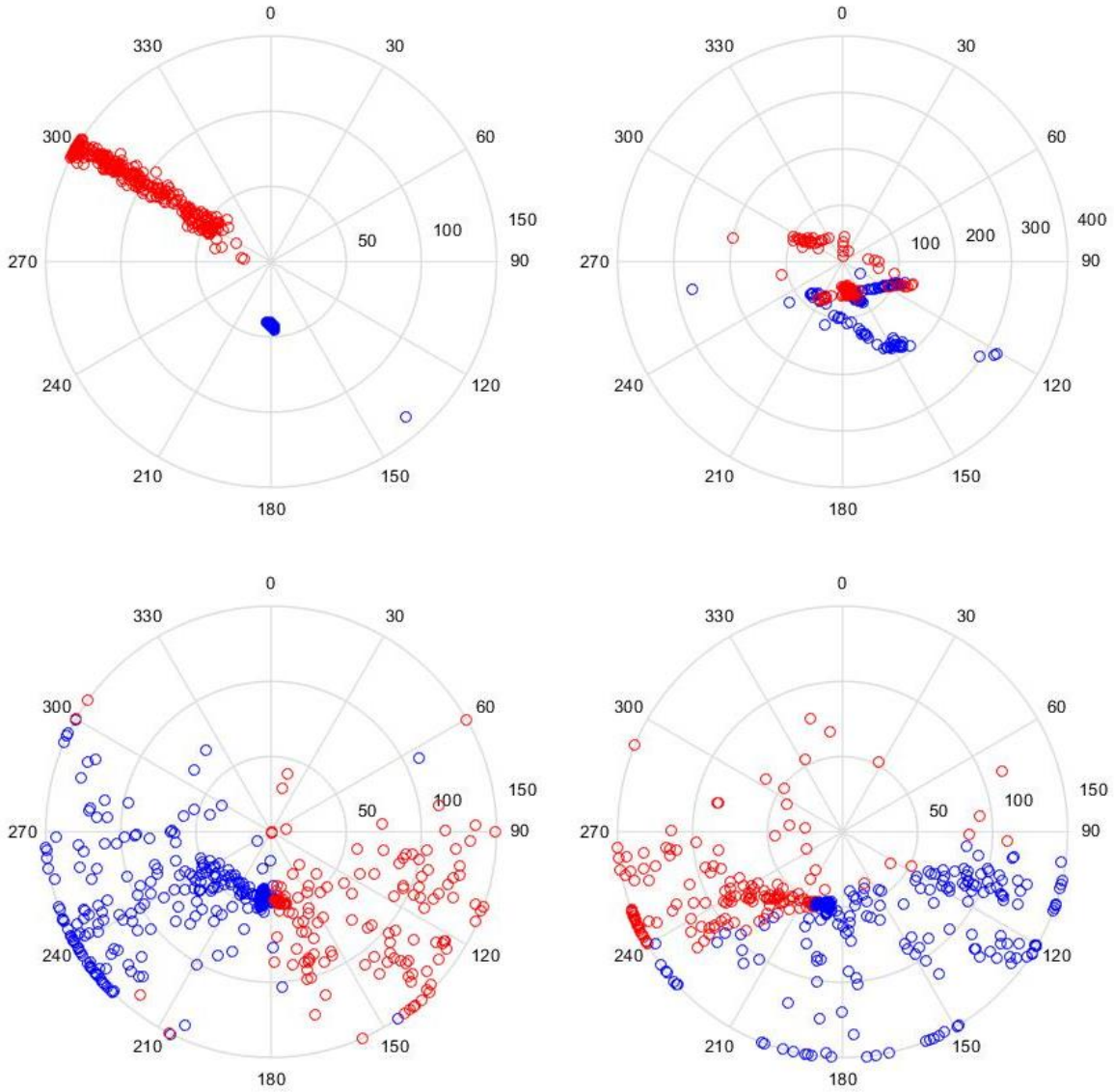


Figure 47 Pseudorotation plots from the MatLab script for Ac-(4S)-FPro-OMe, with use of ^{1}H-^{19}F couplings. For fitting two conformers.

Fitting with the Barfield correction

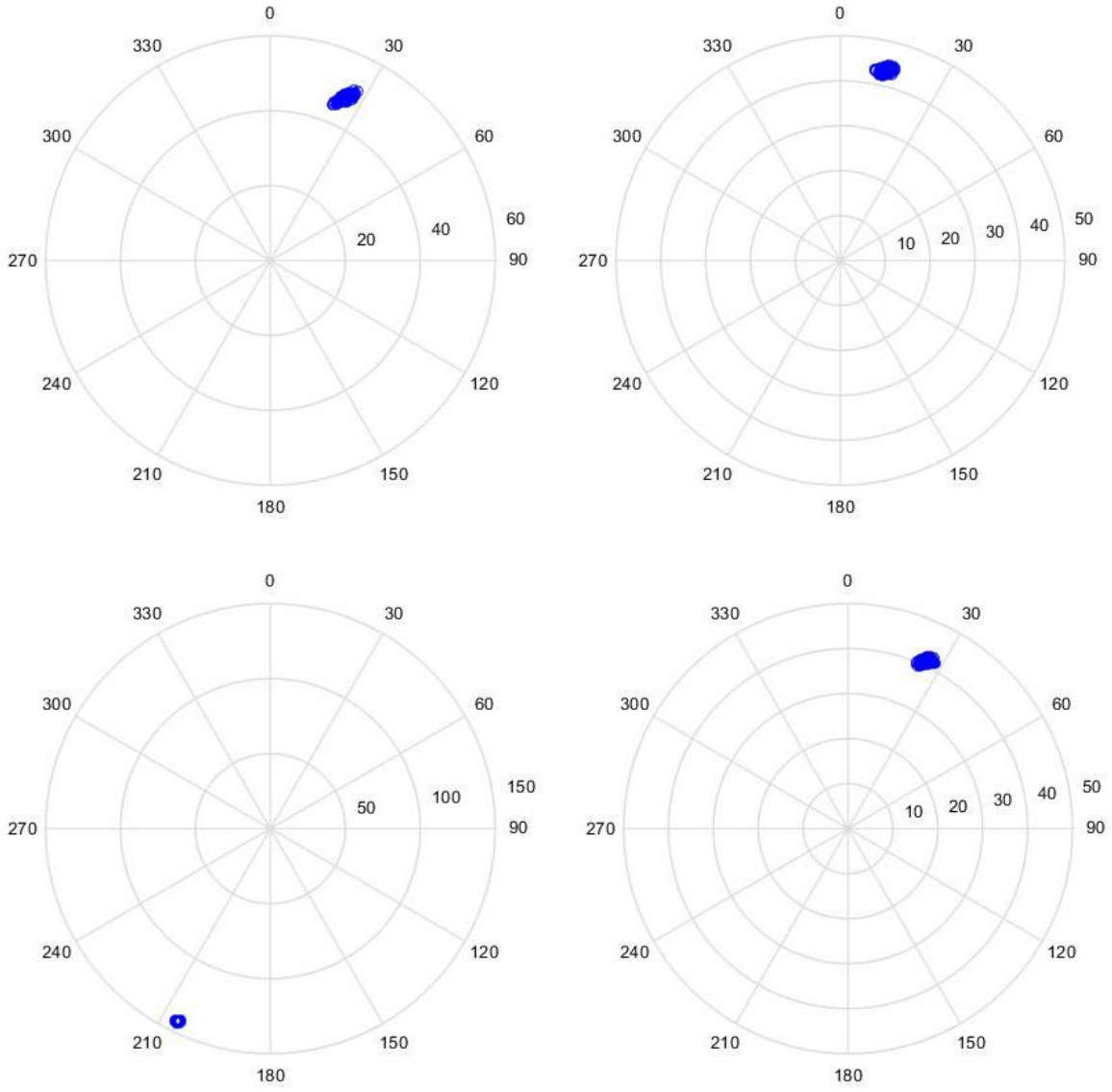


Figure 48 Pseudorotation plots from the MatLab script for Ac-(4R)-FPro-OMe, with Barfield correction, without use of ^{1}H-^{19}F couplings. For fitting one conformer.

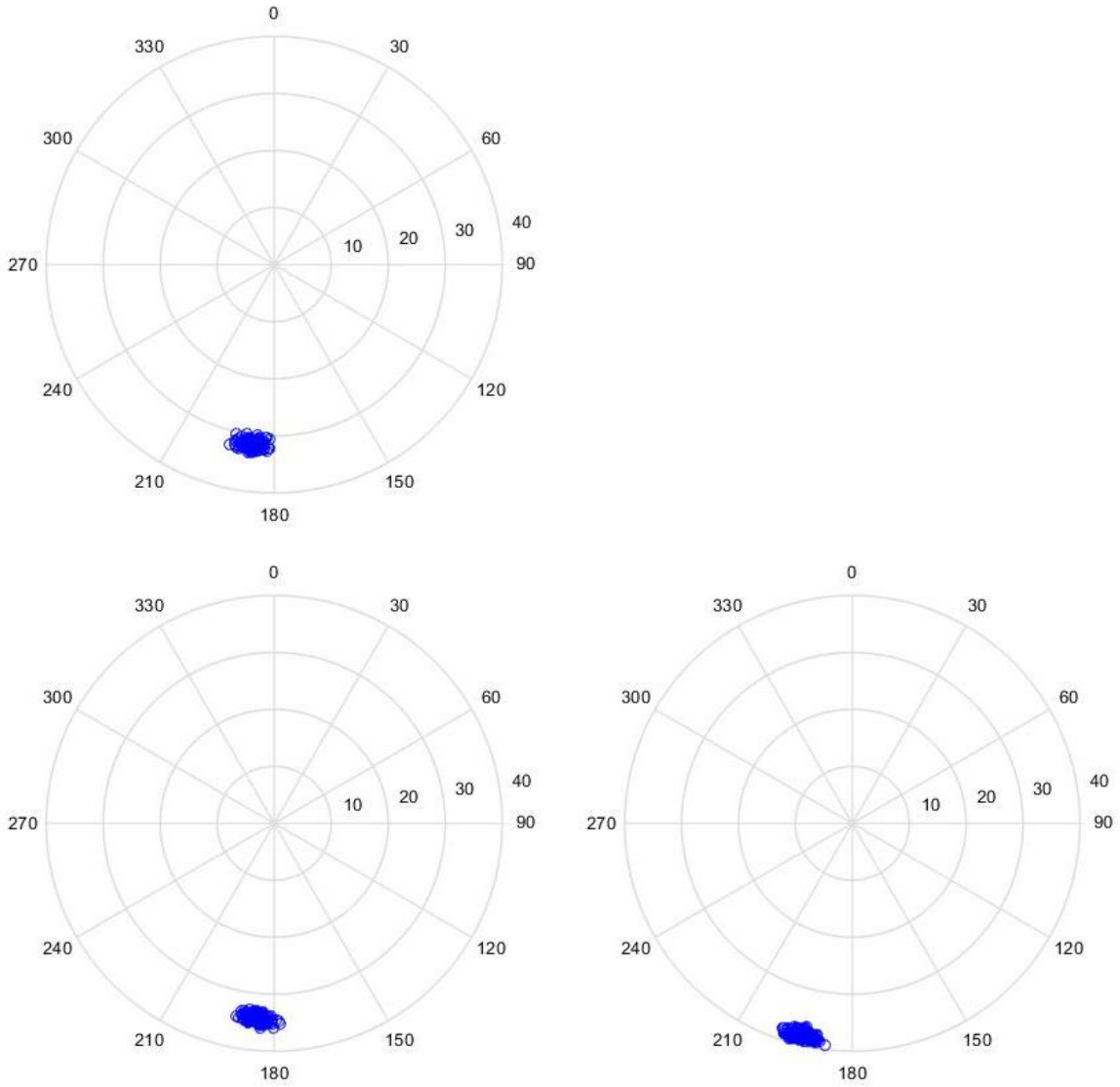


Figure 49 Pseudorotation plots from the MatLab script for Ac-(4S)-FPro-OMe, with Barfield correction, without use of ^{1}H-^{19}F couplings. For fitting one conformer.

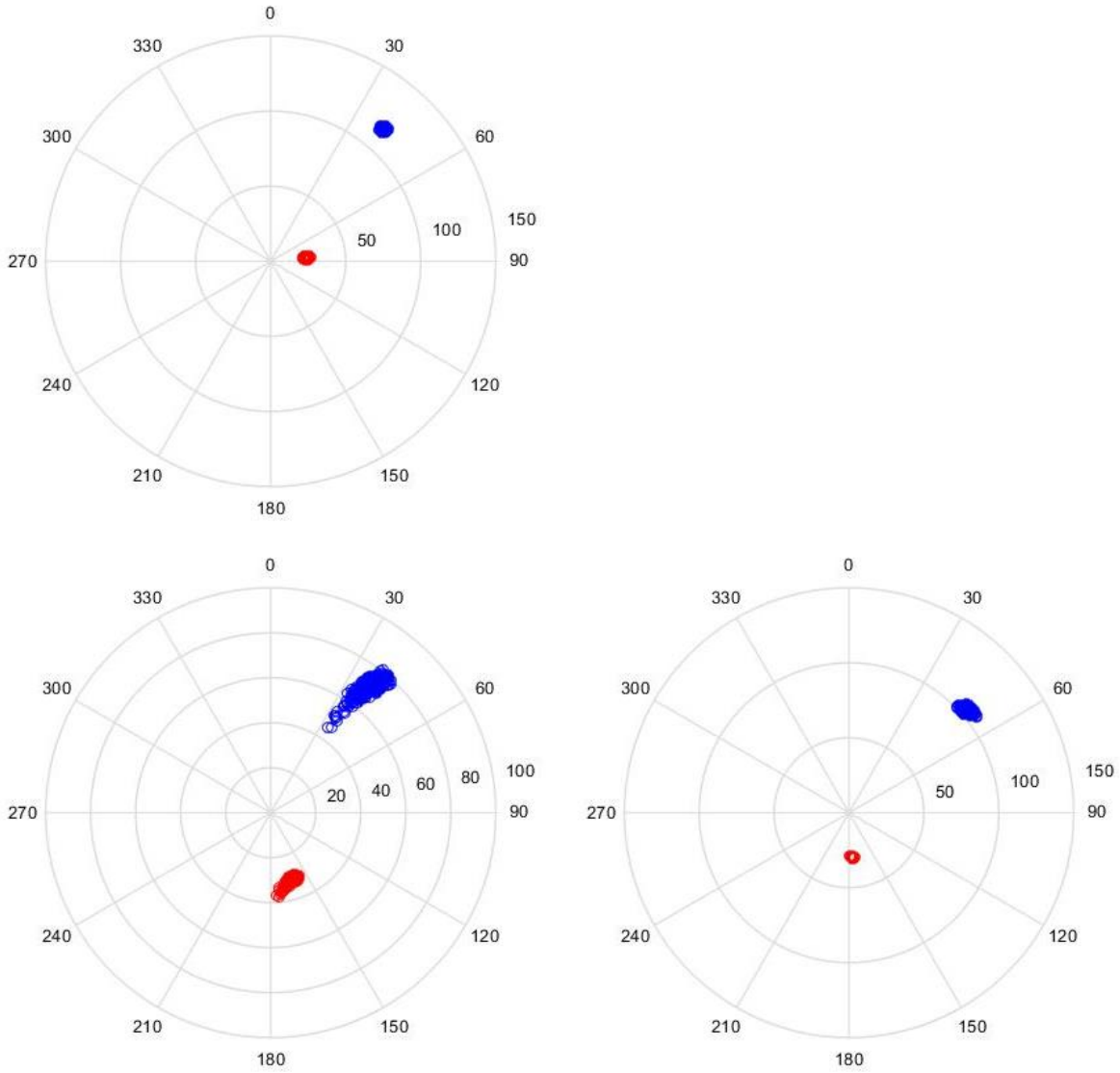


Figure 50 Pseudorotation plots from the MatLab script for Ac-4,4-F_2Pro-OMe, with Barfield correction, with use of ^{1}H-^{19}F couplings. For fitting two conformers.

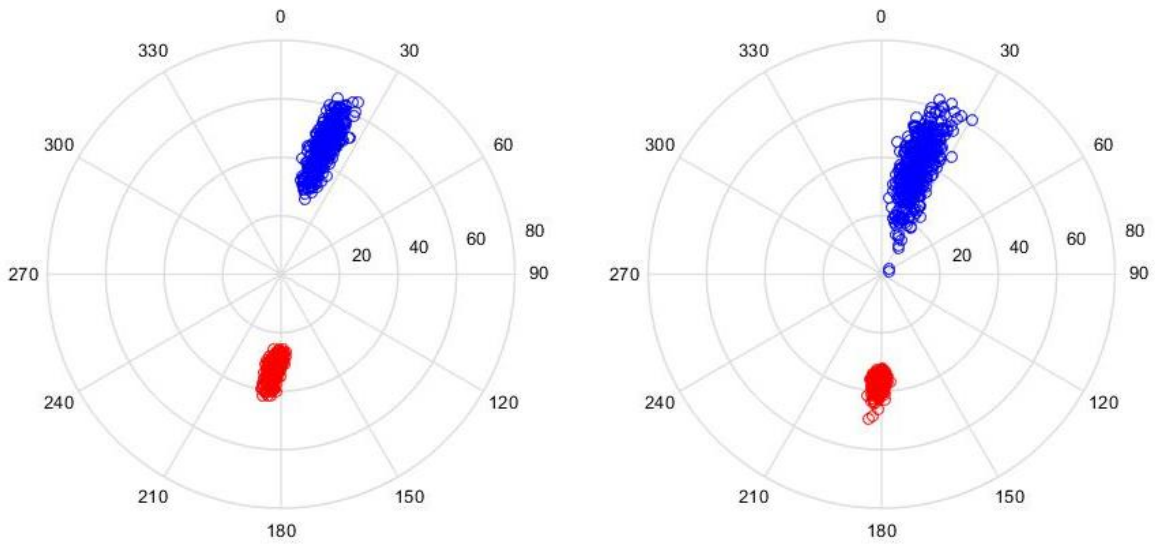


Figure 51 Pseudorotation plots from the MatLab script for Ac-Pro-OMe, with Barfield correction. For fitting two conformers. Left: trans $CDCl_3$. Right: cis D_2O.

