

Impact of Azidohomoalanine Incorporation on Protein Structure and Ligand Binding

Florian Lehner,^[a] Denis Kudlinzki,^[a, d] Christian Richter,^[a] Henrike M. Müller-Werkmeister,^[c] Katharina B. Eberl,^[c] Jens Bredenbeck,^[c] Harald Schwalbe,^{*,[a]} and Robert Silvers^{*,[a, b]}

The impact of the incorporation of a non-natural amino acid (NNAAs) on protein structure, dynamics, and ligand binding has not been studied rigorously so far. NNAAs are regularly used to modify proteins post-translationally in vivo and in vitro through click chemistry. Herein, structural characterisation of the impact of the incorporation of azidohomoalanine (AZH) into the model protein domain PDZ3 is examined by means of NMR spectroscopy and X-ray crystallography. The structure and dynamics of the apo state of AZH-modified PDZ3 remain mostly unperturbed. Furthermore, the binding of two PDZ3 binding peptides are unchanged upon incorporation of AZH.

The interface of the AZH-modified PDZ3 and an azulene-linked peptide for vibrational energy transfer studies has been mapped by means of chemical shift perturbations and NOEs between the unlabelled azulene-linked peptide and the isotopically labelled protein. Co-crystallisation and soaking failed for the peptide-bound holo complex. NMR spectroscopy, however, allowed determination of the protein–ligand interface. Although the incorporation of AZH was minimally invasive for PDZ3, structural analysis of NNAAs-modified proteins through the methodology presented herein should be performed to ensure structural integrity of the studied target.

Introduction

Genetic code expansion (GCE) and the incorporation of non-natural amino acids (NNAAs) through naturally occurring translation machinery have become two very important approaches to site-selectively label proteins for detailed biophysical investigations, both in vitro and in vivo.^[1–3] GCE utilises orthogonal aminoacyl tRNA synthetase (aaRS)/tRNA pairs that recruit the NNAAs during mRNA translation and recognise the amber stop codon. The second and more simple approach to incorporate NNAAs into proteins is based on amino acid analogues that compete as substrates for the naturally occurring amino acids, leading to a gene-defined partial replacement of the natural amino acid by the added NNAAs.^[2] This approach was first described in 2002 for the residue-specific incorporation of the methionine surrogate azidohomoalanine (AZH) by using a

methionine auxotrophic *Escherichia coli* strain, in which AZH served as a substrate of the natural methionyl tRNA synthetase (MetRS), leading to incorporation of AZH instead of methionine.^[4]

Incorporation of the azide group into proteins allows post-translational modifications commonly used in click chemistry (Figure S1 in the Supporting Information).^[5] Bioorthogonal (BONCAT), fluorescent (FUNCAT) and nitrilase-activatable non-canonical amino acid tagging (NANCAT) are among AZH-based click chemistry techniques that have been developed to visualise proteins by using the azide moiety for labelling.

BONCAT was used in proteomics for the time-resolved profiling of protein synthesis and for the screening of stimulation-induced protein synthesis.^[6,7] In vivo, BONCAT was used to analyse AZH effects on larval zebrafish to show that, apart from metabolic incorporation into newly synthesised proteins, there was no apparent toxic effect of AZH and no impact on simple behaviour of the model organism.^[8] Further in vivo studies have been applied in mice, by labelling the proteome through intraperitoneal injections with AZH.^[9] Application to pregnant dams results in no blatant disruptions in development followed by normal parturition and growth of pups. A nitrile-substituted precursor of AZH has been used for cell-selective labelling of proteomes with NANCAT to label only cells that express nitrilase to produce AZH, resulting in the incorporation of AZH into the nascent proteins.^[10] Visualisation and profiling of the nascent proteome occurred through click chemistry labelling with fluorescent probes or affinity tags.


Another application of azide groups in proteins and, in particular, AZH was presented recently. The azide functional group allows probing site-specific structural changes of proteins and

[a] F. Lehner, Dr. D. Kudlinzki, Dr. C. Richter, Prof. Dr. H. Schwalbe, Dr. R. Silvers
Organic Chemistry and Chemical Biology, Goethe University Frankfurt
Max-von-Laue-Strasse 7, 60438 Frankfurt am Main (Germany)
E-mail: schwalbe@nmr.uni-frankfurt.de

[b] Dr. R. Silvers
Present address:
Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology
Cambridge, MA 02139 (USA)
E-mail: silvers@mit.edu

[c] Dr. H. M. Müller-Werkmeister, K. B. Eberl, Prof. Dr. J. Bredenbeck
Institute for Biophysics, Goethe University Frankfurt
Max-von-Laue-Strasse 1, 60438 Frankfurt am Main (Germany)

[d] Dr. D. Kudlinzki
German Cancer Consortium, DKTK, German Cancer Research Center, DKFZ
Im Neuenheimer Feld 280, 69120 Heidelberg (Germany)

 Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201700437>.

the study of protein–ligand binding through changes in the azide absorption frequency in IR spectroscopy applications.^[11,12] In combination with the chromophore azulene, AZH acts as a sensor to form a donor–sensor pair to detect vibrational energy transfer (VET) from the azulene group through the polypeptide by measuring changes to the azide absorption frequency by means of Vis-IR pump–probe experiments.^[13] VET in peptides has been reported in earlier studies by Botan et al.^[14] They used an azobenzene-based heater attached to a peptide 3₁₀-helix and carbonyl groups as local internal thermometers, in combination with selective ¹³C=O isotope labelling for spatial resolution.

