THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Probing peptidylprolyl bond cis/trans status using distal 19F NMR reporters

OPEN ACCESS

RESEARCH ARTICLE

Probing peptidylprolyl bond *cis*/*trans* status using distal ^{19}F NMR reporters

Patrick M. Killoran,[a]‡ George S.M. Hanson,[c]‡ Sanne J. M. Verhoork,[a] Madeleine Smith,[a] Davide Del Gobbo,[a] Lu-Yun Lian,[b] Christopher R. Coxon[c]*

[a]   Dr P. M. Killoran, Dr S. J. M. Verhoork, M. Smith, D. Del Gobbo
        School of Pharmacy and Biomolecular Sciences
        Liverpool John Moores University
        Liverpool, Merseyside, L3 3AF, UK

[b]   Prof. L-Y Lian
        Institute of Systems, Molecular and Integrative Biology
        The University of Liverpool
        Crown Street, Liverpool, L69 7ZB, UK

[c]   G. S. M. Hanson, Dr C. R. Coxon
        EaStChem School of Chemistry
        The University of Edinburgh
        Joseph Black Building, David Brewster Road, Edinburgh, EH14 4AS, UK
        E-mail: chris.coxon@ed.ac.uk

‡   Authors contributed equally and can both list as first author.

        Supporting information for this article is given via a link at the end of the document.

Abstract: A method for measuring peptidylprolyl bond *cis-trans* conformational status in peptide models is described, using 4-fluorophenylalanine (4FPhe) as a distal reporter for ^{19}F NMR. The %*cis*-Pro population was measured for peptides of the general structure Ac-X-Pro-Z-Ala-Ala-4FPhe (X and Z are proteinogenic amino acids) at pH 7.4, and provided conformational populations consistent with literature values obtained by more complex methods. This approach was applied to probe the prolyl bond status in pentapeptide models of the intrinsically disordered C-terminal region of α-synuclein, which mirrored the preferences in the Ac-X-Pro-Z-Ala-4FPhe models. Advantageously, the ^{19}F reporter group does not need to be adjacent to or attached to proline to provide quantifiable signals and distal 4-fluorophenylalanines can be placed so as not to influence prolyl bond conformation. Finally, we demonstrated that the prolyl bond status is not significantly affected by pH when there are ionisable amino acid residues at the carboxyl side of proline, which makes ^{19}F NMR an invaluable tool with which to study proline isomerism at a range of pHs and in different solvents and buffers.

Introduction

The peptidylprolyl bond is a key facilitator of conformational change and peptidylprolyl bond *cis-trans* isomerism is often the slow-phase or rate-limiting step in protein folding, with a ~80 kJ/mol energy barrier for interconversion.[1] Standard backbone amide bonds, comprised of non-cyclic canonical amino acids have a very high thermodynamic preference for a *trans* conformation.[2] However, with its sidechain forming part of the backbone, proline contains a tertiary amide in peptides and proteins, which makes the difference in energy between the *cis* and *trans* forms lower than for standard amino acids, affording a higher population of the *cis* conformation (Figure 1A) at proline compared with non-prolines. These features result in unique properties that are fundamental to the three-dimensional structure of peptides and proteins. Indeed, the prevalence of proline

residues in intrinsically disordered protein motifs (approx. double the average [3]) has prompted further investigation into the role of prolines in adopting partially folded states and promoting disorder or compaction.[4] Indeed, the peptidylprolyl *cis-trans* isomerases CypD and FK506 binding proteins (FKBPs) have been shown to interact directly with the acidic proline-rich C-terminal region of α-synuclein (αSyn), which forms a constituent part of the protein aggregates in Lewy bodies in a range of synucleinopathies e.g. Parkinsons Disease.[5, 6] FKBPs were shown to accelerate the rate of aggregation, and aggregation was abolished by removal of the proline residues from the C-terminus. However, CypD binding to soluble αSyn prevented its aggregation, whilst the addition of CypD to preformed αSyn fibrils led to fibril disassembly. In both cases, this behaviour was linked to an increased rate of isomerase-catalysed peptidylprolyl isomerisation. Moreover, Pro-to-Ala mutations at C-terminal region are reported to accelerate the rate of uncatalysed aggregation,[7] highlighting a key role of proline conformation in Parkinson's Disease pathology.
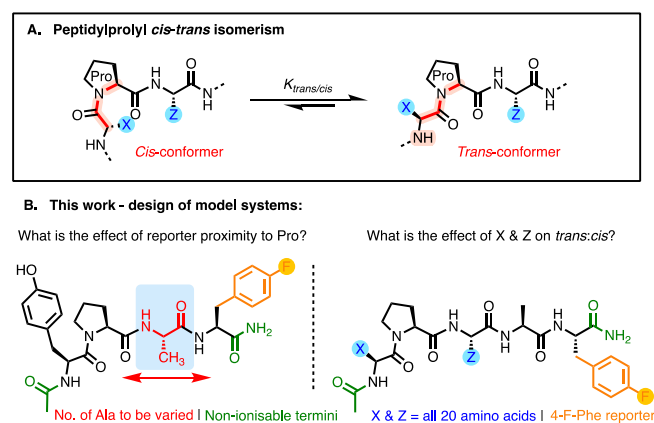

Figure 1. Peptidylprolyl *cis-trans* isomerism and the design of model systems containing fluorine reporter groups for ^{19}F NMR to measure prolyl bond status.