6.7.6. Fitting fractions C^{γ} exo/endo for Ac-4,4-F_2Pro-OMe based on RMSD

The obtained RMSD plots for this procedure can be found here. ^{1}H-^{19}F couplings were used for this fitting.

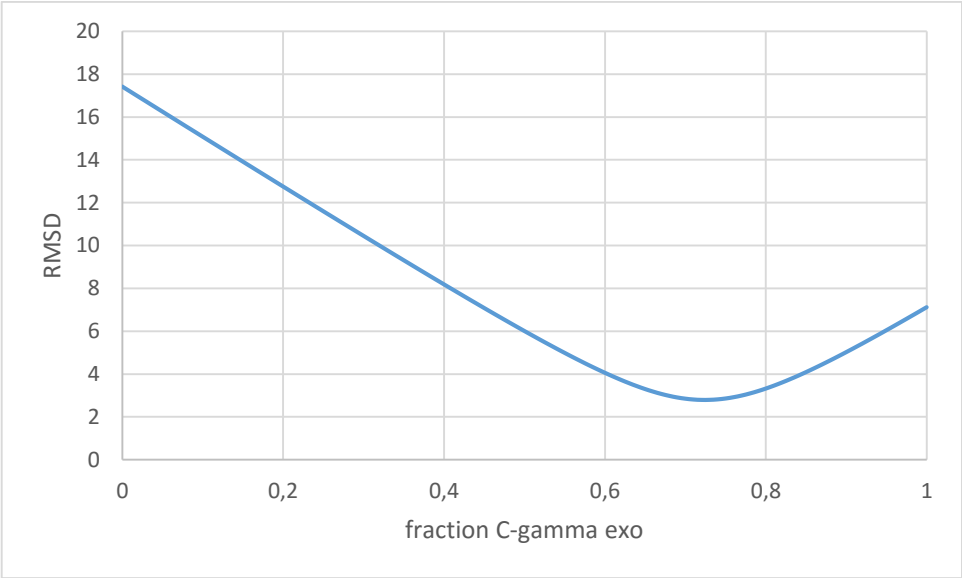


Figure 52 trans Ac-4,4-F_2Pro-OMe in $CDCl_3$.

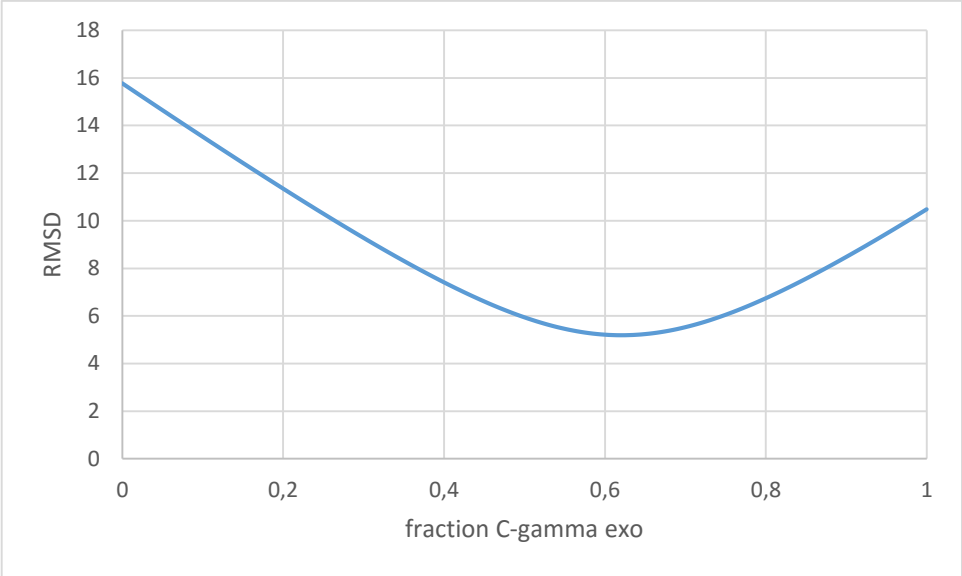


Figure 53 cis Ac-4,4-F_2Pro-OMe in $CDCl_3$.

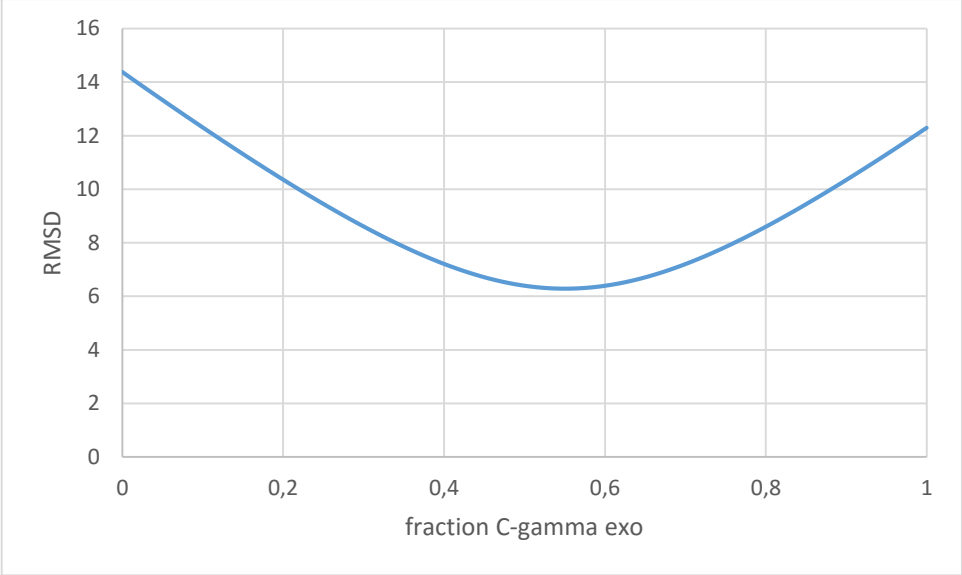


Figure 54 cis Ac-4,4-F_2Pro-OMe in D_2O.

6.8. Thermodynamics and kinetics

6.8.1. T_1 *values*

In this section, T_1 values of all the molecules in $CDCl_3$ are reported, which makes clear that these are indeed very similar for the *cis* and *trans* conformer.

Table 29 T_1 values (in s) for all $CDCl_3$ samples at 298 K and 308 K. By integration of ^{19}F signals for Ac-(4R)-FPro-OMe and Ac-(4S)-FPro-OMe or by integration of methyl ^1H signals for Ac-Pro-OMe and Ac-4,4-F2Pro-OMe.

	Ac-Pro-OMe	Ac-(4*R*)-FPro-OMe	Ac-(4*S*)-FPro-OMe	Ac-4,4-F_2Pro-OMe
$T_{1,trans}$ 298 K	2.7308 ± 0.0492	2.1890 ± 0.0024	2.1142 ± 0.0058	2.2335 ± 0.0260
$T_{1,cis}$ 298 K	2.7547 ± 0.0582	2.2270 ± 0.0099	2.0087 ± 0.0086	2.1625 ± 0.0457
$T_{1,trans}$ 308 K	3.0344 ± 0.0357	2.6118 ± 0.0030	2.4604 ± 0.0109	2.6558 ± 0.0435
$T_{1,cis}$ 308 K	3.0373 ± 0.0544	2.6468 ± 0.0079	2.3752 ± 0.0151	2.6688 ± 0.0305

6.8.2. Linear regression for determining ΔH° and ΔS°

6.8.2.1. $CDCl_3$ samples

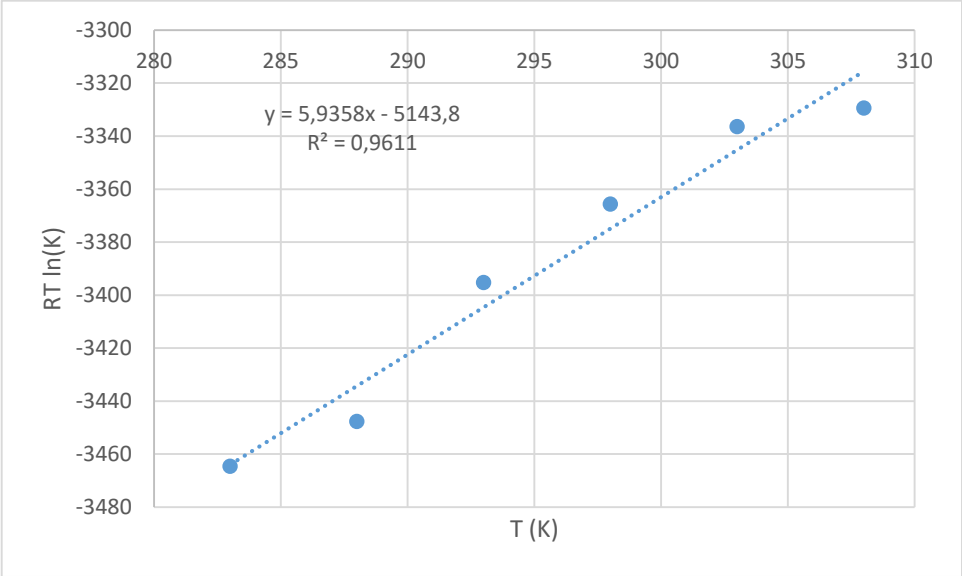


Figure 55 RT ln(K) as a function of temperature for Ac-Pro-OMe in $CDCl_3$.

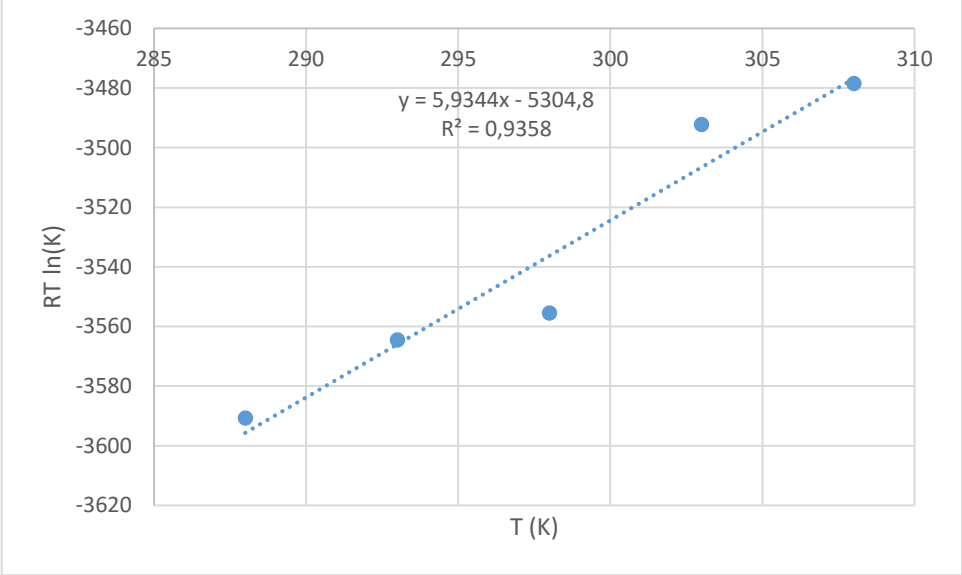


Figure 56 RT ln(K) as a function of temperature for Ac-(4R)-FPro-OMe in $CDCl_3$.

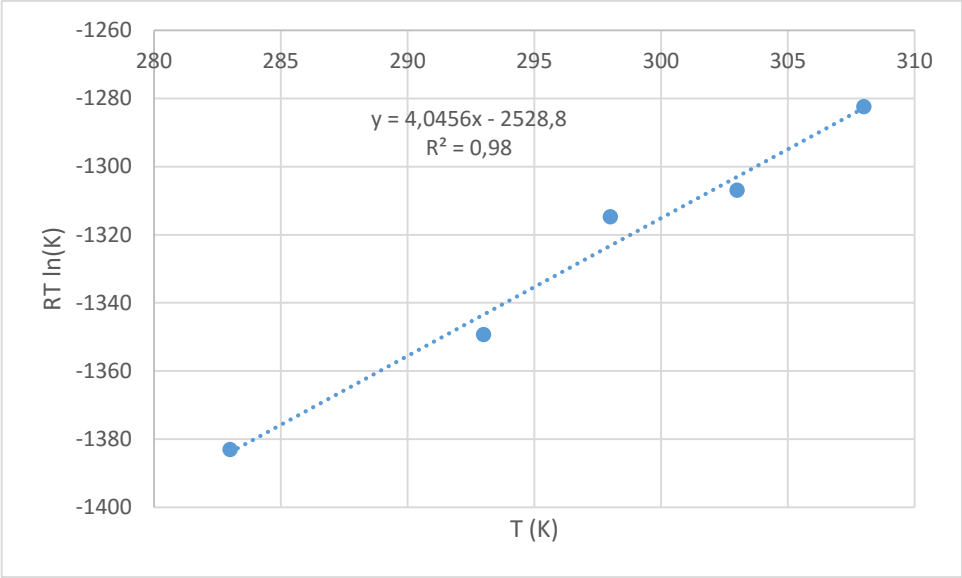


Figure 57 RT ln(K) as a function of temperature for Ac-(4S)-FPro-OMe in $CDCl_3$.

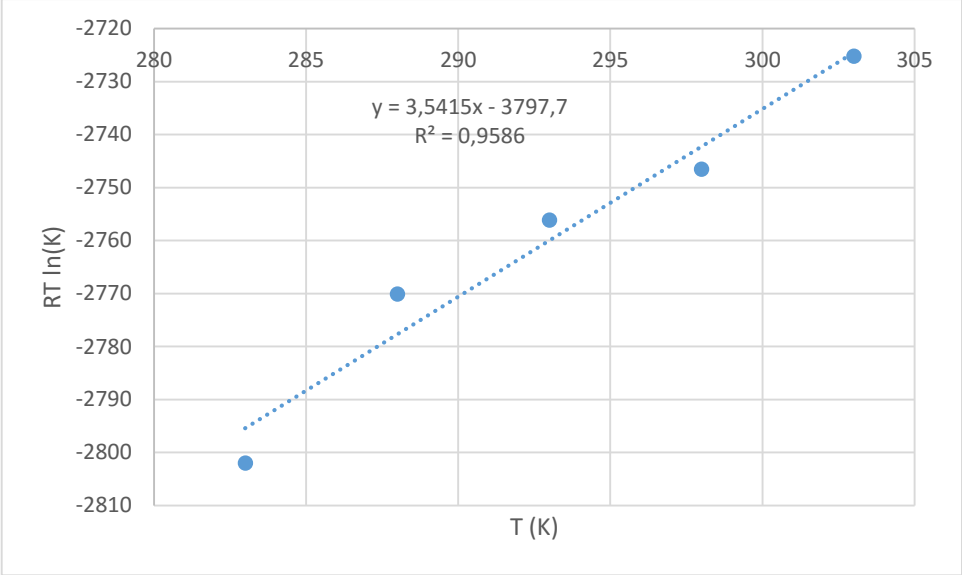


Figure 58 RT ln(K) as a function of temperature for Ac-4,4-F_2Pro-OMe in $CDCl_3$.

6.8.2.2. D_2O samples

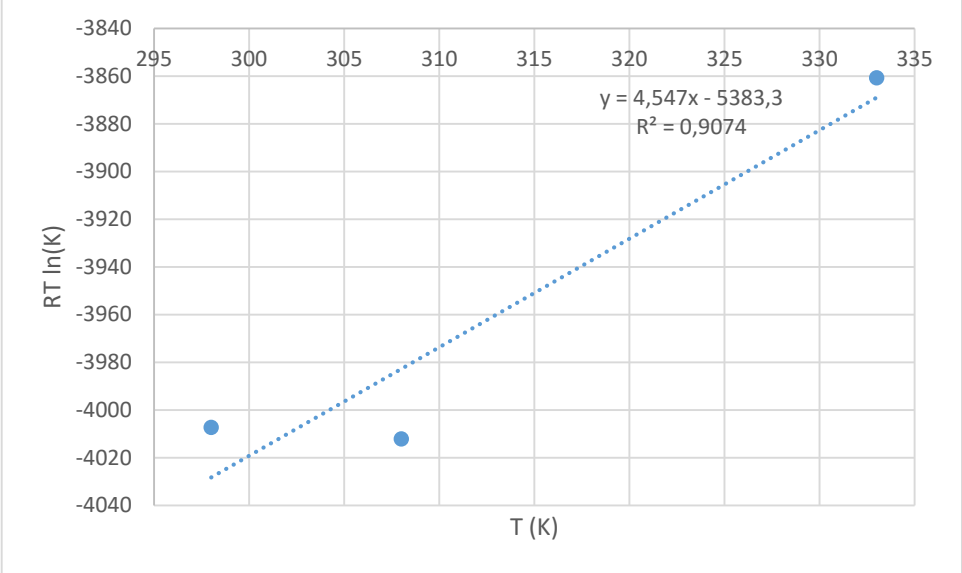


Figure 59 RT ln(K) as a function of temperature for Ac-Pro-OMe in D_2O.