In addition to AZH, homopropargylglycine (HPG) and homoallylglycine (HAG), which can be incorporated by the natural translation machinery, more than 150 genetically encoded NNAAs with a wide spectrum of functional groups are already known and have been part of a great variety of studies on proteins and biocatalysis, which reveals the high importance of NNAAs for structural and functional investigations of proteins.^[3,15]

To date, the extent and impact of the incorporation of NNAAs on structure and dynamics of proteins have not been thoroughly investigated by high-resolution structure-determination methods, such as X-ray crystallography and NMR spectroscopy. To the best of our knowledge, only two X-ray structures have been published, by Debela et al., in which AZH and HPG have been incorporated into human annexin V (PDB ID: 2XO2 (AZH), PDB ID: 2XO3 (HPG)), but no characterisation of the changes to dynamics or binding has been documented. However, high-resolution investigation of structure and dynamics of these modified proteins are necessary to differentiate between artificially introduced changes to the systems and experimentally relevant results.

Herein, we report on the design and conduct of such detailed studies, starting from investigating the incorporation efficiency of NNAAs, characterising the effect of AZH incorporation into the well-studied protein PDZ3 domain of the post-synaptic density protein 95 (PSD-95) and the impact on protein binding to the native-like Ac-NYKQTSV (L₇) peptide binding motif and a modified AzuCa-KQTSV (L_{5A}) peptide ligand. Azu-

lene-1-yl-acetic acid (AzuCa) is introduced as a vibrational energy donor to enable future studies of VET.^[13] The PDZ3 domain of PSD-95 from *Rattus norvegicus* was the starting point of our analysis, as summarised in Figure 1. The incorporation of AZH was performed at position 341, to replace I341 with methionine (I341M), resulting in the PDZ3 mutant PDZ3_{I341M}. Expression in methionine auxotrophic cells allowed us to express the AZH mutant PDZ3_{I341AZH}.

We show that AZH has a minimally invasive impact on the structure and dynamics of PDZ3 in the ligand-free and peptide-bound states. In addition, we observe that ligand modification with AzuCa does not interfere with ligand binding.

Results and Discussion

AZH incorporation

AZH was incorporated into PDZ3 by using a methionine auxotrophic *E. coli* strain B834(DE3) regularly used to incorporate selenomethionine into samples designated for X-ray crystallographic studies.^[16–19] PDZ3_{I341AZH} was expressed by using AZH as a supplement to the regularly used M9 medium, whereas PDZ3_{I341M} and PDZ3_{WT} were overexpressed in *E. coli* BL21(DE3). Initial expression of PDZ3_{I341AZH} under glucose-limited conditions (0.2%), to reduce costs for isotopic labelling, led to AZH incorporation of only around 43%, as estimated by the average intensity ratios between well-separated resonances in the ¹H,¹⁵N HSQC spectra of the same amino acids belonging to either PDZ3_{I341AZH} or PDZ3_{I341M} (see Figure S2). The backbone amide signal of isotopically labelled M341 was also detected; this should not occur due to the utilisation of the methionine auxotrophic *E. coli* strain, which has previously been successfully tested for methionine auxotrophy. After several test expressions, we established an efficient protocol to express isotopically labelled PDZ3_{I341AZH} (either ¹⁵N or ¹³C and ¹⁵N) with AZH incorporation yields of about 94%, as validated by 2D NMR spectroscopy. The key factor for the increased incorporation of AZH was a required twofold increase in glucose levels during expression (0.4%).