The natural populations of specific prolyl bond conformers is affected by i) the presence of electron-donating/withdrawing or bulky groups at the proline 3- or 4-positions, which controls the pyrrolidine *endo/exo* ring pucker through steric and/or stereoelectronic effects,[8] and ii) the nature of neighbouring amino acid sidechains, and that these trends are consistent with the observed *cis*-populations in fully folded proteins.[9-13]

Most of the studies described have used NMR-based approaches to detect and/or quantify prolyl conformers. NMR techniques provide structural information at the molecular level and can be useful in understanding the finer details of protein and peptide folding at specific sites, particularly with respect to proline isomerism.[14,15] Due to the relatively slow exchange rates between *cis* and *trans* conformers, ^{1}H NMR analysis of X-Pro systems typically affords two sets of signals, allowing the measurement of *cis:trans*.[9-12, 16] However, this can create complexity in the analysis and in many cases, 2D NMR approaches are required to resolve specific resonances. Compared with the use of ^{1}H NMR for studies of peptide conformation, ^{19}F NMR exhibits very little-to-no spectral crowding and will detect separate signals for the different forms that are in chemical exchange. The fluorine-19 isotope (^{19}F) exhibits higher local environmental sensitivity compared with the ^{1}H nucleus, providing defined and distinct chemical shifts for individual fluorinated molecules, isomers or conformers.[17] In addition, it is only slightly less sensitive compared with the ^{1}H nucleus. Therefore, ^{19}F NMR analysis of fluorinated peptides enables far simpler identification and quantification of conformer populations and over a wider range of timescales as compared with traditional methods.

The use of a variety of fluorinated amino acids in conjunction with ^{19}F NMR is a valuable tool to study biological processes and several applications have been reported.[18-21] Prolyl bond conformational status in proteins has been measured previously using a combination of fluorinated prolines and ^{19}F NMR. In particular, a number of groups have used fluorinated amino acids derivatives to study *cis-trans* isomerism by ^{19}F NMR. Thomas and colleagues studied 3*S* and 3*R*-fluoroproline measured the kinetic and thermodynamic parameters of proline *cis-trans* isomerisation,[22] whilst Renner performed a similar analysis of 4*S*, 4*R* and 4,4-difluoroprolines in Ac-(F)Pro-OMe using ^{19}F NMR.[23] It was clear from this work (and others) that the 4*R*-fluoroproline has a greater preference for the *trans*-Pro conformation, whilst the 4*S*-diastereomer has a relatively increased *cis*-Pro population compared to the parent proline. These studies highlighted the connection between the fluorine configuration and *cis:trans* conformer populations, albeit these observations were only in only in simple model systems. Kawahara was able to incorporate 4*R*-fluoroproline into collagen polypeptides and study triple helix assembly and the relationship to proline conformation using ^{19}F NMR.[24] Torbeev demonstrated that 4,4-difluoroproline, which has nearly the same *cis-trans* preference as proline but an increased rate of isomerisation, caused pronounced destabilization of macroglobulin; indicating that the *cis-trans* equilibrium and isomerisation dynamics of a single proline affect the rate of protein folding.[25] Increasing the number of equivalent fluorine atoms enhances the signal-to-noise ratio in ^{19}F NMR analysis. Trifluoromethylproline was incorporated into a simple Ac-TfmPro-OMe model and used with ^{19}F NMR to provide an accurate reporter of *cis:trans* and to determine the $K_{trans/cis}$.[26] Whilst Tressler an co-workers developed perfluoro-*tert*-butyl 4-

hydroxyproline (9 fluorine atoms) as a highly-sensitive ^{19}F NMR reporter for proline isomerism, permitting quantification at concentrations as low as 200 nM.[27] However, these studies also highlighted that fluorine atoms exert an influence on the conformation of prolines, and fluoroproline, is therefore, a biased proline conformational reporter.[8] 4-Fluorophenylalanine has been used to study the formation of oligomers in the misfolding of prion proteins using ^{19}F NMR.[28] Multiple 4-fluorophenylalanine residues have also been incorporated into a single protein, rat intestinal fatty acid binding protein (IFABP), to monitor the folding process and providing information on the conformational status at various locations simultaneously.[29] 4-Fluorophenylalanine was also used as a reporter of the enzyme catalysed *cis*-to-*trans* isomerisation rate of [4-fluoro-Phe]bradykinin measured by magnetisation-transfer ^{19}F NMR.[30] However, most of the reported examples of using 4-fluorophenylalanine as a ^{19}F NMR reporter in peptide and protein systems have focused on its applications in probing global conformational effects and little attention has been paid to its use for probing prolyl bond status. Here we examine the use of 4-fluorophenylalanine-as a reporter to measure the relative *cis:trans* prolyl bond conformer ratio in X-Pro-Z systems contained in pentapeptide models. Moreover, we explore how the proximity of this reporter group and the nature of proximal amino acids impacts upon peptidylprolyl *cis:trans* populations (Figure 1B).
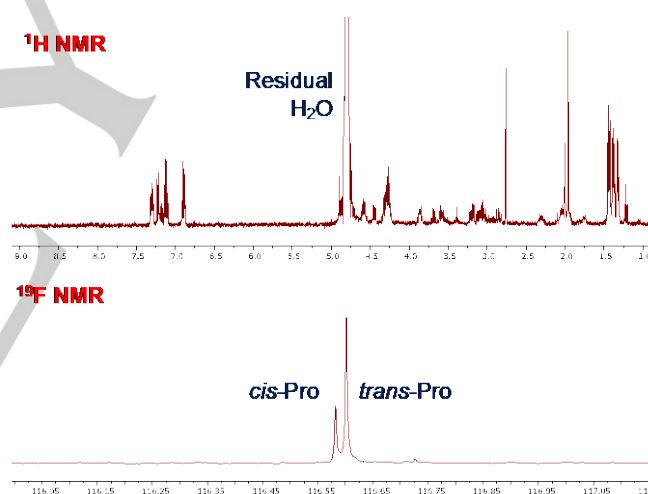

Figure 2. Comparaison of an example ^{1}H NMR (top) and ^{19}F NMR (bottom) spectrum of peptide Ac-Tyr-Pro-Ala-Ala-Ala-4FPhe. Experiments were acquired at 2 mg/mL in D_2O, room temperature.