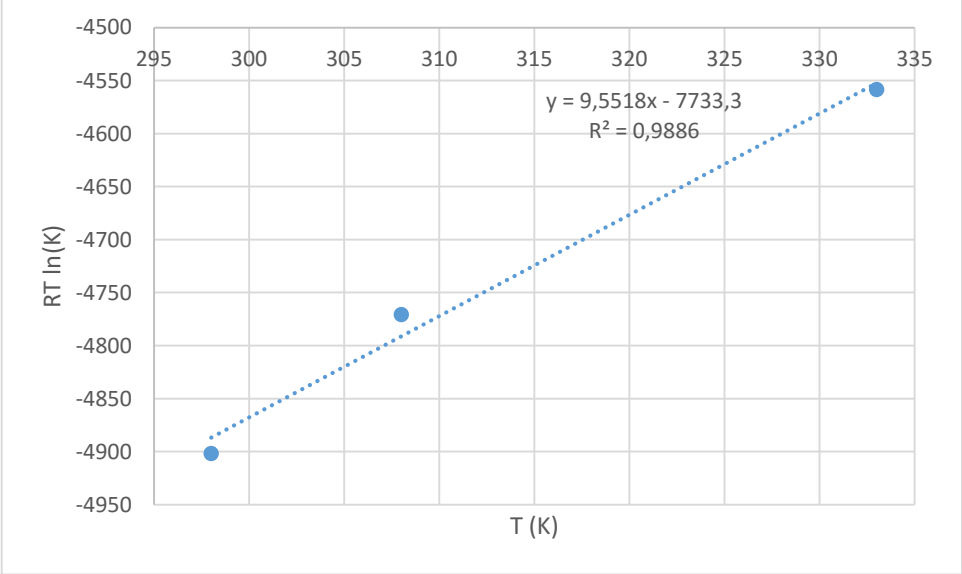


Figure 60 RT ln(K) as a function of temperature for Ac-(4R)-FPro-OMe in D_2O.

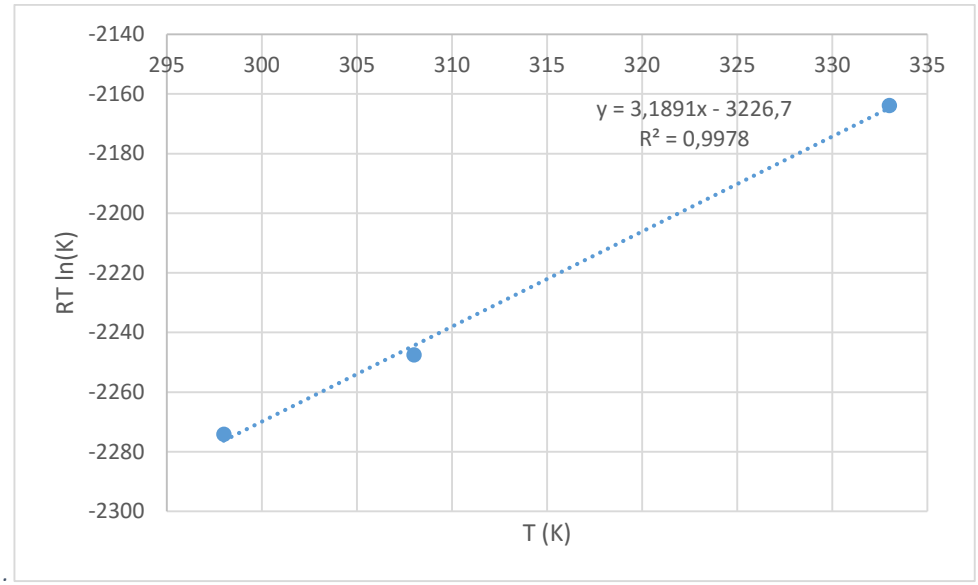


Figure 61 RT ln(K) as a function of temperature for Ac-(4S)-FPro-OMe in D_2O.

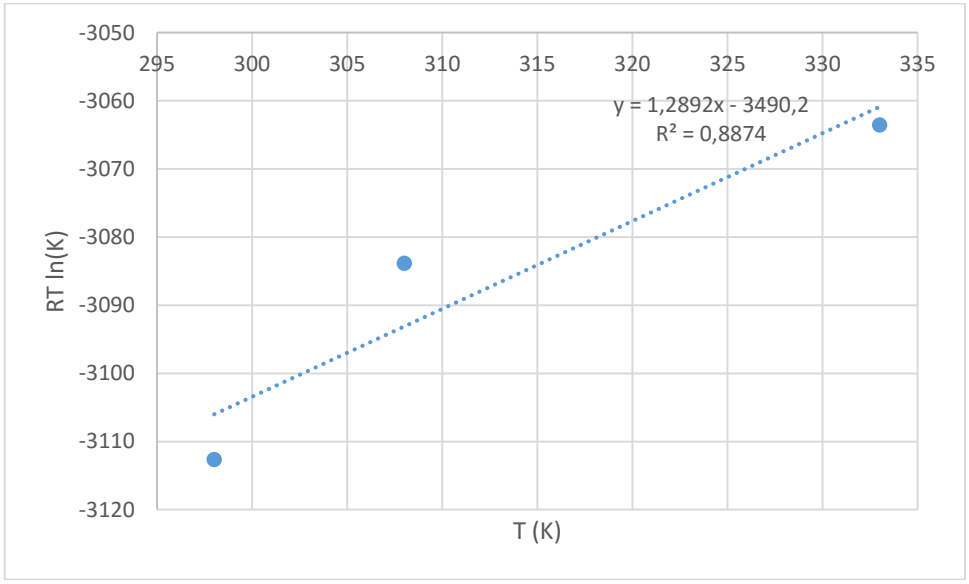


Figure 62 RT ln(K) as a function of temperature for Ac-4,4-F_2Pro-OMe in D_2O.

6.8.3. *Linear regression for determining* E_a *and* A *(Arrhenius equation)*

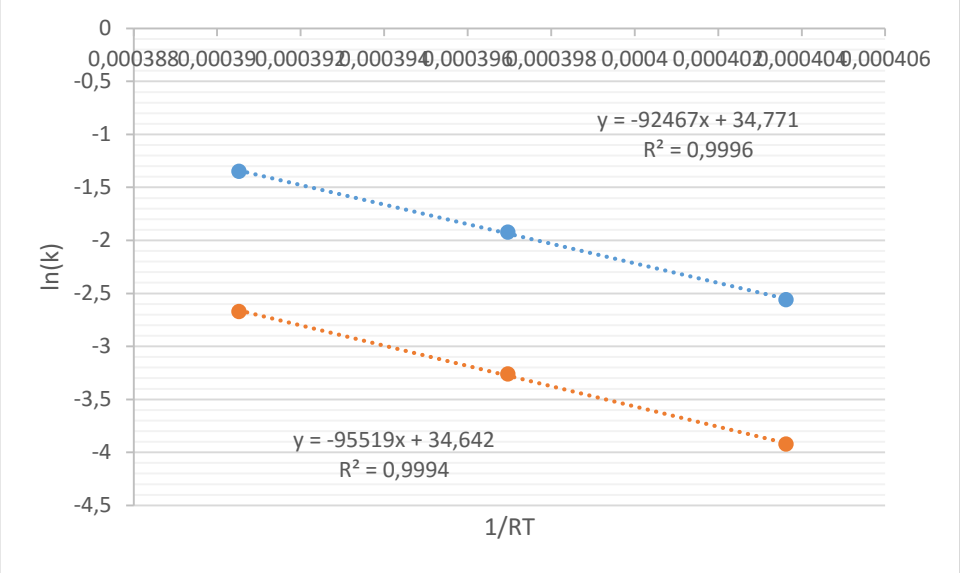


Figure 63 ln(k_1) as a function of 1/RT (blue) and ln(k_{-1}) as a function of 1/RT (orange) for Ac-Pro-OMe in $CDCl_3$.

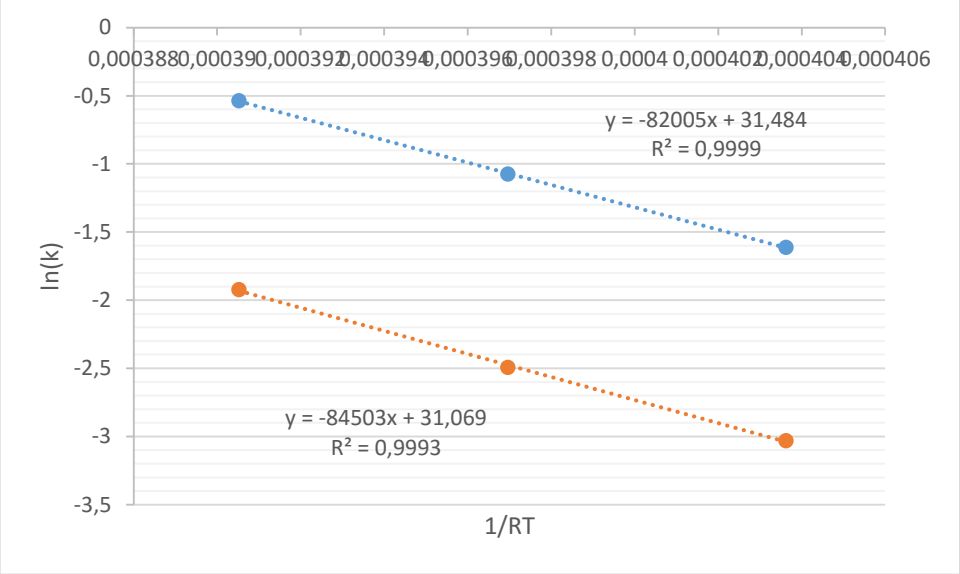


Figure 64 ln(k_1) as a function of 1/RT (blue) and ln(k_{-1}) as a function of 1/RT (orange) for Ac-(4R)-FPro-OMe in $CDCl_3$.

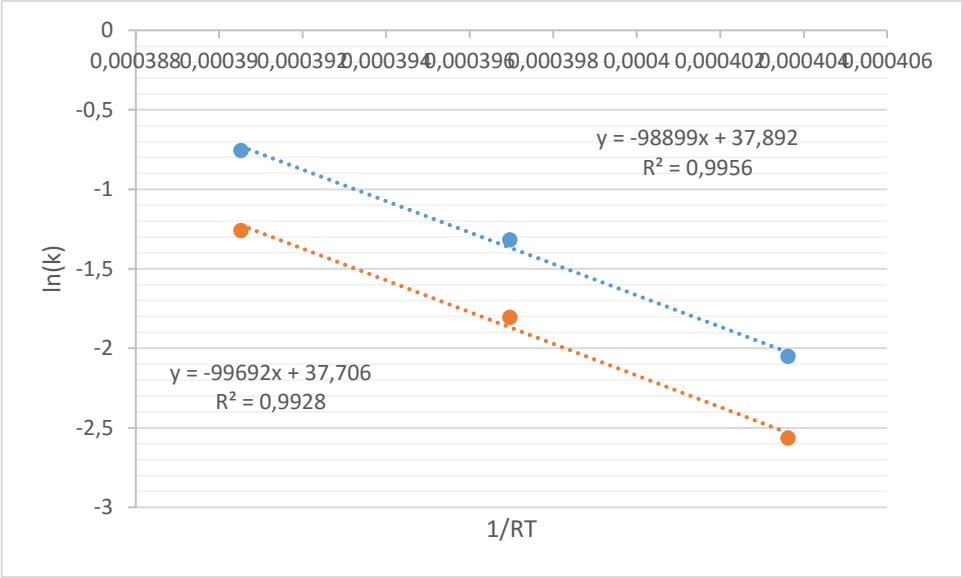


Figure 65 ln(k_1) as a function of 1/RT (blue) and ln(k_{-1}) as a function of 1/RT (orange) for Ac-(4S)-FPro-OMe in $CDCl_3$.

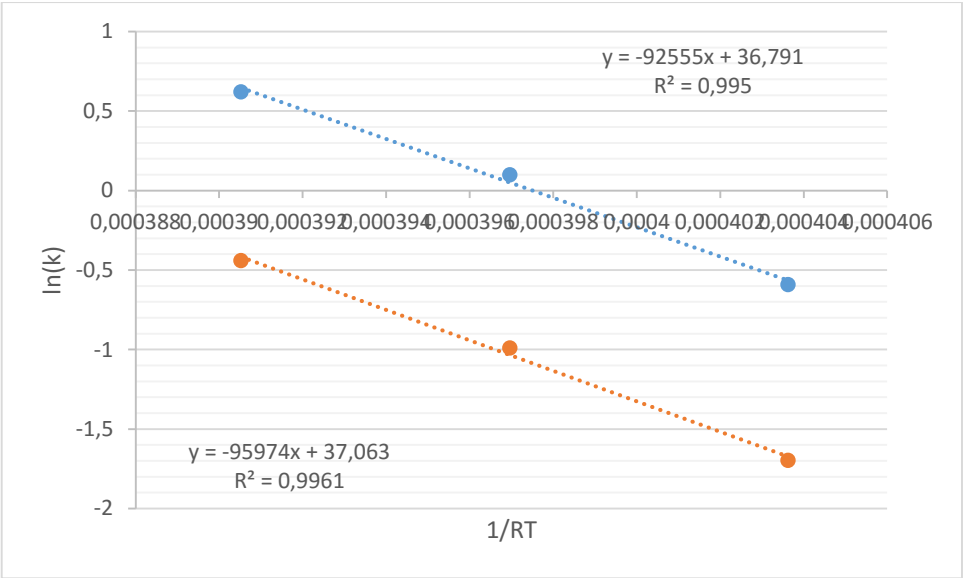


Figure 66 ln(k_1) as a function of 1/RT (blue) and ln(k_{-1}) as a function of 1/RT (orange) for Ac-4,4-F_2Pro-OMe in $CDCl_3$.

6.8.4. Linear regression for determining $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ (Eyring equation)

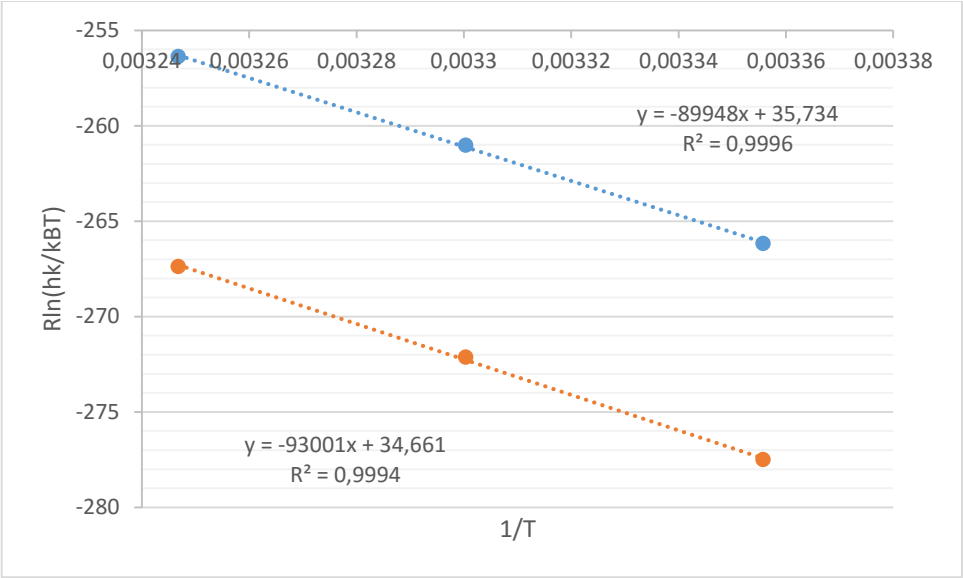


Figure 67 R ln(hk_1/k_BT) as a function of 1/T (blue) and R ln(hk_{-1}/k_BT) as a function of 1/T (orange) for Ac-Pro-OMe in $CDCl_3$.

