The influence of commonly used tags on structural propensities and internal dynamics of peptides

Maria Bräuer¹ · Maria Theresia Zich¹ · Kamil Önder² · Norbert Müller^{1,3}

Received: 11 January 2019 / Accepted: 19 February 2019 / Published online: 29 April 2019
© The Author(s) 2019

Abstract

The influence of biotin and fluorescein tags attached to the N-terminus of peptides on their structural propensities is assessed by NMR methods. While the small peptides investigated are highly mobile with no uniquely preferred conformation, the introduction of the tags, in particular hydrophobic ones, clearly shows an influence on NMR parameters such as chemical shifts and relaxation properties, which are not restricted to the nearby residues, but also affect distant parts. Thus, long-range effects on structural propensities become evident and are cause for concern with respect to the interpretation of weak interaction tests, which rely upon the assumption that tags do not exert influence on intermolecular interactions.

Graphical abstract

Keywords Peptides · NMR spectroscopy · Conformation · Fluorescence tags · Affinity tags

Introduction

The covalent attachment of tag groups to peptides and proteins for either purification, immobilization, or spectroscopic detection is a widely applied practice in biomolecular research. For example, affinity tags are included in

cloned sequences to facilitate purification of overexpressed recombinant proteins or to improve solubility or to facilitate folding of the target protein [1, 2]. Tags are also employed in Western blotting, enzyme-linked immunosorbent assay (ELISA), or other biological assays [3]. Probably, the most common application of tags is the labelling of biomolecules with fluorescent dyes, for analytical purposes based on fluorophore-based techniques including confocal microscopy, fluorescence resonance energy transfer (FRET), luminescence resonance energy transfer (LRET), or microscale thermophoresis (MST) [4–7].

It is often tacitly assumed that such tags have no or negligible influence on the secondary and tertiary structures of the fused peptides or proteins as well as on their conformational dynamics. Especially small tags, like hexahistidine, FLAG, Strep II, or chitin-binding protein (CBP), are assumed to have no or only slight impact on protein function and, therefore, often remain attached at the N- or

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00706-019-02401-x>) contains supplementary material, which is available to authorized users.

✉ Norbert Müller
norbert.mueller@jku.at

¹ Institut für Organische Chemie, Johannes Kepler Universität Linz, Altenbergerstraße 69, 4040 Linz, Austria

² ProComCure BioTech GesmbH, 5000 Anif, Austria

³ Faculty of Science, University of South Bohemia, Branišovská 31, 37005 České Budějovice, Czech Republic

C-termini of proteins, even during structure determination studies. Larger affinity tags including glutathione-*S*-transferase (GST) and maltose-binding protein (MBP), although increasing solubility [8] need to be removed before structure determination using X-ray or NMR and have also been found to modify intermolecular interactions significantly [9]. While affinity tags introduced by recombinant DNA methods can be removed, e.g., by breaking linker sequence encoded between the tag and the target protein using specific proteases [3], tags introduced via chemical derivatization, like most fluorescence tags, are normally not removed. Although small tags are claimed to have minimal or no impact on the secondary structure of bio-macromolecules, there are some reports of opposite experiences using fluorescent as well as affinity tags like fluorescein, biotin [10–13], or hexa-histidine [14–17]. Most reports about tags affecting structures or functions can be found for hexa-histidine tagged bio-macromolecules, probably because it is the most frequently used tag for the purification of recombinant proteins [18]. The influence of short affinity tags like hexa-histidine and FLAG on crystallisation [14, 19] and on the dynamics in the picosecond range of His-tagged myoglobin used in infrared vibrational echo spectroscopy [15] are prominent examples of such reports.

Conversely, not many reports for effects of biotinylation on structure and function could be found in the literature. It is frequently used to attach peptides to solid surfaces as, for example, in surface plasmon resonance [20–22]. Biotin labelling at a specific lysine residue of human lysozyme did not show any alteration of structure or effects on the binding behaviour [10]. Only a slight decrease in enzymatic activity of the biotinylated sample was observed and explained by the steric hindrance of the biotin moiety preventing binding in spatial proximity to the introduced tag. Another recently published report about a structure modification owed to the biotinylation of aminoacyl-tRNA describes altered growth behaviour at higher temperatures (37 °C), lower heat stability, and less efficiency than the unlabelled counterpart [11].