Results and Discussion

Employing 4-fluorophenylalanine as a reporter group, the initial aim was to understand how the proximity of the 4-fluorophenylalanine residue to the proline, affected the relative *cis:trans* ratio and how far away this could be moved within a peptide sequence and still provide distinguishable ^{19}F NMR resonances for each conformer. The Tyr-Pro motif exhibits a relatively high *cis*-Pro population of around 20-39% in, for example, Ac-Ala-**Tyr-Pro**-Ala-Lys-NH2,[9] Ac-Gly-**Tyr-Pro**-Gly[12] and Ac-Thr-**Tyr-Pro**-Asn,[13] depending on other proximal residues.[9,12,13] Therefore, this was used as a starting point for model design and optimisation, and a series of peptides containing Ac-Tyr-Pro-(Ala)$_n$-4FPhe (where n = 0 – 7) were

synthesised using standard microwave-assisted solid phase peptide synthesis, sequentially moving the 4-fluorophenylalanine reporter one residue at a time away from proline. Peptides were prepared as C-terminal amides and were N-terminally capped as an acetamide to avoid ionisable termini influencing the *cis:trans* through electrostatic interactions with one another or ionisable sidechains, as noted in other reports.[31,32]
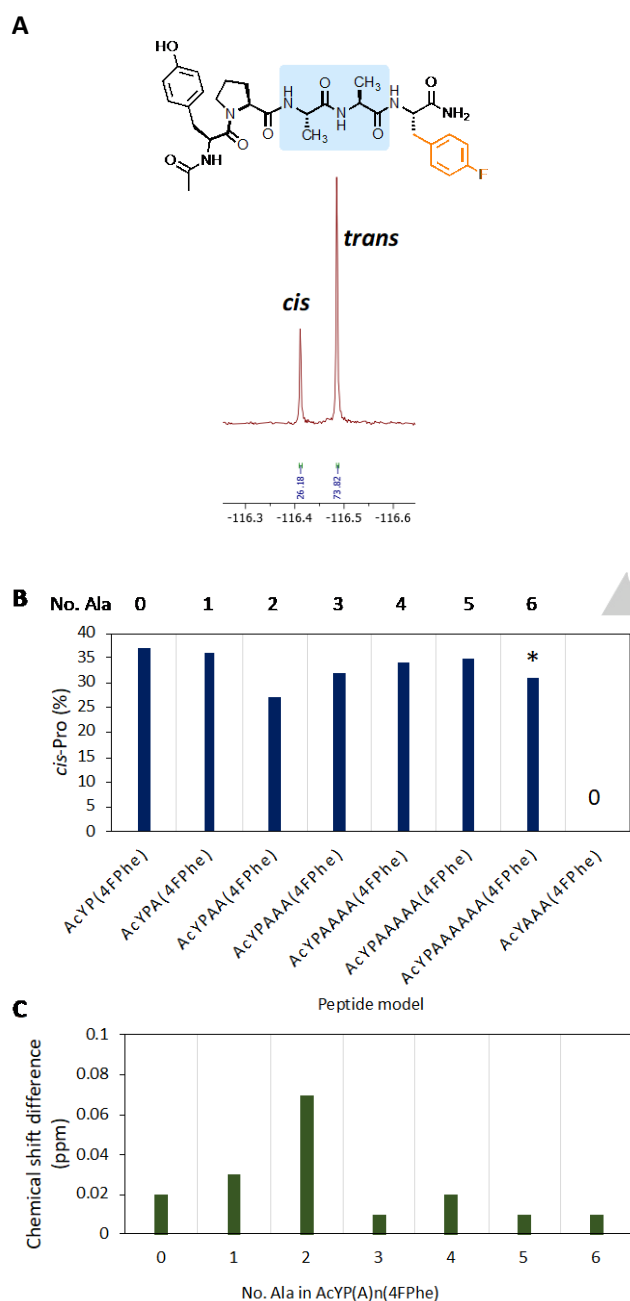
A

B

C

Figure 3. A) An example ^{19}F NMR spectrum of peptide Ac-Tyr-Pro-Ala-Ala-(4FPhe) (2 alanine spacer) exhibiting prolyl bond *cis-trans* conformers. Probing the effect of proximity of the ^{19}F NMR reporter (4FPhe = 4-fluorophenylalanine) on **B)** measured *cis*-Pro populations in model peptides using ^{19}F NMR, and **C)** chemical shift difference between resonances. * Poorly resolved: not accurately quantifiable but still distinguishable ^{19}F NMR spectrum.