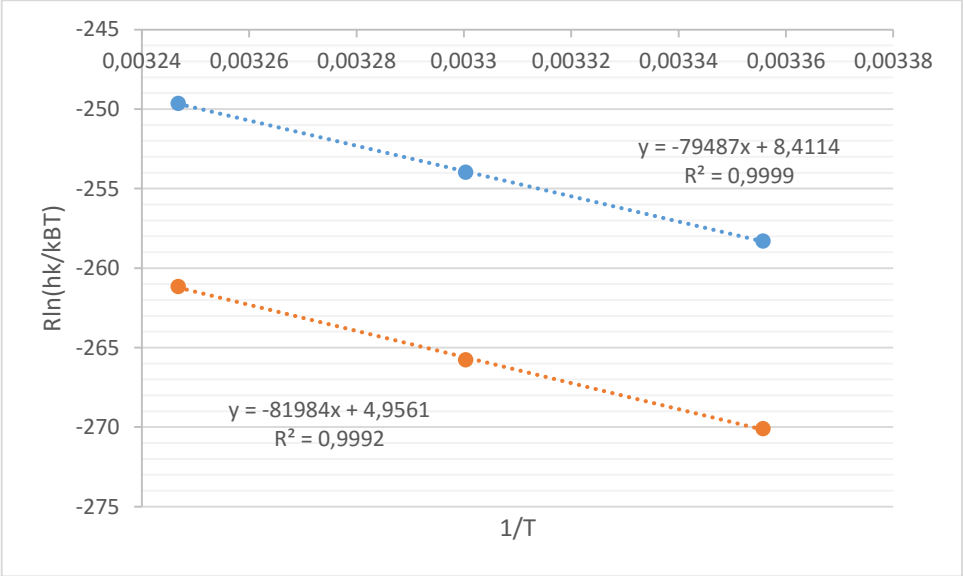


Figure 68 R ln(hk_1/k_BT) as a function of 1/T (blue) and R ln(hk_{-1}/k_BT) as a function of 1/T (orange) for Ac-(4R)-FPro-OMe in $CDCl_3$.

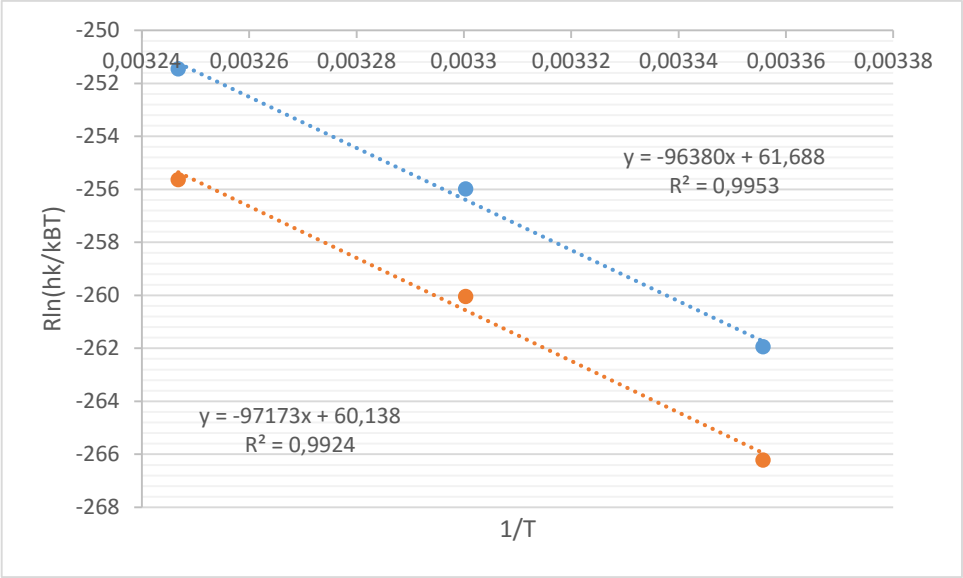


Figure 69 R ln(hk_1/k_BT) as a function of 1/T (blue) and R ln(hk_{-1}/k_BT) as a function of 1/T (orange) for Ac-(4S)-FPro-OMe in $CDCl_3$.

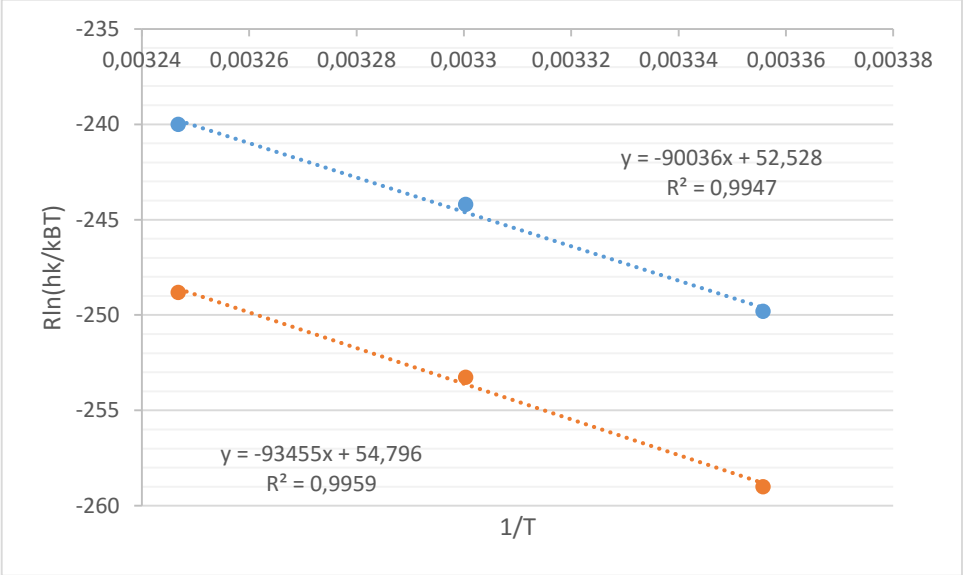


Figure 70 R ln(hk_1/k_BT) as a function of 1/T (blue) and R ln(hk_{-1}/k_BT) as a function of 1/T (orange) for Ac-4,4-F_2Pro-OMe in $CDCl_3$.

6.9. Spectra

In the following section all 1D ^{1}H (without ^{19}F decoupling), 1D ^{19}F (with ^{1}H decoupling) and 2D ^{1}H-^{13}C HSQC spectra for all samples can be found.

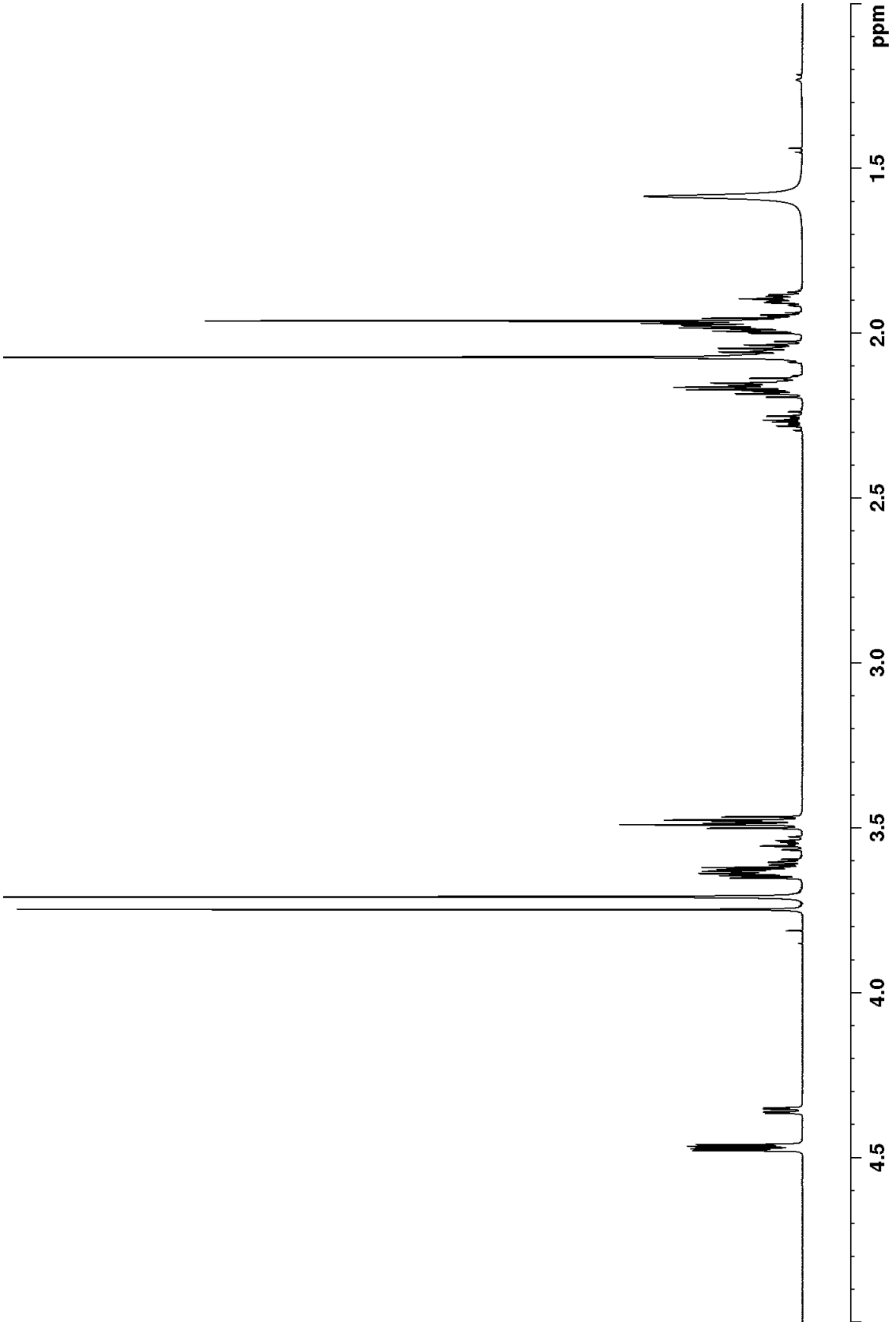


Figure 71 1D ^{1}H spectrum without ^{19}F decoupling of Ac-Pro-OMe in $CDCl_3$. 298 K. 700 MHz.

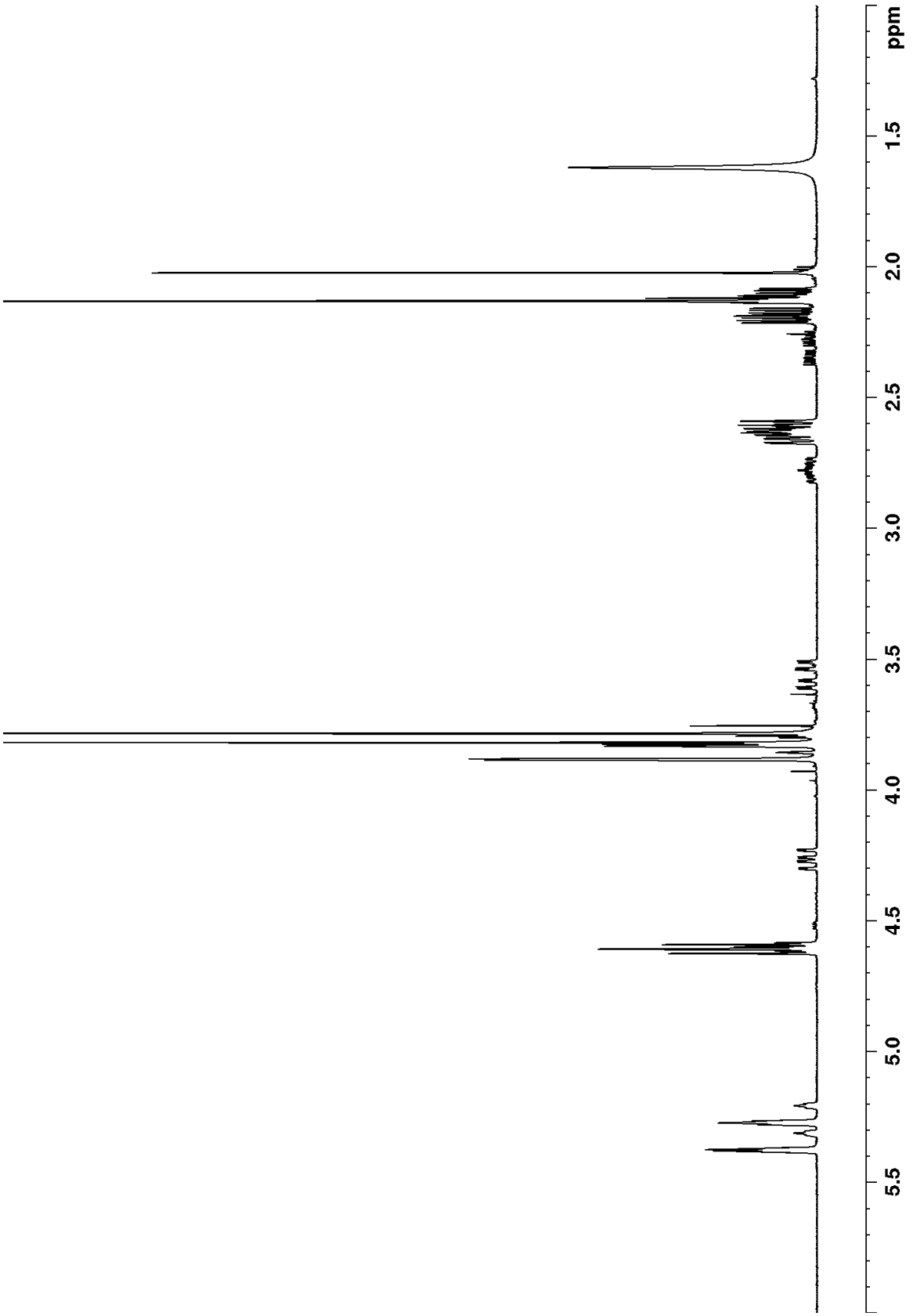


Figure 72 1D ^{1}H spectrum without ^{19}F decoupling of Ac-(4R)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

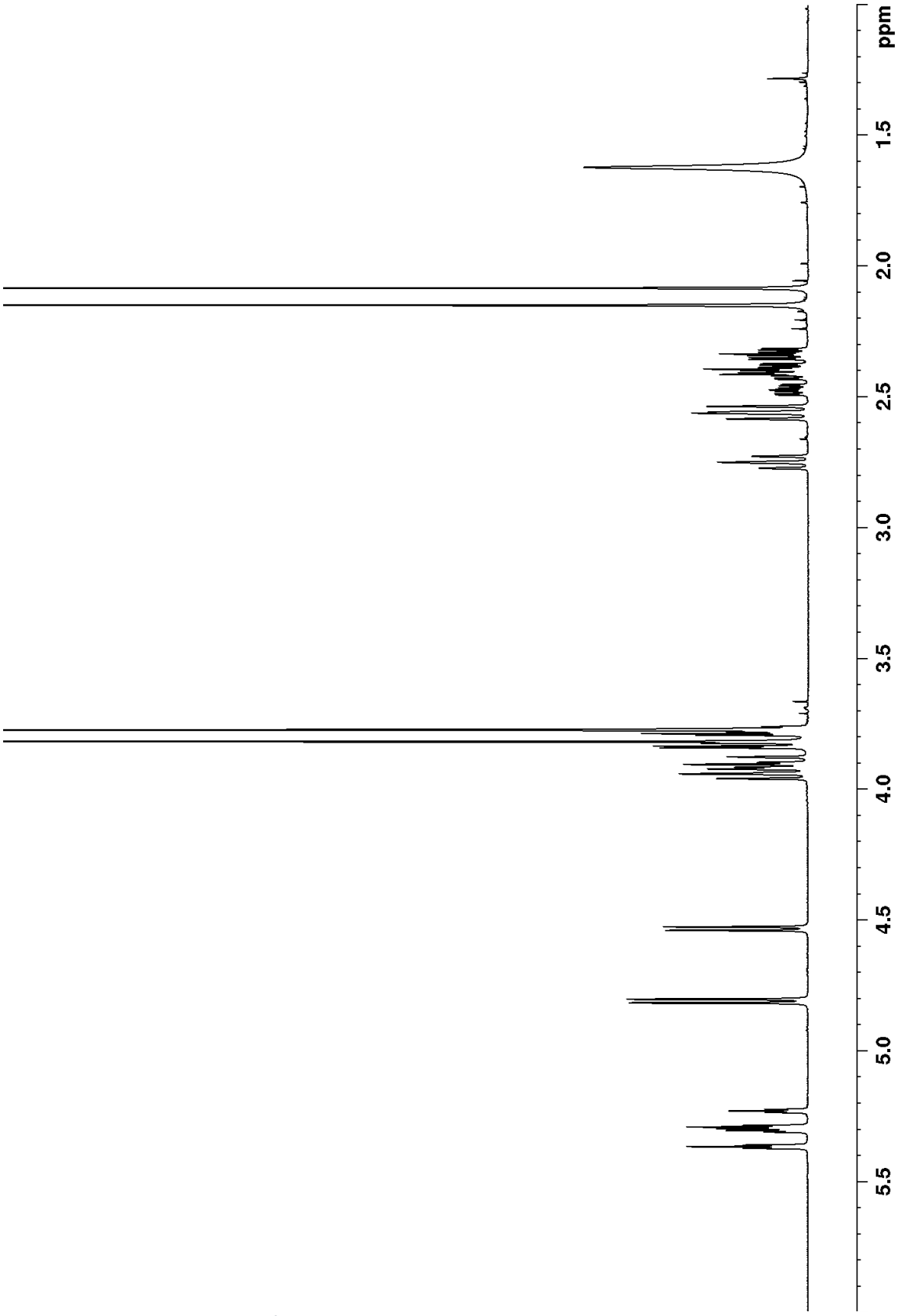


Figure 73 1D ^{1}H spectrum without ^{19}F decoupling of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 700 MHz.