Fluorescein is frequently used as fluorophore in fluorescence microscopy, immunofluorescence assays, and flow cytometry or fluorescence polarization [23]. Fluorescent tags apparently can impair the functionality of DNA mismatch repair proteins when attached at the C-terminus [24]. Studies on MreB, the actin homologue in bacteria, which is essential for cell wall synthesis, showed that an alpha helix reported in an earlier study was actually induced by a fluorescent tag [25].

For fluorescence polarization assays, especially in the context for high-throughput screening, it has been suggested to determine a competitive binding curve between the labelled probe and the corresponding unlabelled molecule [23]. If the binding curve fits, indicating similar affinity and competitiveness, it can be assumed that the tag of the labelled

compound does not influence the binding behaviour. This procedure was used in the development of a high-throughput fluorescence polarization anisotropy assay. Although NMR experiments had earlier confirmed that labelled and unlabelled peptides were binding in the same manner, they also revealed small changes in the chemical shifts, differential broadening of several peaks, and a slightly higher binding affinity of the fluorescein isothiocyanate (FITC)-labelled peptide derivative [26]. As reasons for the changes, the authors suggested additional interactions of the protein with either the hydrophilic fluorescein moiety or the aliphatic aminohexanoic (Ahx) linker. The introduction of a spacer between the tag and the first amino acid of the peptide sequence is usually considered sufficient to prevent the influences of the bulky fluorophore on conformation. Short linker sequences can also prevent side reactions upon the acidic cleavage of peptides produced by solid-phase peptide synthesis from the resin, as the cleavage reaction using trifluoroacetic acid (TFA) can lead to reactions following the Edman degradation pathway [27], resulting in the formation of fluorescein thiohydantoin including a cleavage of the first amino acid [28].

During our studies of antimicrobial, conformationally labile peptides, some significant differences in the NMR spectroscopic properties between tagged and native peptides were observed. Here, we report experimental results concerning a dodecapeptide attached to an FITC tag and two 22 amino acid peptides attached to biotin. For all peptides, NMR experiments with and without a tag were performed, and their secondary structure propensities as well as local order parameters were compared on the basis of chemical shifts, scalar coupling constants, and relaxation parameters.

In spite of the lower number of signals, the interpretation of peptide NMR data is often less straightforward than for structured proteins. The problem with structure determination of peptides using solution state NMR is the high flexibility of the molecules leading to the co-existence of a plethora of different conformers in solution. Therefore, the experimental NMR parameters represent weighted averages of several conformers rather than a small number of well-defined global folds. The structural propensities are highly dependent on the environment of the peptide; in polar solvents the contributions of intramolecular hydrogen bonds are minor. Owing to the high flexibility of small peptides, crystallisation is accompanied by conformational selection, which is obscuring or biasing the structural information relevant in solution. Solution NMR is considered a most effective technique for the assessment of peptide conformation and conformational dynamics [29, 30].

The peptides investigated are part of a larger quest for antibacterial activity [31], they are: A0 having the sequence RKKRLKLLKRL, with flA0 bearing a fluorescein tag

(with an Ahx spacer) at the N-terminus; S2 having the amino acid sequence RIQALYSKQPKLVERILTKNEQ; and P1 having the sequence DSLDIAELVMELEDEFGTEIPD, with bioS2 and bioP1 respectively being biotinylated at the N-terminus.

Results and discussion

As a prerequisite for this study, the NMR chemical shift of the peptides and their derivatives had to be assigned. This could be achieved to an extent of 84–100%, as summarized in Table 1.

Better than 95% complete assignments could be achieved for the backbone signals of all peptides. For the side chains, all assignments were better than 79%, as summarized in Table 1. For peptide A0, backbone H_N and N signals of the first amino acid were not visible, as well as sidechain residues of arginine and lysine. The backbone signals of peptide fIA0 could be completely assigned, only some overlapping side chain signal assignments of arginine and lysine residues are missing. For peptides S2 and bioS2, resonance again assignments of Arg and Lys residues are incomplete due to overlap, and for peptide S2, the nitrogen resonance of the first amino acid was not detectable. In the assignment of peptide P1, resonances of all atoms of the first amino acid are missing, but bioP1 could be completely assigned. All assignments have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) under the submission numbers 27105, 27114, 27116, 27115, 27118, and 27117.

Secondary chemical shifts are a reliable measure for secondary structure in well-structured proteins. In peptides, they are subject to averaging and can be interpreted as an indicator of shifting conformer distributions. The residues in the vicinity of the tag substitution need to be excluded from the consideration due to direct influence from the tag. Positive $\Delta\delta$ values correspond to higher alpha-helix propensities, negative ones to beta-sheet propensities, except for C_α , where the opposite relation applies. In Fig. 1, the secondary chemical shifts for the peptide A0 and its fluorescein-tagged derivative fIA0 are compared.