The first point to note was the relative simplicity of the ^{19}F NMR spectrum compared with the ^{1}H NMR spectrum for the same model peptide (Figure 2), which significantly deconvolutes the conformational analysis. An advantage of the ^{19}F NMR method is that it functions readily at physiological pH (7.4), whereas ^{1}H NMR approaches often quantify *cis*-Pro in the amide region, which disappears due to exchange above ~ pH 7 or in D_2O (absent in Figure 2). ^{19}F NMR can also report *cis*-Pro populations in any non-fluorinated solvent or buffer without need for water-suppression techniques. Peptide prolyl bond *cis:trans* was analysed by ^{19}F NMR (relaxation delay time (D1) set at 10 s; with [1]H-decoupling) in (1 mg/mL in aqueous 10 mM phosphate buffered saline (10% D_2O), pH 7.4) at 23 °C and afforded two ^{19}F resonances with the major being assigned as the thermodynamically favoured *trans* and minor as *cis* conformer (e.g. Figure 3A). Pleasingly, the control peptide containing no proline, only afforded a single resonance. The signal integrations were consistent with the reported range of *cis*-Pro populations in related systems as mentioned earlier and were largely similar for each probe (Figure 3B), confirming that 4-fluorophenylalanine can report on *cis-trans* conformer populations over a range of amino acids (up to >6 residues away). Models containing six and seven alanine spacer residues exhibited distinct signals but were poorly resolved and not quantifiable. Slightly lower *cis*-Pro populations were noted for cases when the 4-fluorophenylalanine was two (27%) or three (32%) residues away from proline, whilst the model containing 4FPhe adjacent to Pro exhibited a slightly greater *cis*-Pro population (37%). In some cases, aromatic (Aro) amino acids in an Aro-Pro-Aro sequence can increase the propensity for a *cis* prolyl bond conformation but this effect may be less significant when the aromatic residue is electron poor e.g., 4-fluorophenylalanine.[13] Despite the Ala-Ala spacer model appearing to slightly under-report the %*cis* conformer (27%), we later found that this provided a good match with the literature conformer populations (discussed subsequently).

The signal resolution between *cis*-Pro and *trans*-Pro resonances is an important criterion for modelling proline isomerism by NMR. By ^{19}F NMR analysis the chemical shift difference for conformers varied with relative proximity of the reporter to the proline (Figure 3C; also see supporting information, Figure S1). Although it was possible to quantify *cis-trans* prolyl populations up to at least 6 amino acids away from proline, the optimal signal resolution (0.07 ppm) was obtained in the model containing two alanine residues between the 4-fluorophenylalanine and proline (AcYPAA-4FPhe, Figure 3A, C). Therefore, this was used as the basis for our the model system to probe the effects of amino acids X and Z in X-Pro-Z systems.

To demonstrate that this model system could be used to study the influence of amino acids proximal to proline on *cis:trans* ratio, 39 peptides of the form Ac-**X**-Pro-Ala-Ala-4FPhe and Ac-Tyr-Pro-**Z**-Ala-4FPhe were synthesised by systematically replacing the amino acid **X** or **Z** immediately adjacent to proline. As before, peptidylprolyl *cis:trans* conformer ratios were determined by integration of characteristic resonances obtained from ^{19}F NMR (386 MHz, 23 °C, 1 mg/mL aqueous 10 mM phosphate buffered saline (10% D_2O), pH 7.4) analysis (Figure S2). As expected, the relative ratio of signal integration varied according to the amino acid at the proline nitrogen (Figure 4, Table S1) - indicating that this was reporting the *cis:trans* ratio. Some pairs of signals were poorly resolved despite the model optimisation and required line-fitting analysis using the Mnova

line-shape tool. Chemical shifts also were observed to change slightly, and indeed there was a weak correlation for higher *cis*-Pro preferences to also exhibit greater signal resolution in ^{19}F NMR spectra, perhaps due to slower rates of isomerism.

Comparison of the *cis*-Pro populations measured by ^{19}F NMR with those obtained by ^{1}H NMR in related models published by Reimier *et al* [9] and Dyson *et al*, [10] respectively, afforded good agreement (Figure 4A, B and Table S2). The ^{19}F NMR measurements are also consistent with the earlier reports of local conformational effects and the role of neighbouring amino acids upon proline *cis-trans* conformational preferences, albeit in slightly different models. [9,10] In general, the X-Pro motif exhibits the greatest influence over the conformation of proline, particularly when X is an aromatic amino acid: Phe (22.5%), Tyr (26.5%) and Trp (34.5%). This was further enhanced in Tyr-Pro-Aro motifs (Aro = aromatic amino acids: Phe (38%), Tyr (38.5%) and Trp (35.5%)) exhibiting the highest propensity for a *cis* prolyl bond. This also supports the proposal that the Aro-Pro-Aro motif is a key stabiliser of type VI beta turns.[33-35] These observations are attributed to favourable CH-π interactions between relatively acidic proline H_α or H_δ atoms and electron-rich aromatic rings; and aromatic rings interacting with both faces of proline, respectively. [9, 12, 13, 34] However, in both types of model His was found to provide a lower *cis*-Pro population than other aromatic amino acids (His-Pro-Ala (14%) and Tyr-Pro-His (31%), behaving more like a basic sidechain. Indeed, there was a clear trend for electron-rich aromatics to promote higher *cis*-Pro in the order: Trp-Pro > Tyr-Pro > Phe-Pro > His-Pro.[9,13, 36] It was also clear that polar side chains e.g. Ser (12%), Thr (13%) Asp (13.5%) promote a slightly higher *cis*-Pro population in X-Pro systems compared with Ala-Pro (10.5%), while Arg (8%) or Lys (9%) in this position leads to a subtle decrease in the *cis*-Pro population. These results corroborate the observations from Sebak and co-workers in their statistical analysis of *cis*-Pro populations in intrinsically disordered proteins.[15] Another interesting observation was the relatively high %*cis* for the Tyr-Pro-Gly (31%) and Tyr-Pro-Pro (39%) peptides, which are approximately equivalent to the Aro-Pro-Aro *cis*-populations.