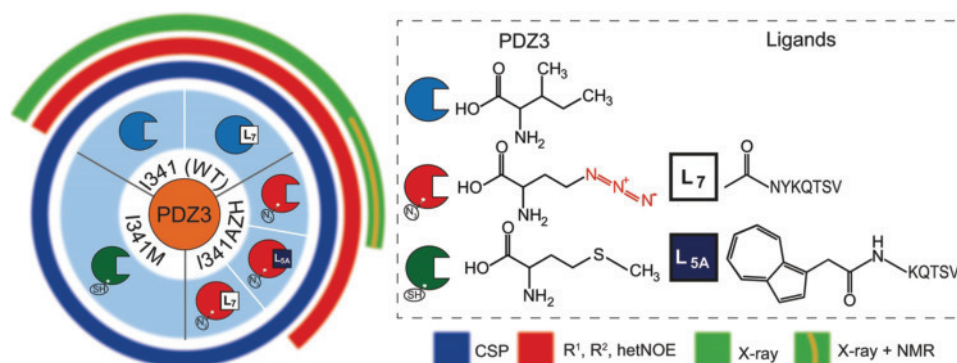


Figure 1. Overview of the investigated proteins derived from PDZ3_{WT} and protein–ligand complexes. Starting from PDZ3_{WT}, the protein variants PDZ3_{I341AZH} and PDZ3_{I341M} were investigated by NMR spectroscopy and X-ray crystallography, in their apo states and together with either the L_{5A} ligand (AzuCa-KQTSV) or the L₇ ligand (Ac-NYKQTSV). Extensive CSP analysis and relaxation data were used for structural and dynamic investigations. X-ray structures of PDZ3_{WT} (apo (PDB ID: 1BFE) and holo (PDB ID: 1BE9)) were available, whereas de novo NMR spectroscopy and X-ray high-resolution structures of PDZ3_{I341AZH} were solved in this work.

Incorporation of isotopically labelled methionine most likely results from residual cobalamin present in the medium, plausibly from rich-medium pre-cultures. The methionine deficiency of B834(DE3) is based on a mutation in *metE*, which deactivates the cobalamin-independent homocysteine methylase that catalyses the transfer of a methyl group from methyltetrahydrofolate to homocysteine in the final step of the methionine synthesis in *E. coli*.^[20,21] It was shown that *E. coli* B834(DE3) cells had normal growth in the absence of methionine, if cobalamin was added to the culture due to a second cobalamin-dependent pathway for methionine synthesis. The enzyme cobalamin-dependent homocysteine methylase (*metH*) enables *E. coli* to produce methionine in the presence of cobalamin. Therefore, the presence of cobalamin has to be limited to achieve a high AZH incorporation yield.

NMR spectroscopic and X-ray structural characterisation of PDZ3_{I341AZH}

Assignment of the backbone resonances of PDZ3_{I341AZH} was performed based on 3D HNCACB, HNHA, HNCB, and ¹H NOESY-¹H,¹⁵N HSQC spectra.^[22–25] Aliphatic side chains were assigned with 3D HCCCONH and 3D ¹H NOESY-¹H,¹³C HSQC spectra.^[26,27] The NNAA AZH within the ¹³C,¹⁵N doubly labelled protein was assigned with 2D ¹³C,¹⁵N doubly filtered ¹H,¹H NOESY and ¹H,¹H TOCSY experiments.^[28–30] A total of 414 of 424 (98%) backbone nuclei were assigned (¹H: 100%, ¹³C: 94%, ¹⁵N: 100%), as well as 360 of 540 (67%) side-chain nuclei (¹H: 78%, ¹³C: 58%, ¹⁵N: 0%). Chemical shifts were deposited into the Biological Magnetic Resonance Data Bank (BMRB ID: 30307).

A total of 2627 distance restraints were extracted from 3D ¹H NOESY-¹H,¹⁵N HSQC, 3D ¹H NOESY-¹H,¹³C HSQC, and 2D ¹H,¹H NOESY spectra. Additionally, a total of 35 ¹H-¹H restraints of the unlabelled AZH were determined from 2D [¹³C,¹⁵N] doubly filtered [¹H,¹H] NOESY, 3D ¹⁵N-edited ¹H NOESY-¹H,¹⁵N HSQC, and 3D ¹³C-edited ¹H NOESY-¹H,¹³C HSQC spectra.^[30] Twenty of the 400 initially calculated conformers were selected based on the CYANA target function. The structure of PDZ3_{I341AZH} (PDB ID: 5W72; Figure 2A and Table 1) obtained by NMR spectroscopy has an average backbone heavy-atom RMSD (R312–A402) mean of (0.69 ± 0.48) Å and an average all heavy atom RMSD (R312–A402) mean of (1.01 ± 0.58) Å (0.85–1.44 Å). The highest residue-specific RMSDs (see Figure S3) are found for the N terminus, which is highly flexible and has therefore been excluded from the average RMSD calculation. Slightly increased residue-specific RMSD values are found for the flexible loop L₁ and also in proximity to the mutation site 341. Although loop L₁ is most likely affected by higher flexibility and conformational exchange (see below), it is unknown whether the slightly higher RMSD around the mutation site is a consequence of the mutation itself, or due to a reduction in NOE contacts as a result of the flexibility of loop L₁, or an intrinsic feature also present in PDZ_{WT}. The AZH side chain is less restrained and points towards the solvent, relative to the side-chain position of I341 in PDZ_{WT} (Figure S4C). Most likely, this is a result of a lack in restraints for the proton-less azide side