There are clearly significant deviations between the free and tagged forms. However, no uniform trends can be derived, since most of the chemical shift indices (CSI) do

not deviate much from zero, with exception of the backbone nitrogen atoms. The major conformation according to the referenced secondary chemical shifts appears to be random coil or dynamic. Remarkably, some of the largest deviations between the tagged and untagged forms occur at the opposite end of the tag position, near the C-terminus, in particular for H_α , H_N , and C_β . This fact indicates a non-local influence of the tag on the conformational distribution. The graphs in Fig. 1 show some significant long-distance effects. However, the data are not consistent with a particular preferred single conformation, rather a shift of the complex conformer distribution.

The CSI-derived secondary structure propensity changes can be compared to the data based on the $^3J_{\text{HNH}\alpha}$ coupling constant data, as shown in Fig. 2. All values of the native (untagged) A0 peptide lie within the random coil range. The fluorescein-tagged peptide shows propensity for an alpha-helical conformation at the N-terminus of the sequence, where the tag is attached (residues R1–L5).

In general, the tagged fIA0 peptide seems to have a slightly higher alpha-helical propensity, but the sequence positions according to Fig. 2 do not correspond to the ones found in the CSI data. It should be noted that coupling constants and chemical shifts have different (non-linear) dependence on torsion angles; therefore, averaging over multiple conformations exhibits different sensitivity to conformational parameters.

Relaxation rates in the rotating frame are sensitive to molecular motions in the kHz range; therefore, little effect is to be expected for the short rapidly tumbling peptides investigated. This was experimentally verified as exemplified in Fig. 3. Any major deviations would be an indication of the formation of aggregates.

By contrast, the order parameters S^2 determined computed using the RCI server (see Fig. 4) indicate that a fluorescein tag at the N-terminus increases the order parameters observed toward the C-terminus of the A0 peptide. This is a remarkable long-range effect, the molecular basis of which is not evident.

Comparing TALOS + predictions of secondary structure propensities, which are based on the chemical shifts of the backbone atoms and referenced to a data base of model structures, we find few secondary structure differences between A0 and fIA0, as shown in Table 2. However, the

Table 1 Overview of degree of NMR (^1H , ^{13}C , ^{15}N) resonance assignments of the investigated peptides at 298 K

CARA assignment report	A0 (%)	fIA0 (%)	S^2 S2 298 K (%)	bioS2 (%)	P1 298 K (%)	bioP1 298 K (%)
All spins	88.0	84.0	92.0	93.7	95.0	100.0
Backbone	95.8	100.0	98.8	100.0	95.4	100.0
Sidechain	85.5	79.0	89.0	91.0	94.7	100.0

Fig. 1 Comparison of secondary chemical shifts of H_α (a), H_N (b), N (c), C_α (d), and C_β (e), and the difference of the secondary chemical shifts of C_α and C_β (f) atoms for peptides A0 (dark blue) and flA0 (orange) with error bars (grey) (colour figure online)

results of two other prediction approaches (ROSETTA and CS23D) show striking differences, as illustrated in Fig. 5.

The N-terminal fluorescein tag appears to induce a helix conformation near the N-terminus, which is absent in the untagged peptide. The two computational approaches agree in this feature, while a “croissant”-like overall bend is only predicted by CS Rosetta.

Therefore, for A0, the available NMR data evaluated by the approaches listed indicate additional order and rigidity is being induced by the fluorescein tag attached to the N-terminus of A0. The preferred secondary structure appears to be shifted toward helical.

For the 22 amino acid peptide P1 and its N-terminally biotin-tagged derivative bioP1, the secondary chemical shift

values are shown in Fig. 6. The secondary chemical shifts of the tagged and untagged peptides lie within the random coil conformation range. They exhibit no significant changes when comparing P1 to the biotinylated derivative. Only the H_N and N resonances near the N-terminus show minor deviations, which are expected due to the local changes in the chemical shielding by the tag attachment. The comparison of the three bonds H_N-H_α coupling constants reveals unexpected and significant deviations in Fig. 7.