A small number of discrepancies were noted between the reported literature *cis*-Pro populations and those measured using ^{19}F NMR. Notably, in our hands when basic amino acids were proximal to proline in Tyr-Pro-Z models we observed significantly higher *cis*-Pro populations (Z = Arg 29%, Lys 29.5%, His 31%) compared with Tyr-Pro-Ala (26.5%) (Figure 4B).[37] Our systems appeared to exhibit higher %*cis* than the corresponding models by Dyson *et al*,[10] wherein the basic amino acids exhibited rather low *cis*-Pro (vs 16, 17 and 13%, respectively,) compared with Tyr-Pro-Ala. However, our data are mostly well aligned with Reimer *et al* for Ac-X-Pro models (X = basic amino acid), who reported *cis*-Pro populations of X = (7.2%), Lys (6.8%) and His (16.5%) at pH 6 but noted that the histidine peptide Ac-Ala-His-Pro-Ala-Lys-NH₂ exhibited a lower *cis*-Pro population (9.5%) at pH 8.

Given that the literature experiments were conducted at different pHs – often to observe non-proline amide NHs for ^{1}H NMR analysis, we were concerned that basic-Pro or Pro-basic motifs may be sensitive to ionisation status or intramolecular electrostatic interactions. To try to understand this better, we resynthesised exemplar peptides related to those from Dyson *et al* but containing 4-fluorophenylalanine (AcYPZD(4FPhe) where Z = ionisable residues D, H, K or R) and measured the %*cis*-Pro at two different pHs (4.1 and 7.4) using ^{19}F NMR.
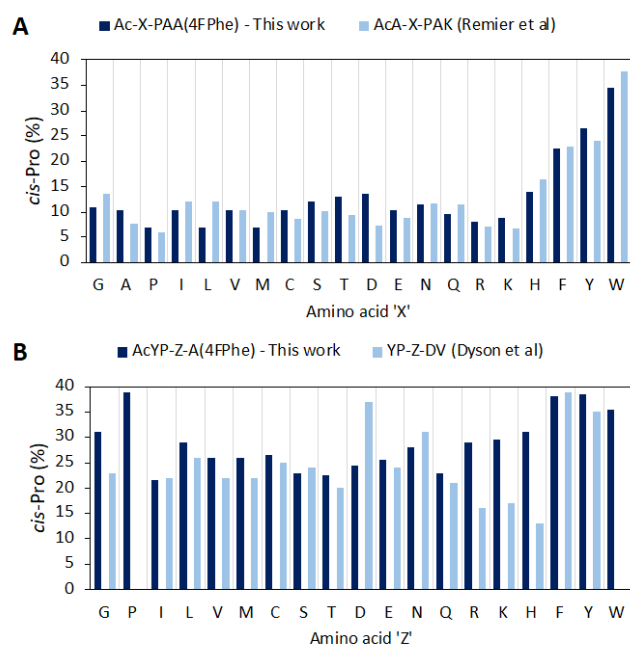

Figure 4. Comparison of ^{19}F NMR derived *cis*-Pro conformational preference (dark blue) with earlier published data (light blue) obtained by ^{1}H NMR analysis. **A**) Comparison with model Ac-Ala-X-Pro-Ala-Lys-NH₂ (Reimer *et al*)[9] at pH 6 where 'X' is at the prolyl amine. **B**) Comparison with model H₂N-Tyr-Pro-Z-Asp-Val-COOH (Dyson *et al*)[10] at pH 4.1 where 'Z' is at the proline carboxamide. Note: blank columns indicate no possible comparison within the literature.

In our analyses, the *cis*-Pro populations were observed to align closely with our earlier models, dominated by the Tyr-Pro motif (Figure 5), and importantly, the prolyl conformer populations were found to be largely insensitive to modest pH changes. This underlines that the use of ^{19}F NMR at pH 7.4 accurately reflects what others have seen at acidic pH. This is a key advantage over ^{1}H NMR methods that are difficult to evaluate at physiological pH. Therefore, the differences between this work and the Dyson work may be related to previous models containing a second adjacent ionisable group (i.e. Asp), which may provide a repulsive interaction. We also demonstrated that the %*cis* was unaffected by changes in salt (NaCl) concentration up to 1.2 M in the buffer solution (see supporting information, Figures S9-12).

We briefly investigated whether the effect of temperature on *cis*-Pro population could be studied using ^{19}F NMR, by heating NMR samples of peptide models Ac-LPAA(4F-Phe)-CONH₂ and Ac-YPGA(4F-Phe)-CONH₂ in pH 7.4 buffer between 10 ºC – 60 ºC in 10 ºC steps. In simple model prolines, this is shown to lead to an increase in the % *cis*-Pro,[23] whereas in small peptides, this was suggested to be a largely insignificant effect.[38] In our hands, resonances relating to *cis* and *trans*-Pro AcLPAA(4FPhe) was found to appear to coalesce towards a single signal with increasing temperature and correspondingly, the % *cis*-Pro after line shaping analysis appeared to slightly increase above 30 ºC (7% at 10 ºC to 12% at 60 ºC). Conversely, the peptide AcYPGA(4FPhe) seemed to trend towards marginally lower % *cis*-Pro (26% at 10 ºC to 22% at 60 ºC) (see supporting information, Figure S7, Table S3). Albeit, in both cases, this was a minor change, this further emphasises that ^{19}F NMR

measurements can perhaps provide additional insight into subtle effects that are challenging to observe by ^{1}H NMR.