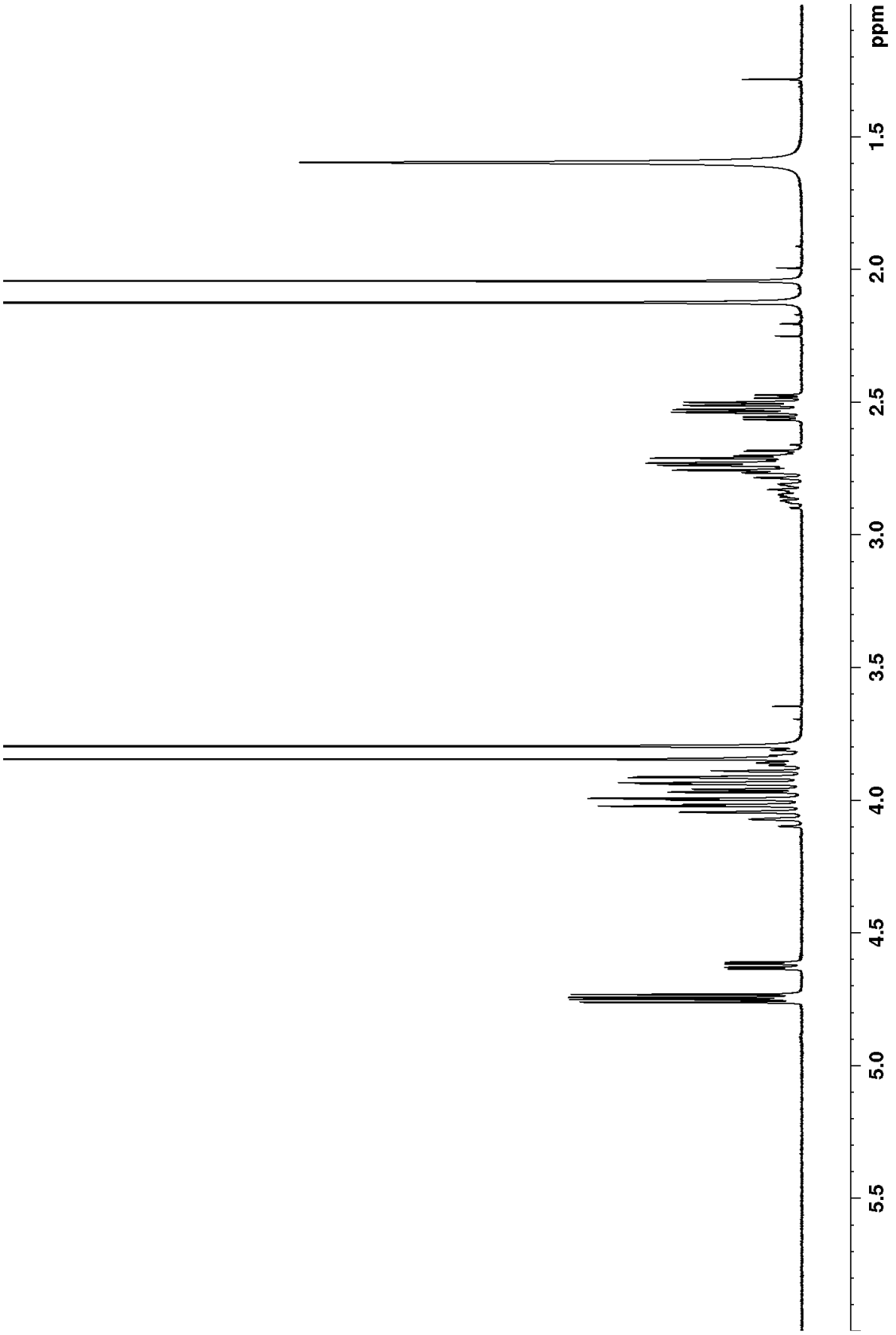


Figure 74 1D ^{1}H spectrum without ^{19}F decoupling of Ac-4,4-F_2Pro-OMe in $CDCl_3$. 298 K. 500 MHz.

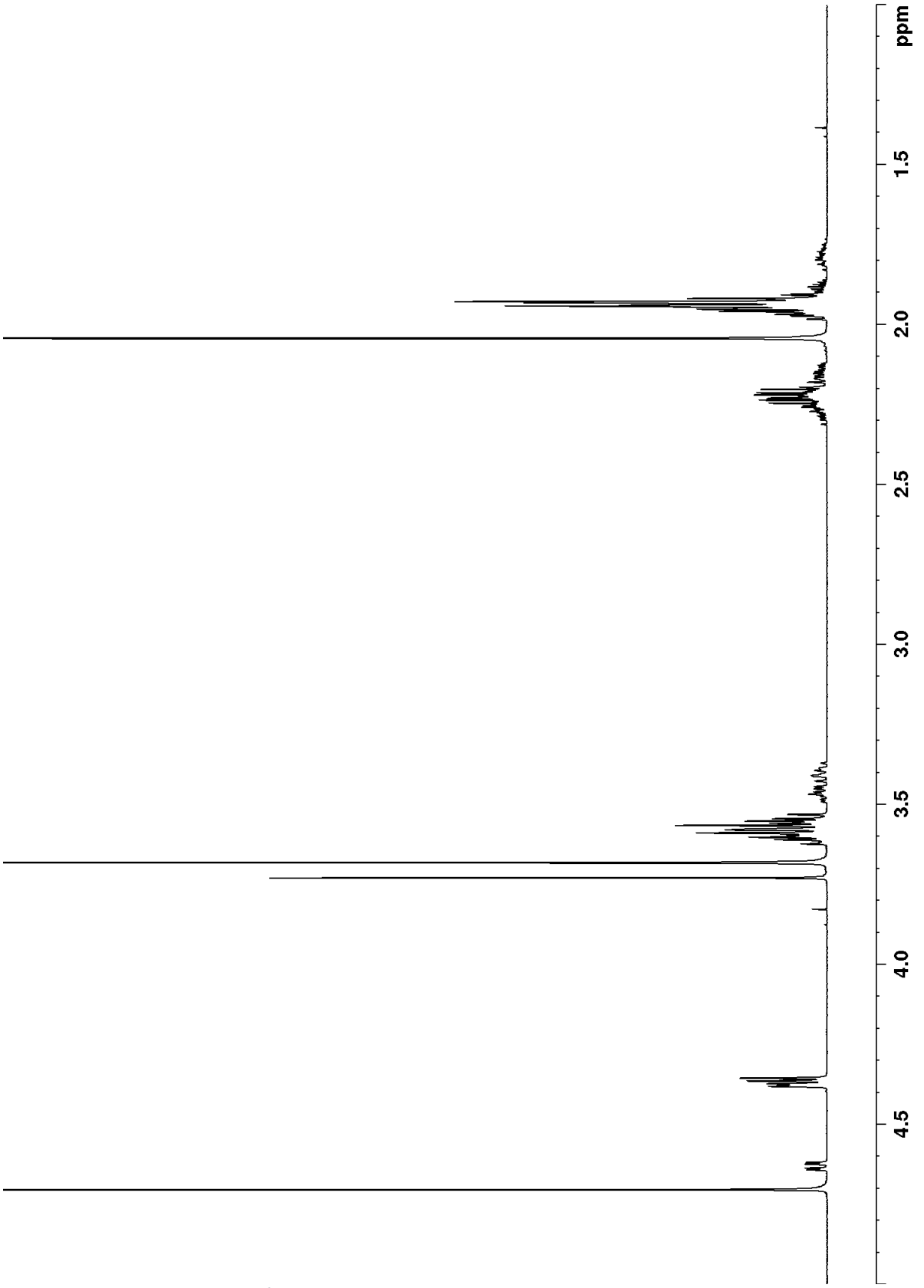


Figure 75 1D ^{1}H spectrum without ^{19}F decoupling of Ac-Pro-OMe in D_2O. 298 K. 500 MHz.

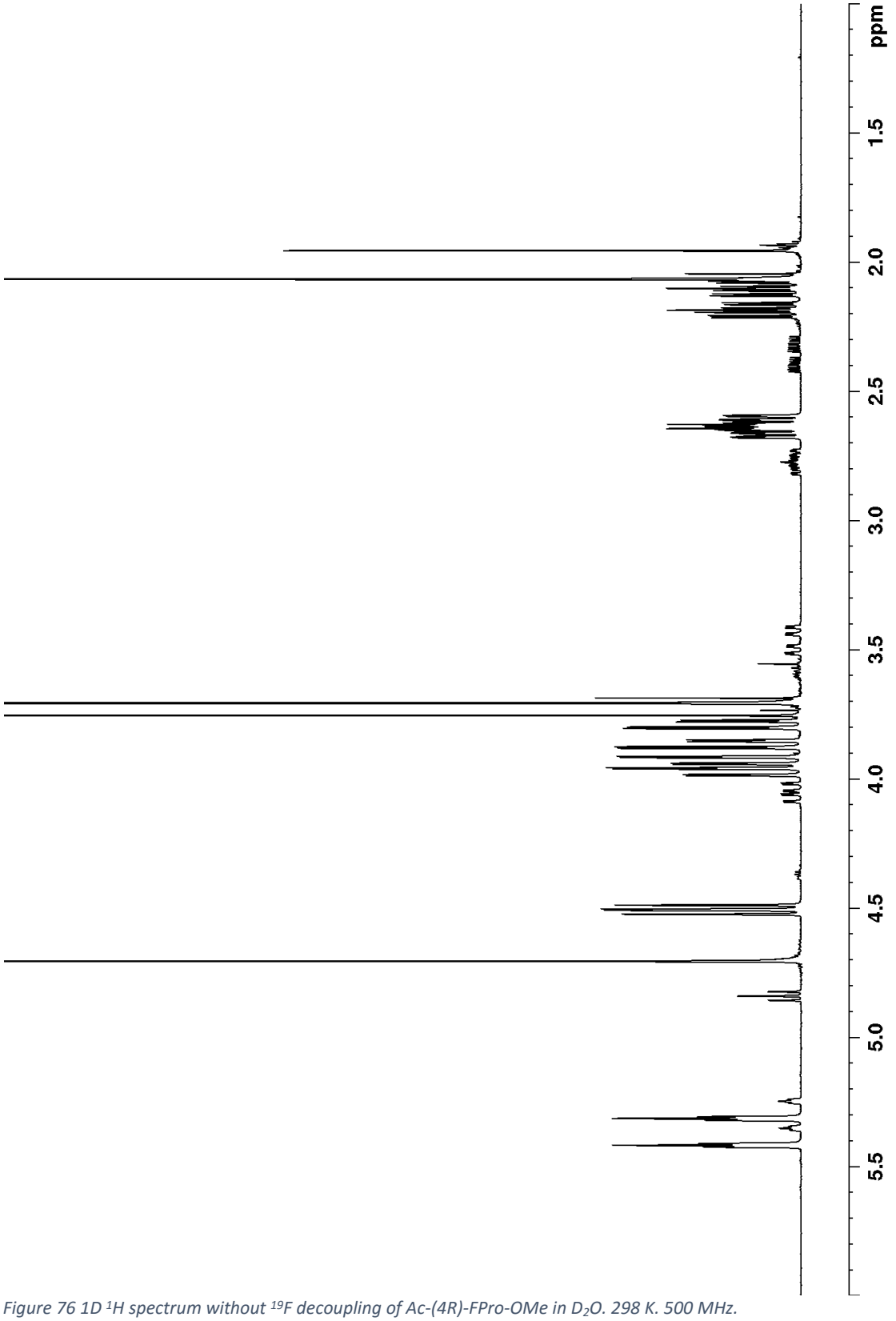


Figure 76 1D ^{1}H spectrum without ^{19}F decoupling of Ac-(4R)-FPro-OMe in D_2O. 298 K. 500 MHz.

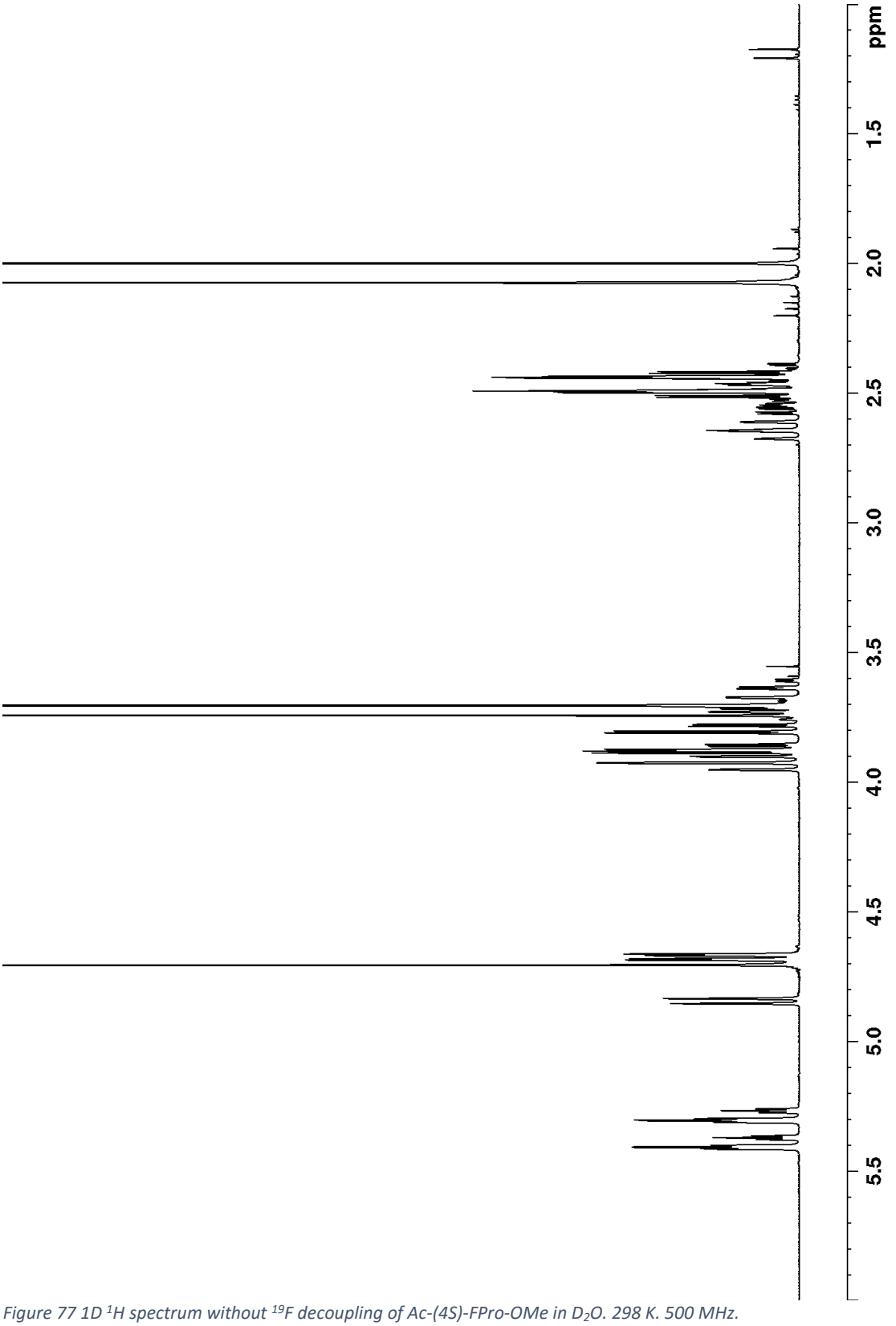


Figure 77 1D ^{1}H spectrum without ^{19}F decoupling of Ac-(4S)-FPro-OMe in D_2O. 298 K. 500 MHz.