Biotinylation apparently slightly favours alpha-helical conformation. The effect is more pronounced than expected from the chemical shift results but not inconsistent with those. Toward the C-terminus, random coil prevails. Relaxation rates in the rotating frame (Fig. 8) exhibit the same

Fig. 2 Coupling constants $^3J_{\text{HNH}\alpha}$ of peptides A0 (dark blue) and flA0 (orange) with error bars (grey). The grey region between 6 and 8 Hz indicates the range, where random coil is the most probable secondary structure; below and above alpha-helical and beta strand-like conformations, respectively, prevail (colour figure online)

Fig. 3 $R_{1\rho}$ relaxation rates of A0 (dark blue) and flA0 (orange) with error bars (grey) determined from the modified TOCSY spectra (colour figure online)

Fig. 4 Order parameter S^2 obtained from the RCI server of flA0 (orange) and A0 (dark blue) indicating higher flexibility at the C-terminus of the sequence for A0 compared to flA0 with a fluorocin tag attached at the N-terminus (colour figure online)

sequence dependence with and without tag; however, the tagged P1 peptide appears to be overall less mobile, while a decrease in the order parameter (Fig. 9) towards the C-terminus upon biotinylation at the C-terminus is observed.

Table 2 TALOS + secondary structure prediction of A0 and flA0

A0	flA0
RKKRLKLLKRL	RKKRLKLLKR LL
LL LLLLL LLLLL	LL LL LL LLLLL EL

As in the preceding case (A0 and flA0), the structure predictions by CS Rosetta and GeNMR for P1 and bioP1 show most striking differences, in spite of the small chemical shift changes, as summarized in Fig. 10.

The predictions for the 22 amino acid P1 peptides show increased alpha-helical propensity upon biotinylation, but to a lesser extent in CS Rosetta. For the biotinylated P1 peptides, the lowest energy structure from CS Rosetta shows some short helical parts, whereas the ten best structures are in coil conformation, but the ensemble structures from GeNMR generally coincide as they all have a helix between residues A6 and D14 (Table 3).

As expected from the previous literature, the biotin tag has a very low influence on structure as evidenced by the small changes in NMR parameters. The significant structural changes predicted by CS Rosetta and GeNMR are difficult to rationalize and may be due to an inherent bias in the data base. This case clearly needs further investigation.

For the peptide S2 and its biotinylated derivative bioS2, the secondary chemical shifts are compared in Fig. 11. Here, the CSI trends are not consistent with the observed changes in coupling constants (Fig. 12); however, this may be caused by small charge differences due to the pH deviation.

In general, again, the biotin tag has no significant influence on structural propensities, which is in accord with the other studies [10]. The differences of secondary chemical shifts of S2 and bioS2 are probably due to the different pH values, since the general trends are comparable. For technical reasons, there was a deviation of -1 pH unit from the S2 to the bioS2 sample, which explains the small near constant offset of all CSI values.

Both S2 peptides tend to alpha-helical conformation at the N-terminus. Around P10, there is an apparent discontinuity, which reflects the greater rigidity caused by the proline, but cannot be interpreted in more detail without further data.

As it is the case of A0 peptides, also the S2 peptide does not show significant changes in rotating frame relaxation rates (Fig. 13). Only at the N-terminal side of bioS2, where

Table 3 TALOS + results for P1 and bioP1

P1	bioP1
DSLIDIAELVM ELEDEFGTEI PD	DSLIDIAELVM ELEDEFGTEI PD
cLLLLLLLLLL LLLLLLLLLL LL	LLLLLHLLL LLLLLLLLLL LL

Fig. 5 The lowest energy structures predicted by CS Rosetta (**a, c**) and CS23D (**b, d**) of A0 (**a, b**) and fA0 (**c, d**)

the biotin tag is attached, the rates are slightly higher compared to the native peptide and big differences appear for residue L12 with a very small relaxation rate, indicating unusually high local mobility, which may be an effect of the proline in position 10.

Based on the order parameters (Fig. 14), the flexibilities of the biotinylated peptides bioS2 and bioP1 are in comparable ranges, but bioS2 seems to be less flexible than S2 in contrast to bioP1, which has higher flexibility than P1, at the N-terminal part of the sequence.

So apparently, the influence of a biotin tag depends on particular properties of the amino acids of each peptide. Peptide P1 is much more acidic than S2, which mainly consists of basic amino acids. This difference can be appreciated by comparing the hydrophobicity plots of the two peptides in Fig. 15. The highly different hydrogen bond donor and acceptor properties of P1 versus S2 may well explain that, due to short lived dynamic contacts of amino acid residues to the biotin tag moiety, different minority conformations may contribute to a different conformation distribution.