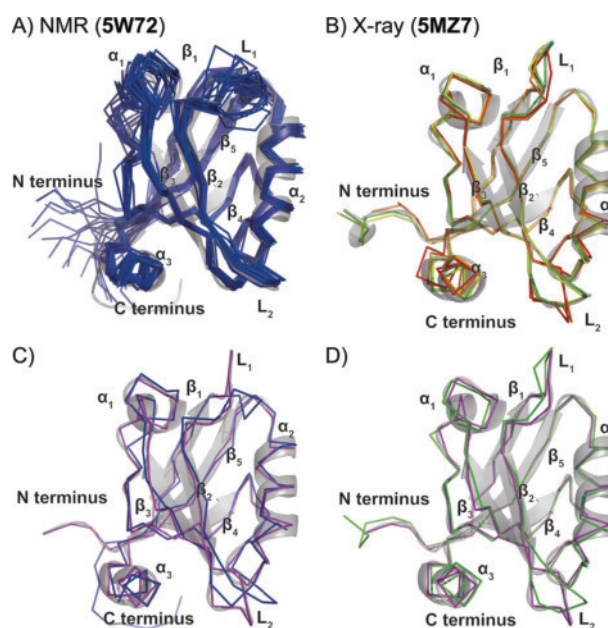


Figure 2. Structures of PDZ3_{I341AZH} obtained by NMR spectroscopy and X-ray crystallography, with a comparison to the PDZ_{WT} crystal structure. A) Structure bundle of PDZ3_{I341AZH} obtained by NMR spectroscopy. B) Crystal structure bundle of PDZ3_{I341AZH} chains A (green), B (yellow), C (orange) and D (red). C) Structure of PDZ3_{I341AZH} obtained by NMR spectroscopy (blue) aligned with the crystal structure of PDZ3_{WT} (PDB ID: 1BFE; violet). D) PDZ3_{I341AZH} X-ray structure (chain A, green) aligned with the crystal structure of PDZ3_{WT} (PDB ID: 1BFE; violet ribbon). Secondary structure elements are visualised in light grey in cartoon shapes. This figure was generated by using the PyMOL Molecular Graphics System, Version 1.7.6.3, Schrödinger, LLC.

Table 1. Structural statistics obtained for the structure bundle determined by NMR spectroscopy.	
Conformational restraints	
distance restraints	2627
intraresidual	1195
sequential ($ i-j = 1$)	555
medium range ($2 \leq i-j \leq 4$)	316
long range ($ i-j \geq 5$)	561
dihedral angle restraints (ϕ/ψ)	136
Restraint violations	
CYANA target function value [\AA^2]	0.87 ± 0.06
RMS distance restraint violation [\AA]	0.0043 ± 0.0007
maximum distance restraint violation [\AA]	0.006
RMS dihedral angle restraint violation [$^\circ$]	0.508 ± 0.06
maximum dihedral angle restraint violation [$^\circ$]	0.627
RMSD	
heavy-atom backbone (R312–A402) [\AA]	0.69 ± 0.48
all heavy atoms (R312–A402) [\AA]	1.01 ± 0.58

chain. We therefore also crystallised PDZ3_{I341AZH} and recorded the X-ray structure to determine the orientation of the AZH side chain.

The crystal structure of PDZ3_{I341AZH} (PDB ID: 5MZ7; Figure 2B and Table S1) was solved at 1.53 Å resolution with four molecules in the asymmetric unit, a crystallographic *R* factor of 18.5% and an *R*_{free} of 22.1%. The molecules are very close to each other, which indicates intermolecular contacts, as listed in

Table S2. These contacts indicate tight crystal packing. Overall, four side chains have multiple conformations in chains A (R313), C (V362, I389) and D (R399). Whereas chains A and B include all residues from G303–A402, in chains C and D, three or four N-terminal amino acids are missing, respectively, probably due to dynamics. Alignment of all chains leads to a backbone heavy-atom RMSD of (0.37 ± 0.25) Å and an all heavy atom RMSD of (0.53 ± 0.40) Å. Major secondary structure elements are well conserved among all chains. Deviations only occur in loops L₁ and L₂; this indicates the presence of conformational heterogeneity (Figure S3). Furthermore, in the C-terminal α_3 -helix, differences are visible that reveal further conformational heterogeneity. For loop L₂, two main conformations occur: one conformation is populated by chains A and B, whereas the other conformation is found for chains C and D. Interestingly, the paired orientation of the loop might result from the difference in flexibility of the N terminus because chains with fewer amino acids visible at the N terminus have one very similar conformation, whereas chains with all amino acids visible populate a different conformation. Overall, chain D shows the highest deviations with respect to the other chains in the asymmetric unit; this indicates that the flexibility of the N terminus affects the orientation of loops L₁ and L₂, as well as the α_3 -helix. In the X-ray structure, the AZH side chain is well defined by the data and points towards the interior, comparable to the side-chain position of I341 in PDZ_{WT} (Figure S4D).