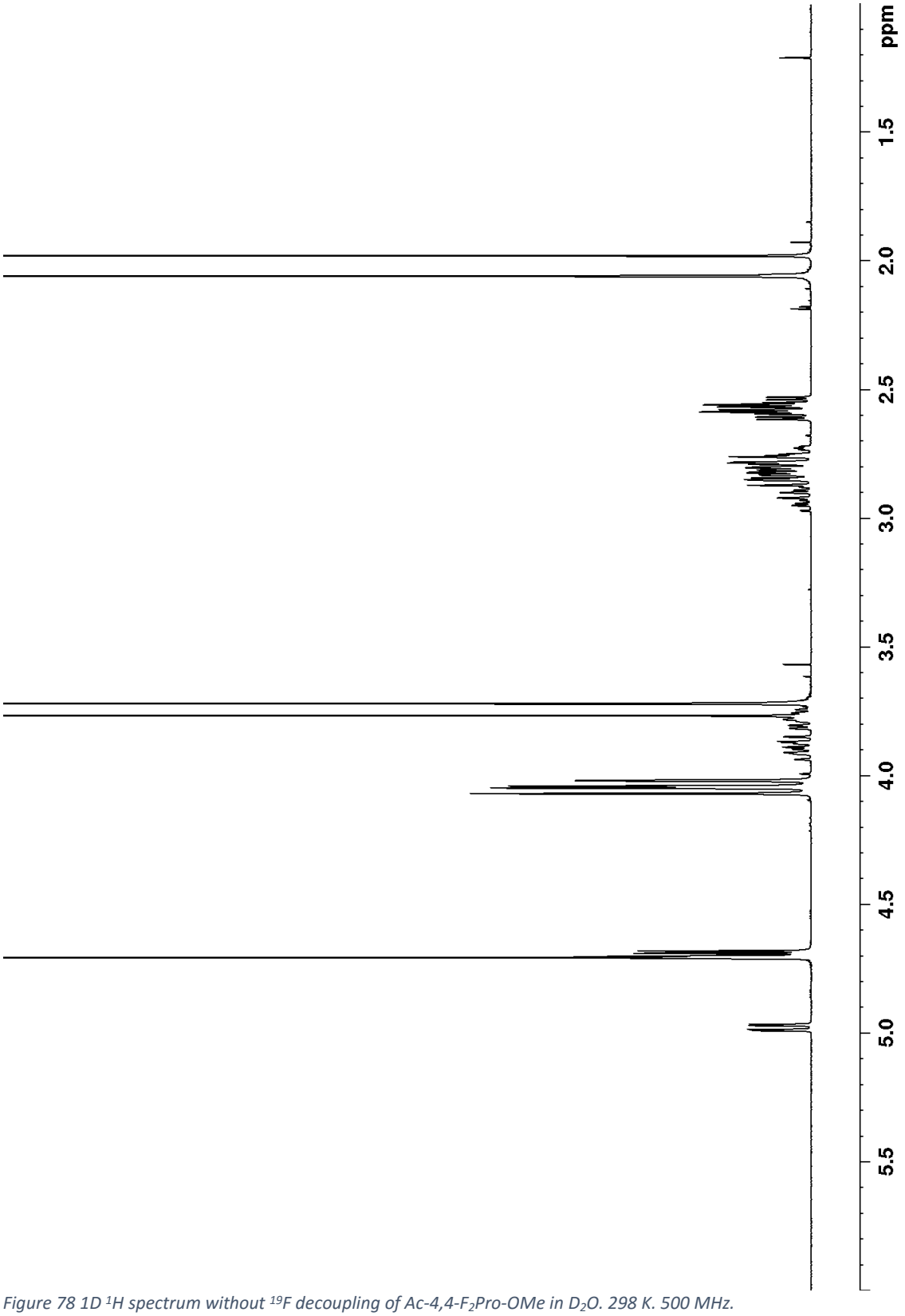


Figure 78 1D ^{1}H spectrum without ^{19}F decoupling of Ac-4,4-F_2Pro-OMe in D_2O. 298 K. 500 MHz.

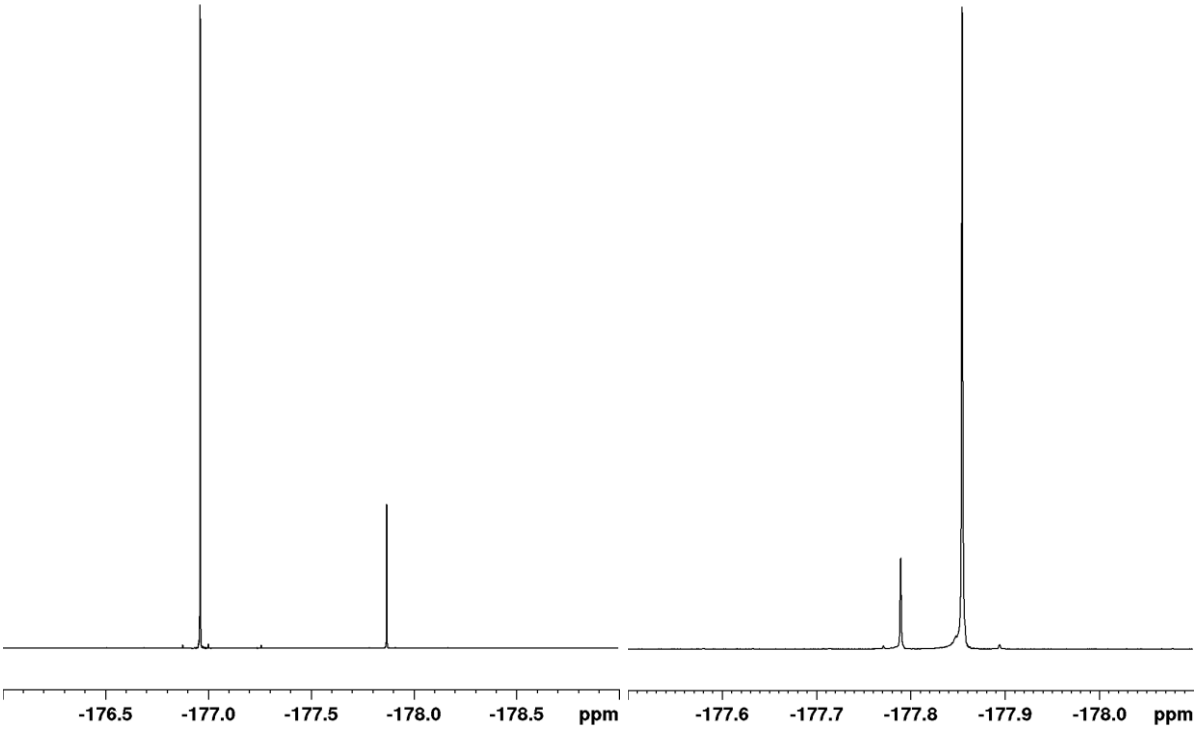


Figure 79 1D ^{19}F spectrum with ^{1}H decoupling of Ac-(4R)-FPro-OMe. Left: $CDCl_3$. Right: D_2O. Both at 298 K and 500 MHz.

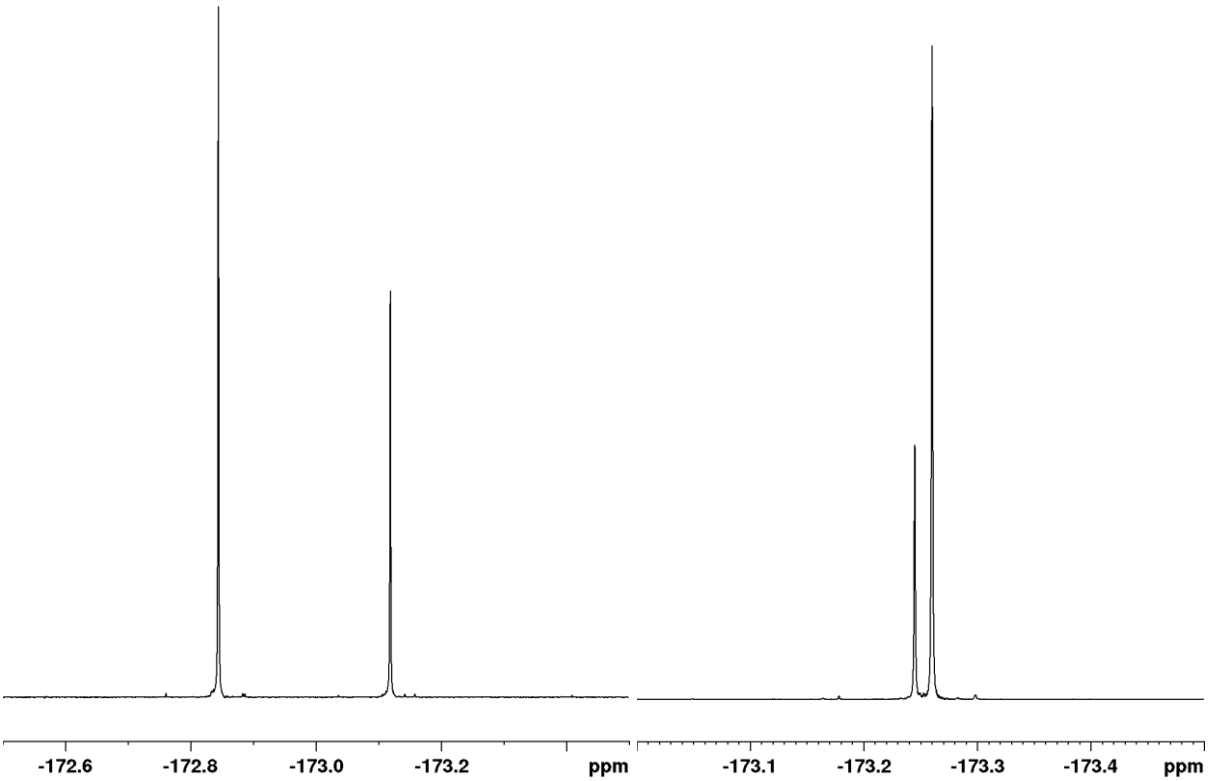


Figure 80 1D ^{19}F spectrum with ^{1}H decoupling of Ac-(4S)-FPro-OMe. Left: $CDCl_3$. Right: D_2O. Both at 298 K and 500 MHz.

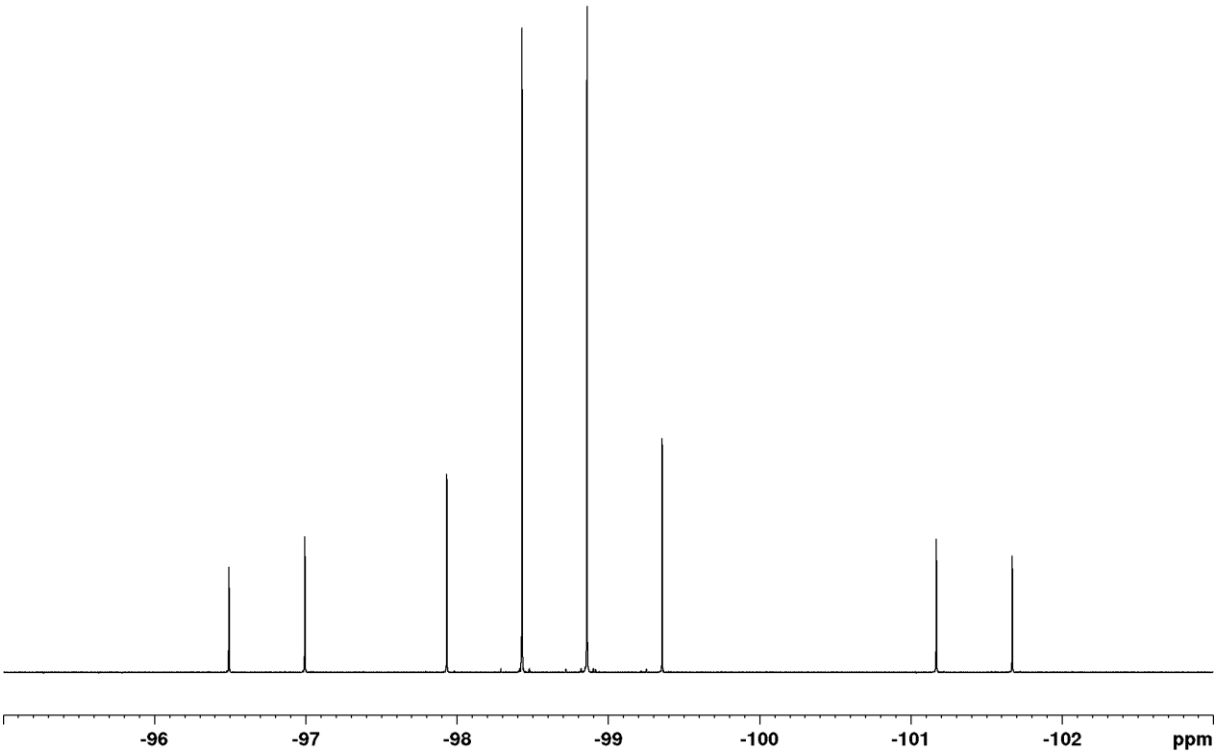


Figure 81 1D ^{19}F spectrum with ^{1}H decoupling of Ac-4,4-F_2Pro-OMe in $CDCl_3$. 298 K. 500 MHz.

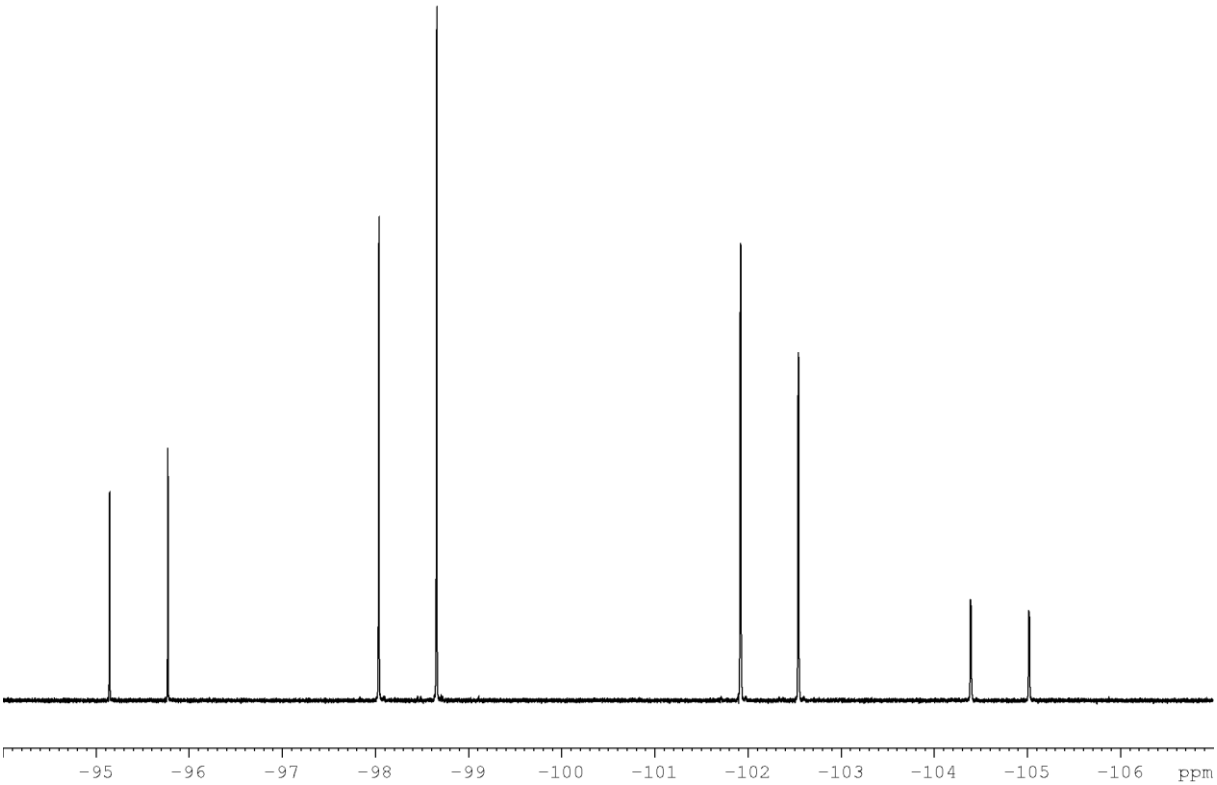


Figure 82 1D ^{19}F spectrum with ^{1}H decoupling of Ac-4,4-F_2Pro-OMe in D_2O. 298 K. 400 MHz.

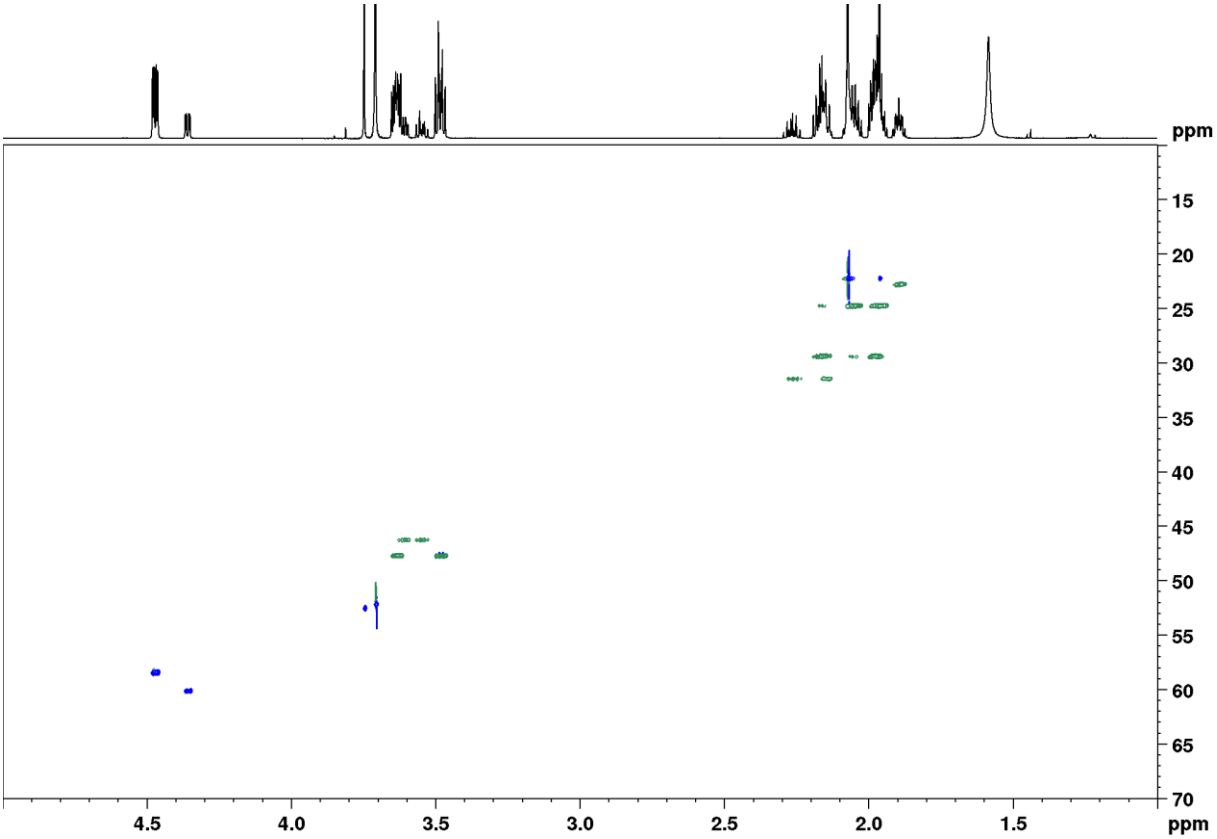


Figure 83 2D ^{1}H-^{13}C HSQC of Ac-Pro-OMe in $CDCl_3$. 298 K. 700 MHz.