α-Synuclein is a generally disordered protein, but can assume stable misfolded protein oligomers and aggregates, which is a hallmark of Parkinson's disease pathophysiology. To demonstrate that the observed prolyl bond populations are relatively conserved in different peptides and are governed mainly by local sequences, we prepared a series of pentapeptide models of the proline-rich (5 proline residues in full length protein) intrinsically disordered C-terminal region of α-synuclein (Figure 6, Table 1). These models contained disorder-promoting amino acids e.g. Q, E, D, N rather than alanines between the proline and ^{19}F reporter. It was clear from these experiments, that the %*cis*-Pro in each model was similar to that observed in our earlier models and the observations of Sebak *et al* in intrinsically-disordered proteins. This indicates that prolyl bond conformation is controlled predominantly by the proline *N*-acyl substituent in these models and that %*cis*-Pro may be generally predictable within an unstructured polypeptide. Importantly, this also highlights that the ^{19}F NMR approaches are broadly useful and sensitive to small changes in conformer populations.
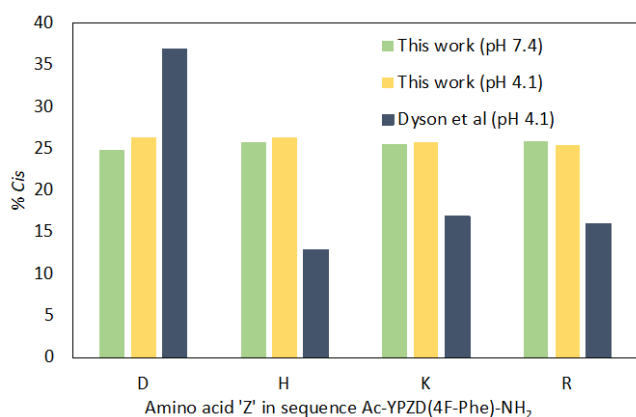

Figure 5. Comparison of ^{19}F NMR derived *cis*-Pro conformer populations for Dyson *et al* [10] models, and the resythesised versions containing 4-fluorophenylalanine, at pH 4.1 and 7.4.
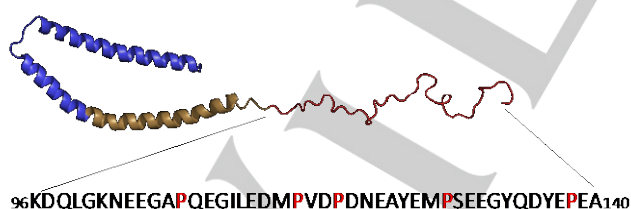

Figure 6. The intrinsically disordered C-terminal region of α-synuclein (PDB: 1XQ8), which consists of amino acids 96 to 140 (proline residues in red). Tetramer sequences, indicated with bold lines, informed the design of the 5 pentapeptide models used in this study.

Table 1. Pentapeptide models of the proline-rich intrinsically disordered C-terminal region of α-synuclein containing 4-fluorophenylalanine. % *cis*-Pro measured by ^{19}F NMR at pH 7.4.

Entry	Sequence	%*cis*-Pro, pH 7.4
1	Ac-APQE(4FPhe)-NH$_2$	8
2	Ac-MPVD(4FPhe)-NH$_2$	12
3	Ac-DPDN(4FPhe)-NH$_2$	10
4	Ac-MPSE(4FPhe)-NH$_2$	6
5	Ac-EPEA(4FPhe)-NH$_2$	11

Conclusion

In summary, the prolyl bond conformational trends reported herein using ^{19}F NMR are broadly similar to the previous observations using different techniques, and importantly, align with the general trends for prolyl bond conformation in folded protein structures – indicating that prolyl bond conformation is predominantly governed by local effects. Furthermore, very similar prolyl bond conformational populations were observed to occur in short models of the intrinsically disordered C-terminal region of α-synuclein, indicating that these prolyl bond populations are conserved in flexible peptides and proteins.

This work has demonstrated that incorporating fluorinated reporter groups *via* 4-fluorophenylalanine is a reliable way to quantify prolyl bond status in peptide models. If positioned carefully, it does not influence the prolyl bond status and advantageously, does not need to be adjacent or attached to the proline to provide quantifiable and distinguishable signals, providing flexibility in model design. Furthermore, Fmoc-L-4-fluorophenylalanine is commercially available at low cost, which makes this approach useful for a wide range of researchers. The simplicity of the ^{19}F NMR read-out compared to the corresponding ^{1}H NMR spectrum and the ability to quantify *cis*-Pro populations at a range of temperatures, physiological pH and in a variety of solvents and buffers without solvent-suppression and in complex mixtures such as cell lysate, will be of broad and important use in biological chemistry. Specifically, this will be utilised in our future studies as a straightforward tool to understand the consequences of point mutations or the role of specific residues upon *cis-trans* isomerism at specific proline residues in larger, more complex and intrinsically disordered and proline-rich systems.

Experimental Section

Materials

Synthesis of model peptides

Fmoc solid-phase peptide synthesis was performed using automated Fmoc-SPPS on a Liberty Blue microwave-assisted peptide synthesiser (CEM). SPPS was conducted using Rink amide ProTide resin (0.56 mmol/g, 0.1 mmol, 179 mg). Fmoc protected amino acids were used in 5 equivalents (0.2 M in DMF) with DIC (1 M stock solution in DMF; 10 equiv.) and Oxyma Pure (1 M stock solution, 5 equiv) as activators. Standard coupling procedures employed single coupling of each amino acid (2.5 min, 90 °C). Amino acids bearing thermally-sensitive protecting groups, for example, Fmoc-L-His(Trt)-OH and Fmoc-L-Cys(Trt)-OH, were coupled under milder conditions (50 °C for 10 min). Fmoc deprotection was performed by treatment of the resin with piperidine (20% *v/v* in DMF, 4 mL). Following the removal of the final Fmoc-protecting group, peptides were acetylated (where relevant) using acetic anhydride (20% in DMF) at 37 °C for 15 min and repeated once, followed by washing the resin with DMF three times and drying under vacuum. The resin was washed several times with Et$_2$O (5 mL) and cleaved from the resin using a cocktail of TFA/TIPS/H$_2$O (95/2.5/2.5) for 2 h at 37 °C or 4 h at RT in a disposable fritted syringe. Afterwards, the suspension was filtered and precipitated in ice-cold Et$_2$O. The resulting suspension was centrifuged (5 min at 5000 rpm) and the pellet was resuspended with Et$_2$O (10 mL). This was centrifuged again (5 min 5000 rpm) and the Et$_2$O was discarded. Crude peptide mixtures were purified using RP-HPLC (Agilent 1260), equipped with a Waters Xbridge peptide BEH C18 Prep column (19 × 100 mm, 5 μm, 130 Å) using a linear gradient from 5–95 % MeOH in water containing 0.1 % TFA over 20 minutes at a flow rate of 8 mL/min.