Independent of the method used for the structural investigation, the overall protein fold of PDZ_{341AZH} is highly conserved. The X-ray structure appears slightly more compact; this is most likely to be due to crystal-packing effects. As discussed previously, in the structure determined by NMR spectroscopy, the AZH side chain points towards the solvent, whereas the AZH side chain in the X-ray structure points towards the interior. Additionally, there is only one conformation of the L₂ loop in the structure determined by NMR spectroscopy, whereas there are two conformations in the X-ray structure; this shows that different conformations seem to be populated in solution and in the crystal state.

Comparison of PDZ_{341AZH} and PDZ_{WT} shows only minor changes in structure and dynamics

Structure alignment of wild-type PDZ3 (PDZ_{WT}, PDB ID: 1BFE) with PDZ_{341AZH} shows that the overall structural fold is highly conserved (Figure 2C and D). Differences between PDZ_{WT} and the mutant structure occur in the loop regions. In the X-ray structure, two conformers of PDZ_{341AZH} match the conformation present in PDZ_{WT} whereas the remaining conformers of PDZ_{341AZH} populate a different L₂ loop orientation. These structural differences result from conformational flexibility, as indicated by high *B* factors in the wild-type crystal structure and the lack of assignable resonances for NMR spectroscopic analysis because of exchange broadening. The flexibility of the N and C termini is also conserved in the PDZ_{WT} and PDZ_{341AZH} structures, but best illustrated by the NMR spectroscopy structure bundle. Minor structural differences are observed around the mutation site, especially in α_1 -helices and subsequent

amino acids, as well as at the end of the α_2 -helices. High *B* factors in the α_2 -helices in the PDZ_{WT} crystal structure indicate that these deviations arise from internal flexibility as well, which is also represented by our NMR structure bundle and are not an effect of AZH incorporation.

The observation of a highly flexible N terminus in the structure determined by NMR spectroscopy is in agreement with missing amino acids of the N terminus in the X-ray structures of PDZ_{341AZH} and PDZ_{WT}. This finding supports conservation of the N-terminal flexibility of PDZ_{WT} upon AZH incorporation.

To map the changes in chemical shifts upon the introduction of AZH, we performed chemical-shift perturbation (CSP) analysis based on 2D ¹H,¹⁵N HSQC spectra of PDZ_{WT} and PDZ_{341AZH} (Figure 3A and B).

The chemical shift changes are mainly located in proximity to the mutation site at position 341. F325 is also marginally affected, which is reasonable because it is located in the previous β_2 -sheet in the binding pocket spatially close to the mutation site. The most, but still moderately, affected amino acid is A343, which is sequentially close to the mutation site. The overall impact of the single mutation on the chemical shifts is very small and the observed CSPs are in good agreement with our structural analysis. Similar CSPs are estimated for the I341M and I341AZH mutations; this suggests that the mutation site causes the majority of perturbations and not the type of mutation itself, although the polarity of the AZH side chain is much higher than that of the hydrophobic aliphatic side chains of isoleucine in the wild-type and S-methyl thioether of PDZ_{341M} (Figure S5).

Furthermore, we investigated changes in dynamics upon incorporation of AZH into the protein by recording ¹⁵N longitudinal and transverse relaxation rates (*R*₁ and *R*₂), as well as by detecting the heteronuclear NOEs (hetNOEs). The relaxation data were used to calculate the residue-specific order parameters, *S*², to compare the dynamics of PDZ_{WT} and PDZ_{341AZH} (Figure 3C and D). High-order *S*² indicates that the majority of residues are well structured in both PDZ_{WT} and PDZ_{341AZH}. Three N-terminal and one C-terminal residues are flexible and conserved in the wild type and the AZH mutant, with minor differences in *S*². The *S*² values of G322 in PDZ_{WT} and PDZ_{341AZH} are slightly decreased, which is in good agreement with the structural information because G322 is located in flexible L₁ loop and the ¹H–¹⁵N resonances of the surrounding amino acids are not visible in the HSQC spectra due to exchange broadening.

Thus, the incorporation of AZH has a small effect on the internal flexibility of the protein, which results in a minor destabilisation of the mutant protein. Similar effects are also observed for PDZ_{341M}; this provides strong evidence that our observation is a typical impact that results from a point mutation at position 341 (Figure S6).