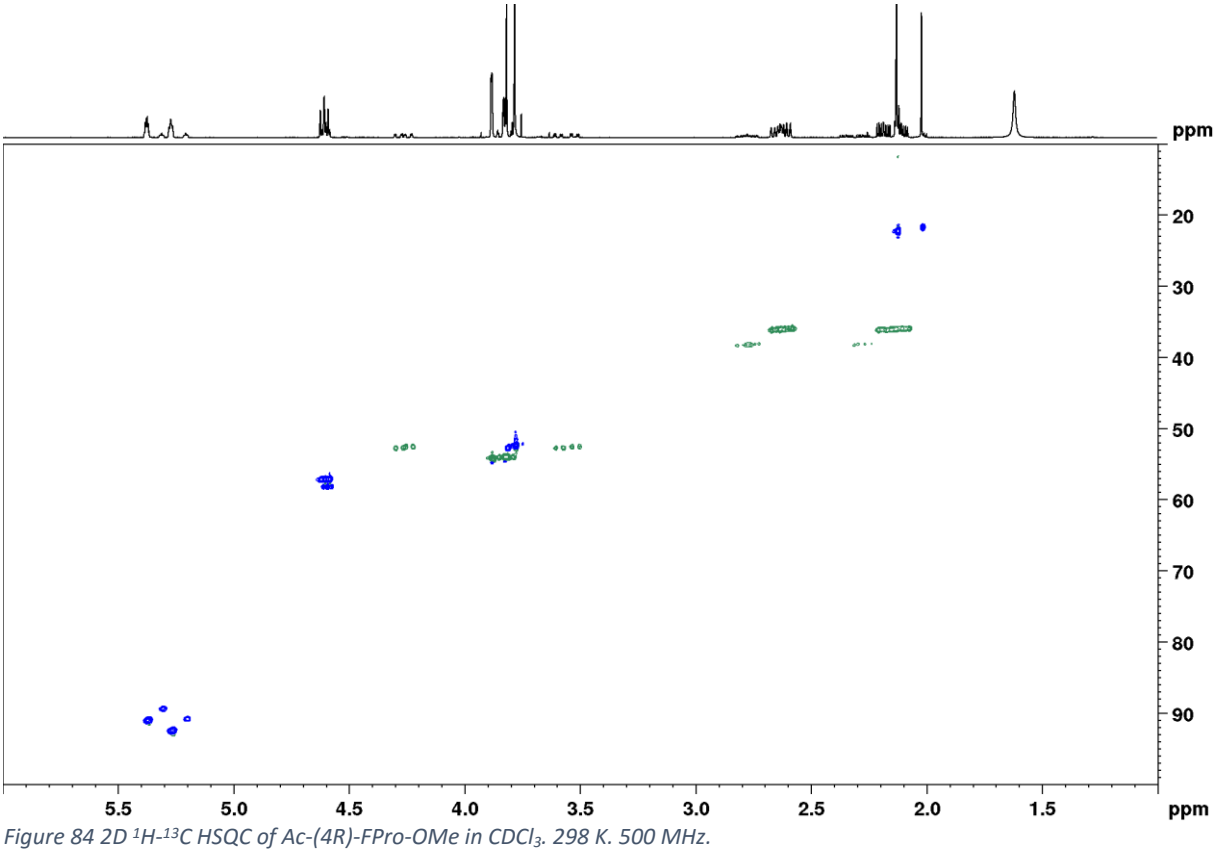


Figure 84 2D ^{1}H-^{13}C HSQC of Ac-(4R)-FPro-OMe in $CDCl_3$. 298 K. 500 MHz.

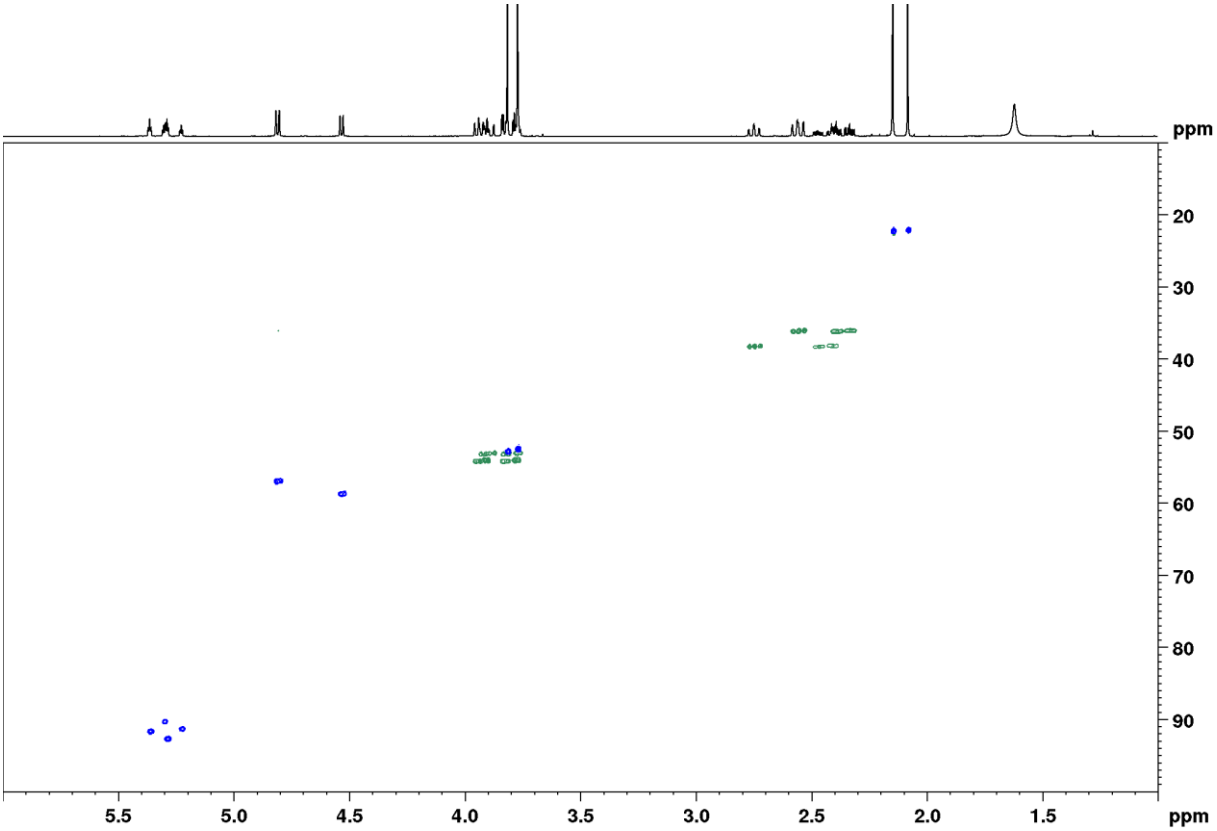


Figure 85 2D ^{1}H-^{13}C HSQC of Ac-(4S)-FPro-OMe in $CDCl_3$. 298 K. 700 MHz.

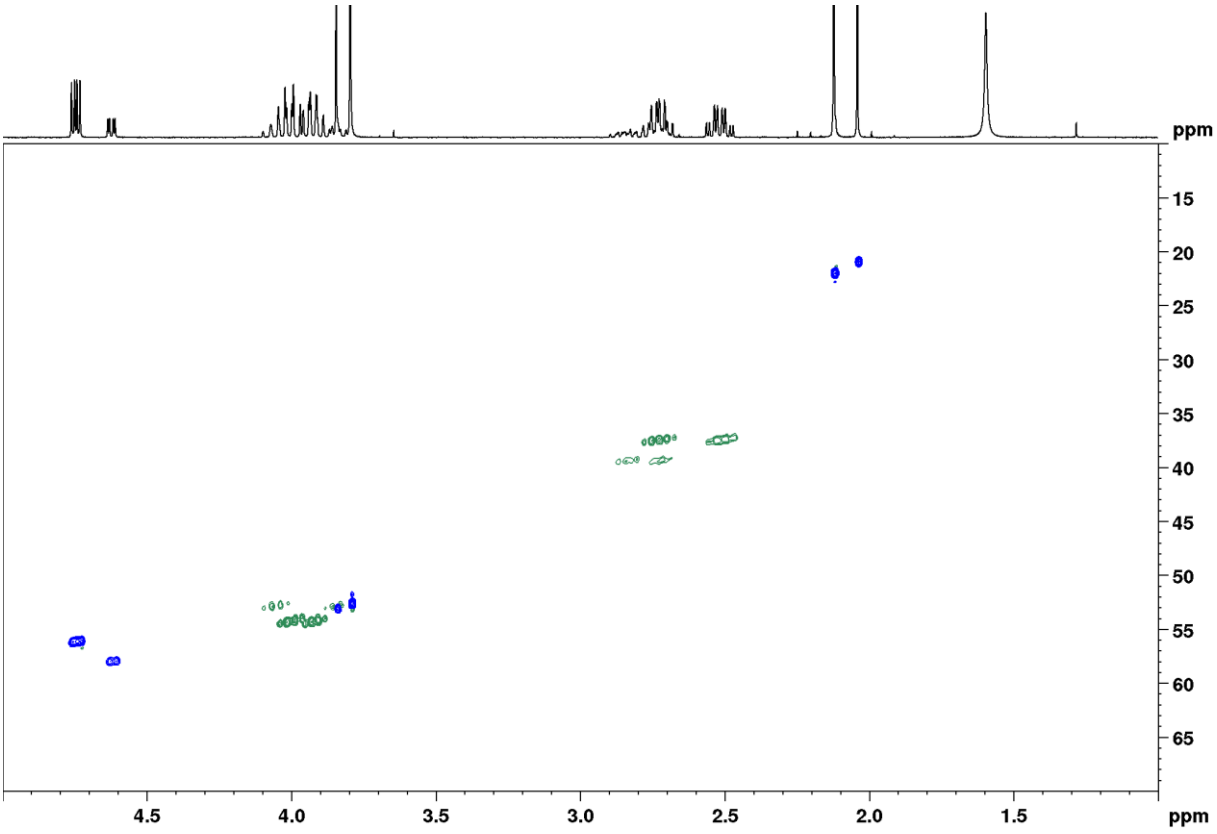


Figure 86 2D ^{1}H-^{13}C HSQC of Ac-4,4-F_2Pro-OMe in $CDCl_3$. 298 K. 500 MHz.

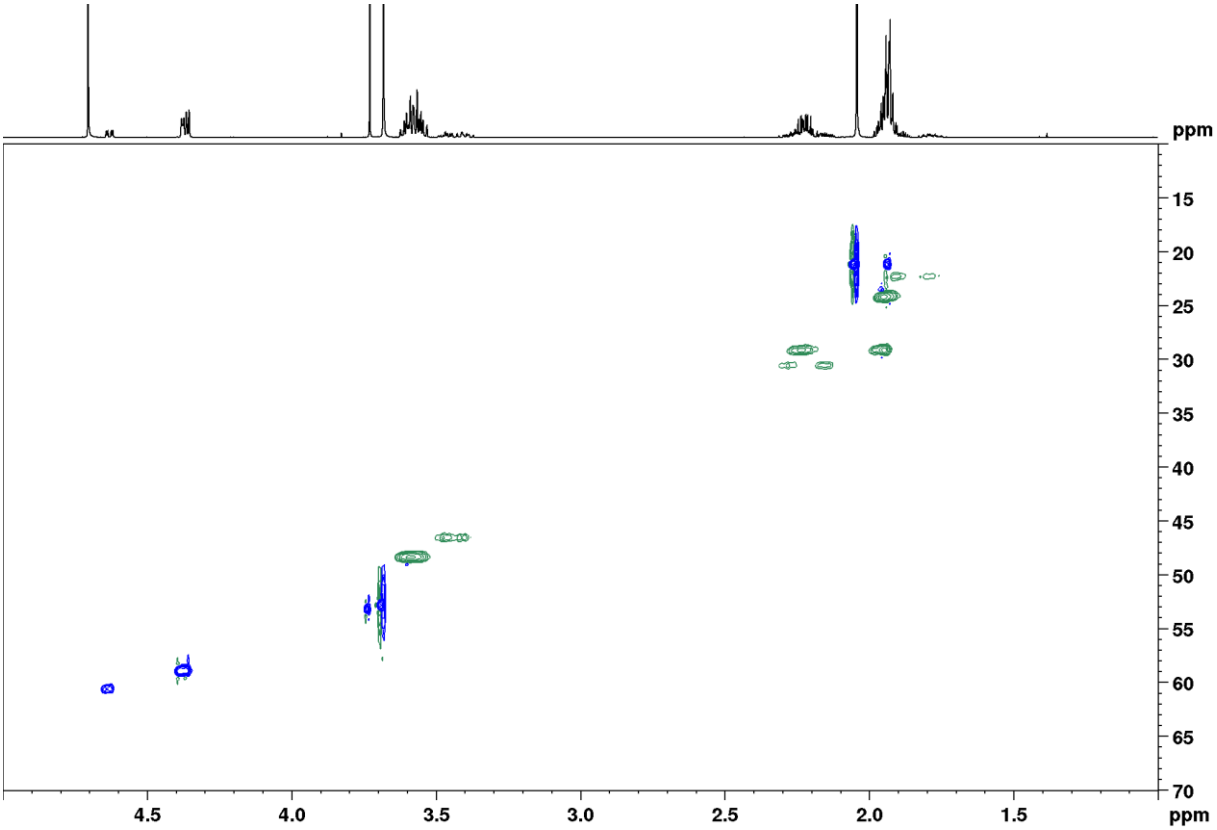


Figure 87 2D ^{1}H-^{13}C HSQC of Ac-Pro-OMe in D_2O. 298 K. 500 MHz.

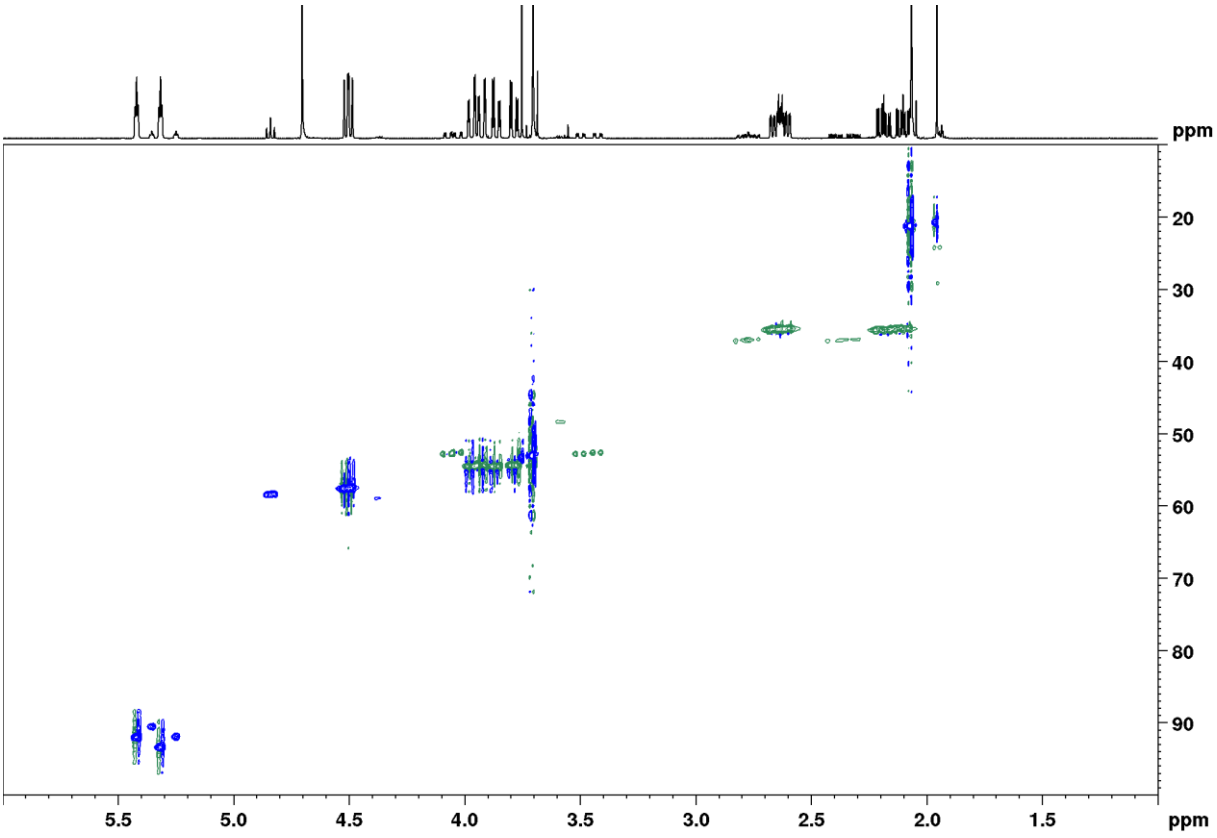


Figure 88 2D ^{1}H-^{13}C HSQC of Ac-(4R)-FPro-OMe in D_2O. 298 K. 500 MHz.