Characterisation of model peptides

High-resolution mass spectrometry method A: performed using an Agilent 6530 accurate mass QToF system. Separation was performed on an Agilent ZORBAX Eclipse Plus C18 Rapid Resolution HD analytical column C18 column (1.8 μm particle size, 2.1 × 50 mm) at 25 °C. The mobile phase of the system consisted of water (with 0.1% TFA, eluent A) and Methanol (with 0.1% TFA, eluent B). the 12.5 min gradient was programmed as follows: 0-1 min 95% A, 1.0-7.0 min to 0% A, 7.0-9.0 min hold 0%, 9.0-9.5 min to 95% A, 9.5-12.5 min hold 95% A. The injection volume was set to 2 μL and the flow rate to 0.5 mL/min, respectively. Monitoring was performed at λ = 215 nm. Mass spectral analysis of peptides was performed in the positive ionisation (ESI+) mode in full scan mode (*m/z* 100-3200). Operating pressures were in the range of 2000-3000 PSI. Electrospray ionisation mass spectrometry was conducted in positive ion mode (*m/z* range: 50 – 3200) using a fragmentor voltage of 150 V, gas temperature of 325 °C (flow 10 L/min) and sheath gas temperature of 400 °C (flow 11 L/min). Exact mass measurements of the products were based on the protonated ions [M+H]+ and were detected as sodiated adducts [M+Na]+.

Analytical HPLC method B (retention times[b]): Separation was performed on a Waters XBridge C18 analytical column (3.5 μm particle size, 4.6 x 50 mm) at 40 °C. The mobile phase of the system consisted of water (with 0.1% TFA, eluent A) and Methanol (with 0.1% TFA, eluent B). The 15 min gradient was programmed as follows: 0-10 min 95% A to 20% A, 10-11 min to 5% A, 11-12 min hold 5% A, 12-13 min to 95% A, 13-15 min hold 95% A. The injection volume was set to 2 μL and the flow rate to 1 mL/min, respectively. Monitoring was performed at λ = 215 nm.

^{19}F NMR analysis of amide bond conformation

General procedure: Purified peptides (1 mg/mL) were dissolved in 1x PBS solution (pH 7.4, 10 mM by phosphate, 10% D$_2$O). The NMR experiments were conducted using a Bruker 400 MHz NMR spectrometer. The experiments employed were 1D ^{19}F NMR: 386 MHz, proton decoupled, 128 scans, D1 = 10 sec, 23 ˚C. The NMR data were processed using TopSpin (v4.0.5) or MNova (v14.2.1-27684).

pH Experiment Procedure: Purified peptides (2 mg/mL) were dissolved in 10% D$_2$O/H$_2$O. The pH was adjusted to 4.1 with HCl, in accordance with the Dyson method,[1,2] or a combination of HCl and NaOH for pH 7.4. The NMR experiments were the same as the general procedure other than being completed in triplicate.

Variable Temperature Experiment Procedure: Purified peptides (2 mg/mL) were dissolved in 10% D$_2$O/H$_2$O. The NMR acquisition parameters were the same as the general procedure other than being conducted over a range of temperatures from 10 ºC to 60 ºC over 10 ºC steps.

Acknowledgements

[1] a) C. Dugave, L. Demange, *Chem. Rev.* **2003**, *103*, 2475−2532; b) G. Fischer, F. X. Schmid, *Biochemistry* **1990**, *29*, 2205−2212; c) F. X. Schmid, L. M. Mayr, M. Mücke, E. R. Schönbrunner, *Adv. Protein Chem.* **1993**, *44*, 25−66.

[2] W. J. Wedemeyer, E. Welker, H. A. Scheraga, *Biochemistry* **2002**, *41*, 14637−14644.

[3] V.N. Uversky, *Intrinsically Disord. Proteins* **2013**, *1*, e24684.

[4] B. Mateos, C. Conrad-Billroth, M. Schiavina, A. Beier, G. Kontaxis, R. Konrat, I.C. Felli, R. Pierattelli, *J. Mol. Biol.* **2020**, *432*, 3093-3111.

[5] M. Gerard, Z. Debyser, L. Desender, P. J. Kahle, J. Baert, V. Baekelandt, Y. Engelborghs, *FASEB J.* **2006**, *20*, 524-526.

[6] J. Torpey, J. Madine, A. Wood, LY Lian. *Sci Rep.* **2020**, *10*, 10159.

[7] J. Meuvis, M. Gerard, L. Desender, V. Baekelandt, Y. Engelborghs, Biochemistry, **2010**, *49*, 9345-9352.

[8] S. J. Verhoork, P. M. Killoran, C. R. Coxon, *Biochemistry* **2018**, 57, 6132-6143.

[9] U. Reimer, G. Scherer, M. Drewello, S. Kruber, M. Schutkowski, G. Fischer, *J. Mol. Biol.* **1998**, *279*, 449-460.