Longitudinal (*T*₁) and transverse (*T*₂) relaxation rates were used to estimate the rotational correlation time, τ_c , of a protein, that is, the time for a protein in solution to rotate one radian.^[31] For PDZ_{WT} and PDZ_{341AZH}, we determined rotational correlation times of (6.77 ± 0.24) and (6.44 ± 0.22) ns, respectively. These values correspond to molecular weights (MWs) of

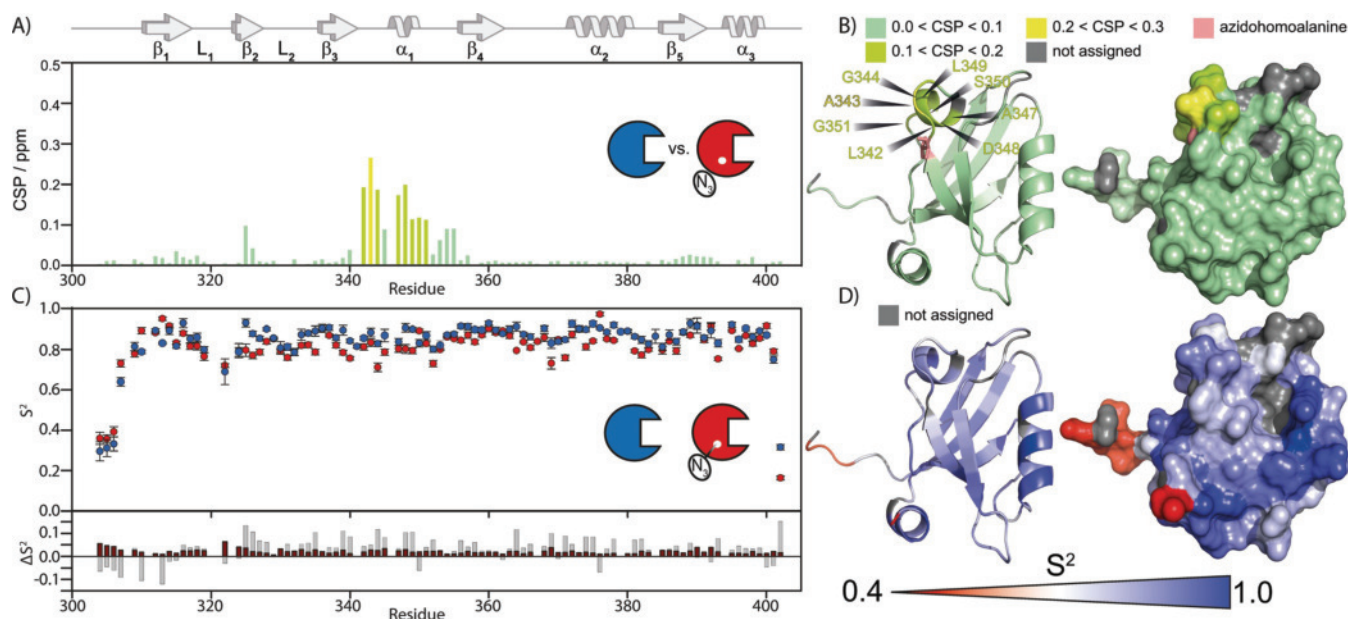


Figure 3. CSPs and order parameter (S^2) comparison of PDZ_{WT} and PDZ_{I341AZH}. A) Absolute combined chemical shift changes of the PDZ_{WT} backbone amide signals upon exchange of I341 (PDZ_{WT}) for AZH (PDZ_{I341AZH}). Secondary structure elements are illustrated at the top in grey. B) Colour-coded CSPs are plotted on the structure of PDZ_{WT} (PDB ID: 1BE9) shown in cartoon and surface representations. Specific amino acids are highlighted to show the most influenced residues. Missing residues are four prolines (308, 311, 346 and 394); two N-terminal glycine residues (302 and 303); and S320, T321, L323, E352 and Q384 due to exchange broadening. C) Order parameters of PDZ3 (blue circles) and PDZ_{I341AZH} (red circles) and their order parameter differences (grey) with errors (red) are plotted versus the amino acid sequence. D) Colour-coded order parameters of PDZ_{I341AZH} are plotted on the structure of PDZ_{WT} (PDB ID: 1BE9) in cartoon and surface representations. Blue illustrates rigid parts of the protein; red indicates flexible parts.

(11.08 ± 0.38) and (10.53 ± 0.36) kDa for PDZ_{WT} and PDZ_{I341AZH}, respectively. Both values are within error of the calculated theoretical MWs of uniformly the ¹⁵N-labelled protein (PDZ_{WT}: 10.914 kDa, PDZ_{I341AZH}: 10.926 kDa). Hence, the incorporation of AZH does not affect the oligomerisation state of PDZ3.

Ligand binding of PDZ_{WT} and PDZ_{I341AZH} remains unaffected upon AZH incorporation

PDZ domains are known for binding to C-terminal tails of their target proteins.^[32] PDZ_{WT} binds to the cysteine-rich PDZ-binding protein (CRIPT) and Doyle et al. showed that the C-terminal peptide derived from CRIPT was the binding motif of PDZ_{WT}.^[33] Therefore, we used the seven (L₇) or five (L_{5a}) C-terminal residues of the CRIPT protein (Figure 1) for our ligand studies to investigate the impact of AZH incorporation on ligand binding, as well as the influence of azulene group modification of the ligand.