The differences between tagged and untagged peptide S2 in the TALOS + computed structural propensities are the highest ones reported in this investigation as shown in Table 4. Again, the structure predictions by CS Rosetta and GeNMR (CS23D) show significant differences, in spite of the small chemical shift changes, as summarized in Fig. 16.

The lowest energy structure predictions of the 22 amino acid peptides S2 and bioS2 result in helical conformations, especially at the N-terminal side, whereby the

helical content is higher in the prediction of GeNMR for the untagged sample, which is in contradiction to the result from CS Rosetta where only the tagged form assumes a partial helical conformation.

The combined results from the NMR data presented above clearly show that N-terminal tags do have a more than local effect on the conformational distribution and dynamics of small disordered peptides. The methodology used has not been developed and optimized for peptides, so the structural implication needs to be interpreted with caution. In particular, one has to bear in mind that the prediction methods (CS Rosetta and GeNMR) used are targeting larger proteins and the chemical shift and secondary structure databases underlying the methods may introduce an inappropriate bias. However, even when ignoring the particular secondary structure predicted, there are clearly differences between the free and tagged forms of the peptides, which cannot be attributed to a methodological artefact as the tags themselves are not represented in the input data or databases. Without further detailed investigations, one may suspect that the subtle effects are over-amplified by the algorithms present in these programs. Concerning the dynamics, data are not fully conclusive and additional experiments determining ^{15}N and ^{13}C relaxation data would be highly beneficial; however, at the given concentrations and solubilities, this would require isotopic labelling, which, since chemical synthesis is required for these peptides, would be prohibitively expensive.

Fig. 6 Comparison of secondary chemical shifts of H_{α} (a), H_N (b), N (c), C_{α} (d), and C_{β} (e), and the difference of the secondary chemical shifts of C_{α} and C_{β} (f) atoms for peptides P1 (dark blue) and bioP1 (orange) with error bars (grey) (colour figure online)

Conclusion

Prediction and experimental assessment of the structural propensities of small peptides in solution is more challenging than for proteins. This is due to their flexibility and shallow energy landscape which causes the co-existence of several conformers appropriately described by a distribution. Introduction of tags may cause changes in the conformer distributions, which give rise to changes in the averaged

chemical shifts and coupling constants observed. Also for the examples investigated, this behaviour has been corroborated to varying extent. Structure prediction programs suggest that higher alpha-helical propensity is often owed to N-terminal tagging in the cases investigated. The more pronounced effects are observed on dynamic parameters, exemplified here by proton relaxation rates in the rotating frame.

Fig. 7 Coupling constants $^3J_{\text{HNH}\alpha}$ of peptides P1 (dark blue) and bioP1 (orange) with error bars (grey). The grey region between 6 and 8 Hz indicates the range, where random coil is the most probable secondary structure; below and above alpha-helical and beta strand-like conformations, respectively, prevail (colour figure online)

Fig. 8 $R_{1\rho}$ relaxation rates of P1 (dark blue) and bioP1 (orange) with error bars (grey) determined from the modified TOCSY spectra (colour figure online)

These findings may have significant consequences for

Fig. 9 Order parameter S^2 of P1 (dark blue) and bioP1 (orange) determined from the RCI-server. Higher values indicate a more rigid local structure (colour figure online)

interpretation of various interaction assays based on tagged peptides and proteins with intrinsically disordered regions. Even when no large shifts in the conformer distributions occur, changes in kinetics may bias biological assays and the interpretation of small differences in particular when dealing with low activity interactions should be reviewed very critically.

Further systematic investigations on the influence of various tags on small bio-active peptides will be required to obtain a conclusive picture and separate general effects from particular properties of specific systems.

Experimental

Peptide samples were custom-synthesized from GenScript (Piscataway, NJ, USA) at a purity between 90 and 99%. For NMR, the lyophilized powders were dissolved in a mixture of $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) to a concentration of 1 to 2 mM. Because of a better solubility, the biotinylated AcpP-derived peptide P1 was dissolved in a 10 mM acetate buffer solution at pH 4.9.

NMR spectra were measured at 298 K on a Bruker AVANCE III spectrometer operating at 700 MHz ^1H frequency, equipped with a triple resonance cryogenically cooled (TCI) probe.