[10] H. J. Dyson, M. Rance, R. A. Houghten, R. A. Lerner, P. E. Wright, *J. Mol. Biol.* **1988**, *201*, 161-200.

[11] C. Grathwohl, K. Wüthrich, *Biopolymers* **1981**, *20*, 2623–2633.

[12] W. J. Wu, D. P. Raleigh, *Biopolymers* **1998**, *45*, 381-394.

[13] K. M. Thomas, D. Naduthambi, N. J. Zondlo, *J. Am. Chem. Soc.* **2006**, *128*, 2216–2217

[14] P. A. Evans, C. M. Dobson, R. A. Kautz, G. Hatfull, R. O. Fox, *Nature* **1987**, *329*, 266-268.

[15] F. Sebák, P. Ecsédi, W. Bermel, B. Luy, L. Nyitray, A. Bodor, *Angew. Chem. Int. Ed.* **2022**, *61*, e202108361

[16] I. R. Kleckner, M. P. Foster, *Biochim. Biophys. Acta Proteins Proteom.* **2011**,*1814*, 942-968.

[17] H. Chen, S. Viel, F. Ziarelli, L. Peng, *Chem. Soc. Rev.* **2013**, *42*, 7971-7982.

[18] W. C. Pomerantz, N. Wang, A. K. Lipinski, R. Wang, T. Cierpicki, A. K. Mapp, *ACS Chem. Biol.* **2012**, *7*, 1345-1350.

[19] F. Evanics I. Bezsonova J. Marsh J. L. Kitevski, J. D. Forman-Kay, R. S. Prosser, *Biochemistry* **2006**, *45*, 14120-14128.

[20] Y. Suzuki, J. R. Brender, K. Hartman, A. Ramamoorthy, E. N. G. Marsh, *Biochemistry* **2012**, *51*, 8154-8162.

[21] G. F. Wang, C. Li, G. J. Pielak, *Protein Sci.* **2012**, *19*, 1686-1691.

[22] C. A. Thomas, E. R., Talaty, J. G. Bann, *Chem. Commun.* **2009**, 3366−3368.

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

[24] K. Kawahara, N. Nemoto, D. Motooka, Y. Nishi, M. Doi, S. Uchiyama, T. Nakazawa, Y. Nishiuchi, T. Yoshida, T. Ohkubo, Y. Kobayashi, *J. Phys. Chem. B* **2012**, *116*, 6908–6915

[25] V. Y. Torbeev, D. Hilvert, *Proc. Natl. Acad. Sci.* **2013**, *110*, 20051−20056.

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

[27] C. M. Tressler, N. J. Zondlo, *J. Org. Chem.* **2014**, *79*, 5880−5886.

[28] S. T. Larda, K. Simonetti, M. S. Al-Abdul-Wahid, S. Sharpe, R. S. Prosser, *J. Am. Chem. Soc.* **2013**, *135*, 10533-10541.

[29] H. Li, and C. Frieden, *Proc. Natl. Acad. Sci.* **2007**, *104*, 11993-11998.

[30] R. E. London, D. G. Davis, R. J. Vavrek, J. M. Stewart, R. E. Handschumacher, *Biochemistry* **1990**, *29*, 10298-10302.

[31] C. Grathwohl, K. Wüthrich, *Biopolymers* **1976**, *15*, 2025-2041.

[32] C. Grathwohl, K. Wüthrich, *Biopolymers* **1976**, *15*, 2043-2057.

[33] J. Yao, H. J. Dyson, P. E. Wright, *J. Mol. Biol.* **1994**, *243*, 754-766.

[34] J. Yao, V. A. Feher, B. F. Esperjo, M. T. Reymond, P. E. Wright, H. J. Dyson, *J. Mol. Biol.* **1994**, *243*, 736–753

[35] M. Brandl, M. S. Weiss, A. Jabs, J. Suhnel, R. Hilgenfeld *J. Mol. Biol.* **2001**, *307*, 357–377.

[36] S. K. Gupta, S. Banerjee, E. N., Prabhakaran, *Bioorg. Med. Chem. Lett.*, **2022**. *76*, 128985.

[37] In Dyson's model system (ref 10) (Ser-X-Pro-Gly-Asp-Val), basic amino acids afforded approximately 100% of the *trans* conformer (at pH 4.1), whilst Reimer *et al* reported a lower than normal but still significant *cis*-Pro population (Arg (7.2%), Lys (6.8%) and His (16.5%)) at pH 6 but noted that the histidine peptide Ac-Ala-His-Pro-Ala-Lys-NH$_2$ exhibited a lower % *cis*-Pro (9.5%) at pH 8.

[38] D. P. Raleigh, P. A. Evans, M. Pitkeathly, C. M. Dobson, *J. Mol. Biol.* **1992**, *228*, 338-342.

Entry for the Table of Contents
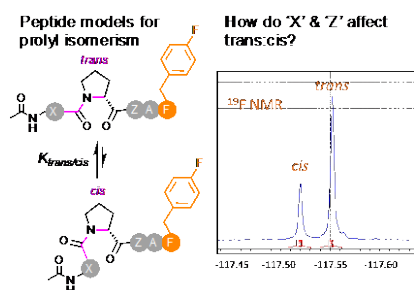
We demonstrate that incorporation of 4-fluorophenylalanine in model peptides provides an accurate and non-biased reporter of peptidylprolyl *cis-trans* conformational status by ^{19}F NMR at physiological pH, even when several residues away from proline. In simple models, the prolyl bond conformational populations are predominantly governed by the amino acids local to proline and are conserved across different model systems.

Institute and/or researcher Twitter usernames: @Dr_CR_Coxon (Dr Coxon); @EdinburghChem (School of Chemistry)