The dissociation constants (K_d) for PDZ_{WT} and PDZ_{I341AZH} with native-like ligand L₇, were determined by isothermal titration calorimetry (ITC; Figure S7) to be (2.87 ± 0.09) and (2.52 ± 0.04) μ M, respectively; this shows that the AZH mutation does not affect ligand binding. The ITC measurements for PDZ_{I341AZH} with ligand L_{5a} yielded a dissociation constant of (14 ± 2) μ M, which was slightly elevated relative to binding of native-like ligand L₇ to both PDZ_{WT} and PDZ_{I341AZH}. These dissociation constants are in good agreement with K_d values reported by Petit et al.,^[32] who determined a dissociation constant of (1.2 ± 0.1) μ M for the binding of a native-like nonamer peptide to PDZ_{WT}. Furthermore, removal of the seven C-termi-

nal residues of PDZ3 (PDZ _{Δ 7ct}) lowers the affinity 21-fold, resulting from removal of Y397 and F400, which are thought to interact with the N-terminal residues of the nonamer ligand. Thus, it is likely that the significantly shorter peptide ligand L_{5a} binds more weakly to PDZ_{I341AZH} with a K_d value comparable to that of the interaction of the nonamer peptide with PDZ _{Δ 7ct}.

We analysed the CSPs of binding of the ligands L₇ and L_{5a} to PDZ_{I341AZH}. The residue-specific CSPs induced by both ligands are very similar, as easily seen in related CSP plots (Figure 4A and B). Most pronounced CSPs are found for residues in β -sheet β_2 and in the L₂ loop region. Also, remote effects can be observed in the α_1 -helix and its flanking residues, which show high CSPs, although they are not spatially close to the bound peptide ligands. Residues in the α_2 -helix experience a change in the chemical environment upon ligand binding that is independent of whether ligand L_{5a} or L₇ is bound to PDZ_{I341AZH}. The ligand binding site of the PDZ domains share a common, highly conserved X Φ G Φ binding motif (here: G₃₂₂LGF₃₂₅), in which X is any amino acid residue and Φ is a hydrophobic residue.^[34] The X Φ G Φ binding motif is situated at the C-terminal end of loop L₁ and beginning of β -sheet β_2 . Accordingly, residues G324 and F325 show strong CSPs upon ligand binding (Figure 4A and B). Residues S320, T321 and L323 cannot be detected due to conformational exchange broadening in both apo PDZ_{I341AZH} and PDZ_{WT}. Importantly, the chemical shifts of these residues are unperturbed upon AZH incorporation (Figure 3A). Furthermore, the X Φ G Φ binding motif of PDZ_{I341AZH} is structurally identical to that of PDZ_{WT} (Figure S4B) and shows significant dynamics, especially in L323 (Figure S4A).

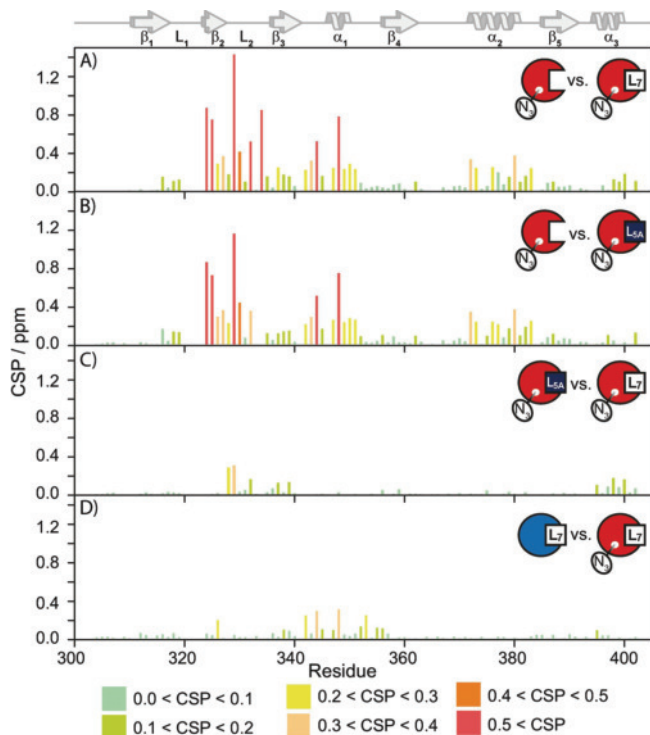


Figure 4. Residue-specific CSPs upon ligand binding to PDZ_{WT} and PDZ_{I341AZH}. A) CSPs upon ligand L_{5A} binding to PDZ_{I341AZH}. B) CSPs induced by ligand L₇ binding to PDZ_{I341AZH}. C) CSPs between PDZ_{I341AZH} bound to either ligand L_{5A} or L₇. D) CSPs between the PDZ_{WT}:L₇ and PDZ_{I341AZH}:L₇ complexes.