For assignment of peptide signals and sequential assignment, total correlation spectroscopy (TOCSY) and nuclear Overhauser spectroscopy (NOESY) were acquired using 60 and 80 ms of mixing time, respectively. For assignment of heteronuclear signals, ^{15}N - ^1H - and ^{13}C - ^1H -heteronuclear single quantum spectroscopy (HSQC) as well as ^{13}C - ^1H -heteronuclear multiple quantum spectroscopy (HMCB) spectra were acquired. The chemical shift assignment procedures were done using the program CARRA [33]. The spectra were referenced using 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) (A0, fA0, P1, bioP1, and S2) or trisilylpropionate (TSP) (bioS2). The pH values of the peptide in $\text{H}_2\text{O}/\text{D}_2\text{O}$ solutions were between 3 and 5 using no further adjustment, which resulted in a pH of 3.3 and 2.97 for A0 and fA0, respectively and a pH of 4.0 and 3.1 for S2 and bioS2, respectively. Only in the case of P1 and bioP1, the pH was adjusted to 4.9. For solutions referenced using TSP, the pH shift was compensated relative to DSS using the following formula [34].

$$\delta_{\text{TSP-1H(ppm)}} = 0.003 - 0.019 * (1 + 10^{5.0-\text{pH}})^{-1}$$

Coupling constants were determined from TOCSY spectra by applying Gaussian line shape deconvolution on 1D rows of well-resolved cross peaks [35]. Rotating frame relaxation parameters were determined from TOCSY-type

Fig. 10 Structure predictions for P1 (**a, b**) and bioP1 (**c–f**) by CSRosetta (**a, c, e**) and GeNMR (**b, d, f**). The ten best structures obtained by the respective method are superimposed in **e** and **f**

spectra using a pulse sequence pulses listed in the Supplementary Information derived from the Bruker standard pulse sequence “mlevesgpph” by insertion of variable spin-lock period, with durations between 0 and 2 s [36].

The assignment tables have been deposited in the Biological Magnetic Resonance Data Bank (BMRB, <http://www.bmrb.wisc.edu>) under the submission numbers 27105, 27114, 27116, 27115, 27118, and 27117 for A0, flA0, S2, bioS2, P1, and bioP1 respectively.

The relaxation rate in the rotating frame $R_{1\rho} = 1/T_{1\rho}$ is very sensitive to molecular conformational processes occurring in the milli- to micro-second time frame. It is, therefore, suitable for detecting changes in conformational dynamics, even when the overall structural propensity is not or only slightly changed. We determined the $R_{1\rho}$ relaxation rates for

HN protons in the peptides using the pulse sequence given in the Supplementary Information. Relaxation rates were evaluated using the program Sparky [37] for the evaluation of 2D cross peaks from the modified TOCSY spectra (see above).

Secondary chemical shifts $\Delta\delta$ are calculated as the difference between observed amino acid chemical shift δ_{obs} and the corresponding random coil value δ_{rc} as $\Delta\delta = \delta_{\text{obs}} - \delta_{\text{rc}}$ [34]. Error estimates correspond to the time-domain resolution (Topspin parameter “fidres”) of the respective experiments.

Using the Karplus relation, three-bond scalar coupling constants 3J are related to torsion angles: $^3J = A \cos 2\theta + B \cos \theta + C$, with θ being the dihedral angle and A , B , and C are constants which are dependent on the involved nuclei

and bond angles [38]. Residues with a $^3J_{\text{HNH}\alpha}$ above 8 Hz are more likely to be in a β sheet conformation, while values below 6 Hz indicate alpha-helical propensity. Between 6 and 8 Hz, random coil conformation is prevalent. In small highly dynamic peptides, we note that the experimentally determined values correspond to weighted averages over a distribution of conformers present rather than a particular rigid geometry. Experimental determination of three-bond scalar coupling constants $^3J_{\text{HNH}\alpha}$ can be an error prone procedure,

due to spectral overlap of the multiplet components and limited resolution. Several approaches have been published [39–41]. For the purpose of comparing secondary structure, propensities depending on tags relative changes are more important than exact J values. Therefore, a fast and robust method was chosen: Gaussian line shape deconvolution on 1D rows of well-resolved cross peaks in TOCSY spectra. The error was estimated as the sum of the standard deviations of the picked peak frequencies used for deconvolution.