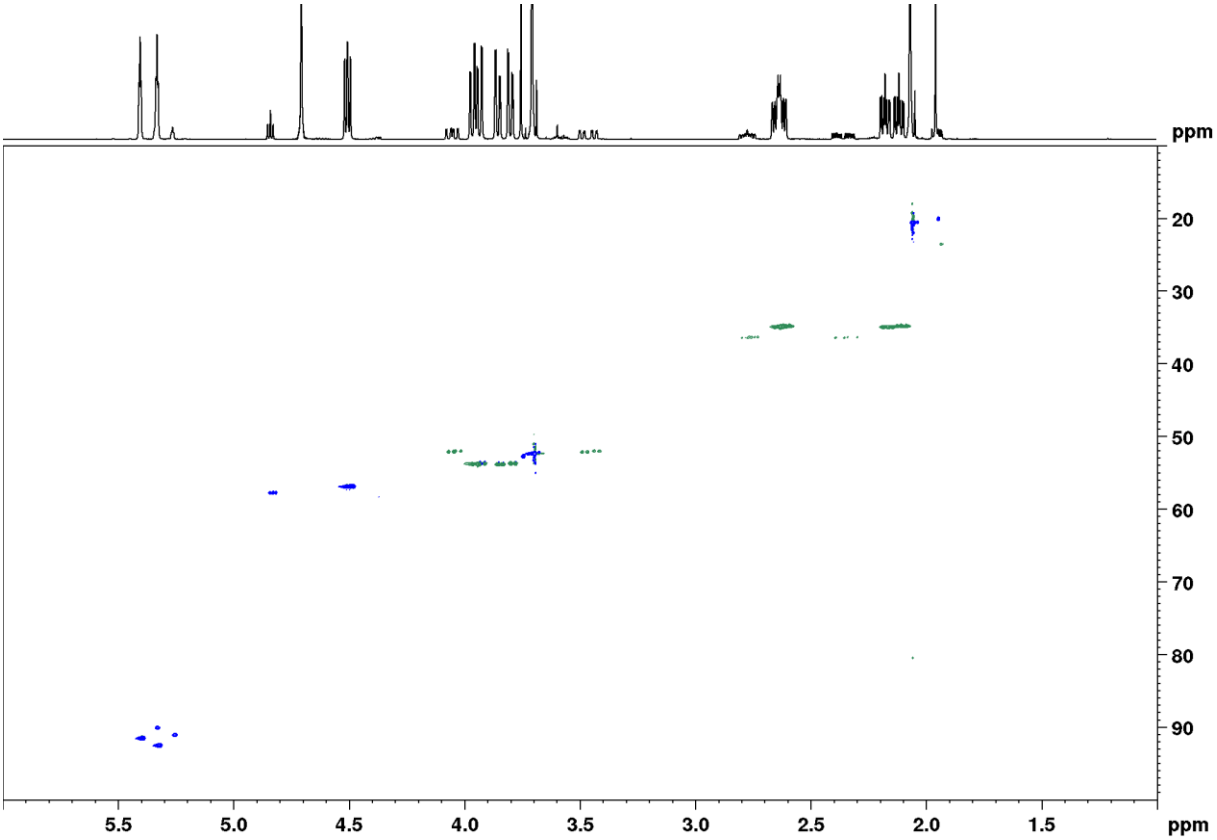


Figure 89 2D ^{1}H-^{13}C HSQC of Ac-(4S)-FPro-OMe in D_2O. 298 K. 700 MHz.

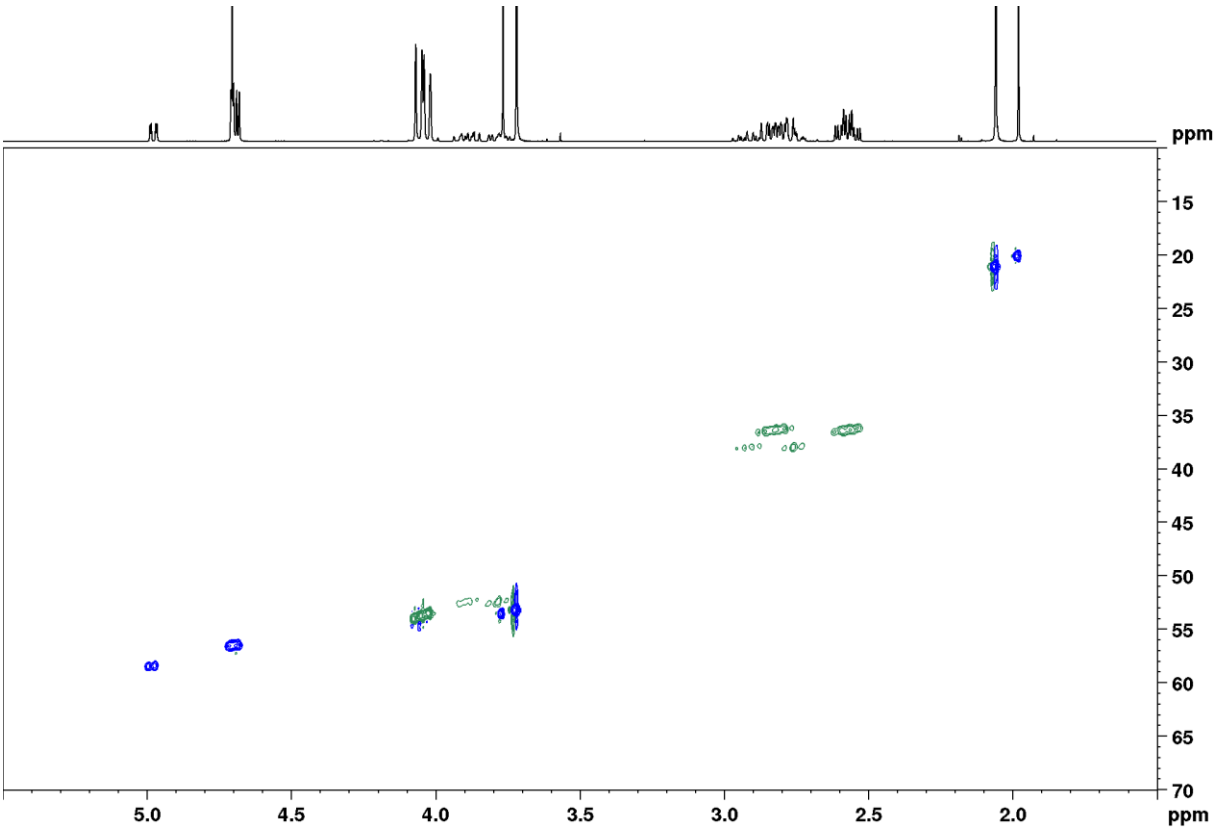


Figure 90 2D ^{1}H-^{13}C HSQC of Ac-4,4-F_2Pro-OMe in D_2O. 298 K. 500 MHz.

7. References

1. Salwiczek, M., et al., *Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions.* Chem Soc Rev, 2012. **41**(6): p. 2135-71.
2. Thomas, K.M., et al., *Proline editing: a divergent strategy for the synthesis of conformationally diverse peptides.* Org Lett, 2005. **7**(12): p. 2397-400.
3. Milner-White, E.J., L.H. Bell, and P.H. Maccallum, *Pyrrolidine ring puckering in cis and trans-proline residues in proteins and polypeptides.* J Mol Biol, 1992. **228**(3): p. 725-34.
4. Renner, C., et al., *Fluoroprolines as tools for protein design and engineering.* Angew Chem Int Ed Engl, 2001. **40**(5): p. 923-5.
5. Bretscher, L.E., et al., *Conformational stability of collagen relies on a stereoelectronic effect.* J Am Chem Soc, 2001. **123**(4): p. 777-8.
6. Newberry, R.W., et al., *Signatures of n-->pi* interactions in proteins.* Protein Sci, 2014. **23**(3): p. 284-8.
7. Newberry, R.W., et al., *n-->pi* interactions of amides and thioamides: implications for protein stability.* J Am Chem Soc, 2013. **135**(21): p. 7843-6.
8. DeRider, M.L., et al., *Collagen stability: insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations.* J Am Chem Soc, 2002. **124**(11): p. 2497-505.
9. Hinderaker, M.P. and R.T. Raines, *An electronic effect on protein structure.* Protein Sci, 2003. **12**(6): p. 1188-94.
10. Adzhubei, A.A., M.J. Sternberg, and A.A. Makarov, *Polyproline-II helix in proteins: structure and function.* J Mol Biol, 2013. **425**(12): p. 2100-32.
11. Horng, J.C. and R.T. Raines, *Stereoelectronic effects on polyproline conformation.* Protein Sci, 2006. **15**(1): p. 74-83.
12. Moradi, M., et al., *Conformations and free energy landscapes of polyproline peptides.* Proc Natl Acad Sci U S A, 2009. **106**(49): p. 20746-51.
13. Kay, B.K., M.P. Williamson, and M. Sudol, *The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains.* FASEB J, 2000. **14**: p. 231-41.
14. Best, R.B., et al., *Effect of flexibility and cis residues in single-molecule FRET studies of polyproline.* Proc Natl Acad Sci U S A, 2007. **104**(48): p. 18964-9.
15. Doose, S., et al., *Probing polyproline structure and dynamics by photoinduced electron transfer provides evidence for deviations from a regular polyproline type II helix.* Proc Natl Acad Sci U S A, 2007. **104**(44): p. 17400-5.
16. Bochicchio, B. and A. Pepe, *Role of polyproline II conformation in human tropoelastin structure.* Chirality, 2011. **23**(9): p. 694-702.
17. O'Hagan, D., *Understanding organofluorine chemistry. An introduction to the C-F bond.* Chem Soc Rev, 2008. **37**(2): p. 308-19.
18. Odar, C., M. Winkler, and B. Wiltschi, *Fluoro amino acids: a rarity in nature, yet a prospect for protein engineering.* Biotechnol J, 2015. **10**(3): p. 427-46.
19. O'Hagan, D. and H. Deng, *Enzymatic fluorination and biotechnological developments of the fluorinase.* Chem Rev, 2015. **115**(2): p. 634-49.
20. Buer, B.C. and E.N. Marsh, *Fluorine: a new element in protein design.* Protein Sci, 2012. **21**(4): p. 453-62.

21. Sharaf, N.G. and A.M. Gronenborn, *(19)F-modified proteins and (19)F-containing ligands as tools in solution NMR studies of protein interactions.* Methods Enzymol, 2015. **565**: p. 67-95.
22. Panasik, N., et al., *Inductive effects on the structure of proline residues.* International Journal of Peptide and Protein Research, 1994. **44**(3): p. 262-269.
23. Hodges, J.A. and R.T. Raines, *Stereoelectronic and steric effects in the collagen triple helix: toward a code for strand association.* J Am Chem Soc, 2005. **127**(45): p. 15923-32.
24. Kim, W., K.I. Hardcastle, and V.P. Conticello, *Fluoroproline flip-flop: regiochemical reversal of a stereoelectronic effect on peptide and protein structures.* Angew Chem Int Ed Engl, 2006. **45**(48): p. 8141-5.
25. Shi, G.-Q. and W.-L. Cai, *d,e-unsaturated b,b-difluoro-a-keto esters: novel synthesis and utility as precursors of b,b-difluoro-a-amino acids.* J Org Chem, 1995. **60**(20): p. 6289-95.
26. Liu, Z., et al., *3-Fluoroazetidinecarboxylic Acids and trans,trans-3,4-Difluoroproline as Peptide Scaffolds: Inhibition of Pancreatic Cancer Cell Growth by a Fluoroazetidine Iminosugar.* J Org Chem, 2015. **80**(9): p. 4244-58.
27. Hommel, U., et al., *Pyrrolidine derivatives and their use as complement pathway modulators*. 2014, Google Patents.
28. Wüthrich, K., *NMR of proteins and nucleic acids*. 1986.
29. Zangger, K., *Pure shift NMR.* Prog Nucl Magn Reson Spectrosc, 2015. **86-87**: p. 1-20.
30. Foroozandeh, M., et al., *Ultrahigh-resolution NMR spectroscopy.* Angew Chem Int Ed Engl, 2014. **53**(27): p. 6990-2.
31. Sinnaeve, D., et al., *A general method for extracting individual coupling constants from crowded 1H NMR spectra.* Angew Chem Int Ed Engl, 2016. **55**: p. 1090-3.
32. Aguilar, J.A., G.A. Morris, and A.M. Kenwright, *"Pure shift" 1H NMR, a robust method for revealing heteronuclear couplings in complex spectra.* RSC Advances, 2014. **4**(16): p. 8278.
33. Boeyens, J.C.A. and D.G. Evans, *Group theory of ring pucker.* Acta Cryst, 1989. **B45**: p. 577-81.
34. Cremer, D. and J.A. Pople, *Molecular orbital theory of the electronic structure of organic compounds. XXIII. Pseudorotation in saturated five-membered ring compounds.* J Am Chem Soc, 1975. **97**(6): p. 1358-67.
35. Altona, C. and M. Sundaralingam, *Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation.* J Am Chem Soc, 1972. **94**(23): p. 8205-12.
36. Hendrickx, P.M.S., *Conformational analysis of flexible molecules*, in *Organic Chemistry*. 2009, Ghent University.
37. Karplus, M., *Contact electron-spin coupling of nuclear magnetic moments.* J Chem Phys, 1959. **30**: p. 11-5.
38. Postova Slavetinska, L., D. Rejman, and R. Pohl, *Pyrrolidine nucleotide analogs with a tunable conformation.* Beilstein J Org Chem, 2014. **10**: p. 1967-80.
39. Diez, E., J. San-Fabian, and J. Guilleme, *Vicinal proton-proton coupling constants I. Formulation of an equation including interactions between substituents.* Molecular Physics, 1989. **68**(1): p. 49-63.

40. Guilleme, J., J. San-Fabian, and E. Diez, *Vicinal proton-proton coupling constants II. Analysis of the effect of interaction between geminal substituents upon vicinal couplings to methyl groups.* Molecular Physics, 1989. **68**(1): p. 65-85.
41. Marshall, J.L., et al., *Reasons for the nonequivalence of the exo-exo and endo-endo vicinal NMR coupling constants in norbornanes.* Tetrahedron, 1976. **32**: p. 537-42.
42. Thibaudeau, C., J. Plavec, and J. Chattopadhyaya, *A new generalized Karplus-type equation relating vicinal proton-fluorine coupling constants to H-C-C-F torsion angles.* J Org Chem, 1998. **63**(15): p. 4967-84.
43. Hendrickx, P.M. and J.C. Martins, *A user-friendly Matlab program and GUI for the pseudorotation analysis of saturated five-membered ring systems based on scalar coupling constants.* Chem Cent J, 2008. **2**: p. 20.
44. Alper, J.S. and R.I. Gelb, *Standard errors and confidence intervals in nonlinear regression: comparison of Monte Carlo and parametric statistics.* J Phys Chem, 1990. **94**(11): p. 4747-51.
45. Hobart, D.B., Jr. and J.S. Merola, *(2S,4R)-4-Fluoro-pyrrolidinium-2-carboxyl-ate.* Acta Crystallogr Sect E Struct Rep Online, 2012. **68**(Pt 8): p. o2490.
46. Siebler, C., et al., *Importance of dipole moments and ambient polarity for the conformation of Xaa–Pro moieties – a combined experimental and theoretical study.* Chem. Sci., 2015. **6**(12): p. 6725-6730.
47. Eberhardt, E.S., et al., *Solvent effects on the energetics of prolyl peptide bond isomerization.* J Am Chem Soc, 1992. **114**(13): p. 5437-9.
48. Eberhardt, E.S., N. Panisik, Jr., and R.T. Raines, *Inductive Effects on the Energetics of Prolyl Peptide Bond Isomerization: Implications for Collagen Folding and Stability.* J Am Chem Soc, 1996. **118**(49): p. 12261-12266.
49. Thomas, C.A., E.R. Talaty, and J.G. Bann, *3S-fluoroproline as a probe to monitor proline isomerization during protein folding by 19F-NMR.* Chem Commun (Camb), 2009(23): p. 3366-8.
50. Jacobsen, N.E., *NMR spectroscopy explained*. 2007.
51. Sinnaeve, D., *Gevorderde structuuranalyse van een recent geïsoleerd natuurproduct met behulp van NMR spectroscopie*. 2005, Ghent University.
52. Cory, D.G. and W.M. Ritchey, *Suppression of signals from the probe in Bloch decay spectra.* J Magn Reson, 1988. **80**: p. 128-32.
53. Thrippleton, M.J. and J. Keeler, *Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra.* Angewandte Chemie, 2003. **115**(33): p. 4068-4071.
54. Kupce, E., J. Boyd, and I.D. Campbell, *Short selective pulses for biochemical applications.* J Magn Reson Series B, 1995. **106**: p. 300-3.
55. Geen, H. and R. Freeman, *Band-selective radiofrequency pulses.* J Magn Reson, 1991. **93**: p. 93-141.