Fig. 11 Comparison of secondary chemical shifts of H_{α} (a), H_N (b), N (c), C_{α} (d), and C_{β} (e), and the difference of the secondary chemical shifts of C_{α} and C_{β} (f) atoms for peptides S2 (dark blue) and bioS2 (orange) with error bars (grey) (colour figure online)

Fig. 12 Coupling constants $^3J_{\text{HNH}\alpha}$ of peptides S2 (dark blue) and bioS2 (orange) with error bars (grey). The light grey region between 6 and 8 Hz indicates the range, where random coil is the most probable secondary structure, below and above alpha-helical and beta strand-like conformations, respectively, prevail (colour figure online)

Fig. 13 $R_{1\rho}$ rotating frame relaxation rates of peptides S2 (dark blue) and bioS2 (orange) with error bars (grey) (colour figure online)

Fig. 14 Order parameters S^2 of peptides S2 (dark blue) and bioS2 (orange) determined from the RCI-server. Higher values indicate a more rigid local structure (colour figure online)

The Random Coil Index (RCI) is calculated from the sequence-specific assigned chemical shifts. The S^2 is a measure for the internal dynamics of a molecule and was

Fig. 15 Hydrophobicity plots with values according to Kyte and Doolittle [32] of peptides P1 (filled circles) and S2 (filled triangles) showing the very different values for the peptides. S2 is basic, while P1 has an amino acid composition resulting in a more acidic character

Table 4 TALOS + results for S^2 and bioS2

S^2	bioS2
RIQALYSKQ PKLVERILTK NEQ	RIQALYSKQP KLVERILTKN EQ
sLLLLLLLLL LLELLLLLLL LLL	LHLHHEEELL EEEEEELLL HL

determined by the program RCI developed by Berjanskii and Wishart [42]. Highly flexible or rigid areas of a peptide or protein can be recognized by comparison with random coil shifts, which usually correspond to flexible parts, while rigid parts deviate from the random coil shifts. From the RCI, the model free order parameter S^2 can be determined using the following relationship: $S^2 = 1 - 0.5 \ln(1 + \text{RCI} * 10.0)$. A value of S^2 near 0 corresponds to unrestricted tumbling; a value near 1 indicates rigidity.

For visualization and modelling of structural propensities using NMR-derived data, the platforms CS Rosetta, TALOS + [43] and CS23D [44] were employed. In principle, the sizes of the investigated peptides are smaller than recommended for these programs; however, we used their prediction to compare the structures of same size, and therefore, only the relative differences are relevant. For the determination of secondary structure propensities, several NMR-derived parameters were used, including measurement of coupling constants ^1H – ^1H TOCSY spectra, mapping of the chemical shift index [45], and using the secondary chemical shift data [34] as input for state-of-the-art prediction methods TALOS +, CS-Rosetta [46], RCI webserver [42], and GenMR [47], which implements the CS23D code. Prediction results from TALOS + indicate secondary structural motifs by the letters L, H, and E for (random) coil, helix, and β -strand, respectively.

Fig. 16 Lowest energy structures predicted by CS Rosetta (**a, c**) and GeNMR (CS23D) (**b, d**) of peptide S2 (**a, b**) and its tagged derivative bioS2 (**c, d**)

Acknowledgements Open access funding provided by Johannes Kepler University Linz. This research was supported by the FFG-Bridge Grant No. 3696376. All NMR experiments were performed at the NMR laboratory of the Upper Austrian–South Bohemian Research Infrastructure Center at JKU Linz, co-financed by the European Union in the context of the Project “RERI-uasb”, EFRE RU2-EU-124/100-2010 (ETC Austria-Czech Republic 2007– 2013, Project M00146).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Kimple ME, Brill AL, Pasker RL (2013) *Curr Protoc Protein Sci* 73(1):9
2. Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) *Protein Expr Purif* 41:98
3. Zhao X, Li G, Liang S (2013) *J Anal Methods Chem* 2013:ID581093
4. Giepmans BNG, Adams SR, Ellisman MH, Tsien RY (2006) *Science* 312:217
5. Goncalves MS (2009) *Chem Rev* 109:190
6. Wombacher R, Cornish VW (2011) *J Biophotonics* 4:391
7. Haraguchi T, Shimi T, Koujin T, Hashiguchi N, Hiraoka Y (2002) *Genes Cells* 7:881
8. Waugh DS (2005) *Trends Biotechnol* 23:316
9. Rathner P, Stadlbauer M, Romanin C, Fahrner M, Derler I, Müller N (2018) *Protein Expr Purif* 146:45
10. Ahn M, De Genst E, Kaminski Schierle GS, Erdelyi M, Kaminski CF, Dobson CM, Kumita JR (2012) *PLoS One* 7:e50192
11. Chang C-Y, Chang C-P, Chakraborty S, Wang S-W, Tseng Y-K, Wang C-C (2016) *J Biol Chem* 291:17102
12. Zhao D, Huang Z (2016) *Bioinorg Chem Appl* 2016:1
13. Klose J, Wendt N, Kubald S, Krause E, Fechner K, Beyermann M, Bienert M, Rudolph R, Rothmund S (2004) *Protein Sci* 9:2470
14. Bucher MH, Evdokimov AG, Waugh DS (2002) *Acta Cryst D* 58:392
15. Thielges MC, Chung JK, Axup JY, Fayer MD (2011) *Biochemistry* 50:5799
16. Perron-Savard P, De Crescenzo G, Le Moual H (2005) *Microbiology* 12:3979
17. Lindbo S, Javad G, Astrand M, Honarvar H, Orlova A, Hober S, Tolmachev V (2016) *Bioconjugate Chem* 3:716
18. Kuo WHK, Chase HA (2011) *Biotechnol Lett* 2011:1075
19. Ledent P, Duez C, Vanhove M, Lejeune A, Fonze E, Charlier P, Rhazi-Filali F, Thamm I, Guillaume G, Samyn B, Devreese B, Van Beeumen J, Lamotte-Brasseur J, Frère J-M (1997) *FEBS Lett* 413:194
20. Hutsell SQ, Kimple RJ, Siderovski DP, Willard FS, Kimple AJ (2010) *Methods Mol Biol* 627:75
21. Gruber HJ, Marek M, Schindler H, Kaiser K (1997) *Bioconjugate Chem* 8:552
22. Kada G, Riener CK, Hinterdorfer P, Kienberger F, Stoh CM, Gruber H (2002) *Single Mol* 3:119
23. Moerke NJ (2009) *Curr Protoc Chem Biol* 1:1
24. Brieger A, Plotz G, Hinrichsen I, Passmann S, Adam R, Zeuzem S (2012) *PLoS One* 7:e31863
25. Margolin W (2012) *J Bacteriol* 194:6369

26. Souza-Fagundes EM, Frank AO, Feldkamp MD, Dorset DC, Chazin WJ, Rossanese OW, Olejniczak ET, Fesik SW (2012) *Anal Biochem* 421:742
27. Edman P (1956) *Nature* 177:4510
28. Jullian M, Hernandez A, Maurras A, Puget K, Amblard M, Martinez J, Subra G (2009) *Tetrahedron Lett* 50:260
29. Bertini I, McGreevy KS, Parigi G (2012) *NMR of biomolecules: towards mechanistic systems biology*. Wiley, London
30. Williamson MP (1993) *Methods Mol Biol* 17:69
31. Volk K, Breunig SD, Rid R, Herzog J, Bräuer M, Hundsberger H, Klein C, Müller N, Önder K (2017) *Biol Chem* 398:125
32. Kyte J, Doolittle RF (1982) *J Mol Biol* 1:105
33. Keller RW (2004) Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. PhD Thesis, ETH Zürich
34. Wishart DS (2011) *Prog Nucl Magn Reson Spectrosc* 58:62
35. Vuister GW, Bax A (1993) *J Am Chem Soc* 115:7772
36. Samson AO, Chill JH, Anglister J (2005) *Biochemistry* 44:10926
37. Goddard TD, Kneller DG (2007) *Sparky 3*. San Francisco: University of California. Available from: <http://www.cgl.ucsf.edu/home/sparky/>. Accessed 30 Nov 2018
38. Cavanagh J, Fairbrother WJ, Palmer AG III, Rance M, Skelton NJ (2007) *Protein NMR spectroscopy*. Elsevier, Amsterdam
39. Griesinger C, Sørensen OW, Ernst RR (1985) *J Am Chem Soc* 107:6394
40. Wang Y, Nip AM, Wishart DS (1997) *J Biomol NMR* 10:373
41. Neri D, Otting G, Wüthrich K (1990) *J Am Chem Soc* 112:3663
42. Berjanskii MV, Wishart DS (2007) *Nucleic Acids Res* 35:W531
43. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) *J Biomol NMR* 44:213
44. Wishart DS, Arndt D, Berjanskii M, Tang P, Zhou J, Lin G (2008) *Nucleic Acids Res* 36:W496
45. Wishart DS, Sykes BD, Richards FM (1992) *Biochemistry* 31:1647
46. Shen Y, Vernon R, Baker D, Bax A (2009) *J Biomol NMR* 43:63
47. Berjanskii M, Tang P, Liang J, Cruz JA, Zhou J, Zhou Y, Bassett E, MacDonell C, Lu P, Lin G, Wishart DS (2009) *Nucleic Acids Res* 37:W670

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.