Differences between ligand-induced CSPs are very small, occurring with significant amplitude only for residues in the loop regions, especially for residues I329 and G330 in the L₂ loop (Figure 4C). These differences arise from the effect of the bulky hydrophobic azulene group of ligand L_{5A}, which has a different impact on the chemical shift environment than that of the asparagine and tyrosine residues of ligand L₇ located at the same positions. Because the C-terminal residues threonine, serine and valine of both ligands contribute most to ligand binding and no significant CSPs between the L₇- and L_{5A}-bound states of PDZ_{I341AZH} at the XΦGΦ binding motif occur, the binding mode of both ligands is highly conserved.

By calculating the CSPs between the ligand L₇ bound states of PDZ_{WT} and PDZ_{I341AZH}, the impact of AZH on the chemical environment of the protein in the ligand-bound state can be estimated (Figure 4D). The affected residues in PDZ_{I341AZH} (N326, L342, G344, G345, A347, D348, E352, L353, K355, G356, E395) are mainly found in spatial proximity to the mutation site overall, with relatively small amplitudes; this demonstrates that PDZ_{I341AZH} binds ligand L₇ in the same binding mode as that of PDZ_{WT} because significant CSPs result from the AZH mutation. The dynamic changes upon ligand binding of the proteins were analysed through relaxation data analysis on PDZ_{I341AZH} with ligand L_{5A} (Figure S8A) and PDZ_{WT} with ligand L₇ (Figure S8B).

The value of S² for PDZ_{I341AZH} bound to ligand L_{5A} indicates a small difference from that of the ligand-free protein. Stabilis-

ing effects occur for amino acids G344 and E352, which are not located in the peptide binding pocket, but these residues are slightly destabilised through mutation. Overall, there are only small changes to the internal flexibility of PDZ_{I341AZH} upon binding to ligand L_{5A}. Similar results were obtained for the PDZ_{WT}:L₇ complex, without any significant changes to the residue-specific flexibility after ligand binding. A comparison of the ligand-bound states of PDZ_{WT} and PDZ_{I341AZH} (Figure S8C) indicates minor destabilisation of the mutant complex, which is in agreement with the slightly increased flexibility of the ligand-free mutant structure.

To gain further insight into the ligand-binding mode of ligand L_{5A}, we performed 2D [¹³C,¹⁵N] doubly half-filtered NOESY (Figure 5) and TOCSY (Figure S9) experiments to assign the signals of ligand L_{5A} in the binding pocket of PDZ_{I341AZH}^[28,29] Furthermore, we performed a 3D [¹³C,¹⁵N] doubly half-filtered ¹⁵N-separated NOESY HSQC (Figure S10) experiment to determine intermolecular contacts between PDZ_{I341AZH} ¹⁵N-¹H protons and the protons of the naturally abundant ligand L_{5A}^[30]

Through the unique patterns of signals of the amino acids of ligand L_{5A}, we could assign all visible ligand signals. These assigned signals were used together with the backbone assignment of the ¹H-¹⁵N resonances of PDZ_{I341AZH} to analyse the 3D [¹³C,¹⁵N] doubly half-filtered ¹⁵N-separated NOESY HSQC experiment. Intermolecular NOE contacts between ¹⁵N-bound protons of the protein and ¹²C/¹⁴N-bound protons of the ligand were assigned (Figure S10). Four NOEs were detected between I327, A376, I377 and L379 of the protein and H^γ of the T3 of ligand L_{5A}. Two further NOEs were assigned to L379 and K380, which showed their proximity to the H^β proton of V5 of ligand L_{5A}. These NOEs confirm the correct orientation of ligand L_{5A} in the binding pocket, observed in the crystal structure (PDB ID: 1BE9) of PDZ_{WT} in complex with its native C-terminal ligand.^[33]

A flexible protocol for tracking NNAA-induced changes

Based on experience gained for this model protein, we developed a protocol to investigate the effects of NNAA incorporation into proteins (Figure 6). Although the list of experiments and scope of applications is not complete and depends on the investigated systems, we believe that this protocol will help to plan and conduct experiments that are feasible for the majority of systems dealing with NNAAs. All experiments are run on the wild-type and mutant proteins to delineate changes related to the incorporation of NNAA.

Basic set: The basic set is comprised of experiments that can be run with little effort (time and money) and are mainly aimed at investigating the NNAA incorporation efficiency, global structure and protein functionality. Performing experiments on each of these three points to confirm NNAA incorporation and to investigate resulting changes is very important. The experiments of the basic set give insights into changes without local information, but allow a first estimation of the effects of NNAA incorporation. Mass spectrometry can be used

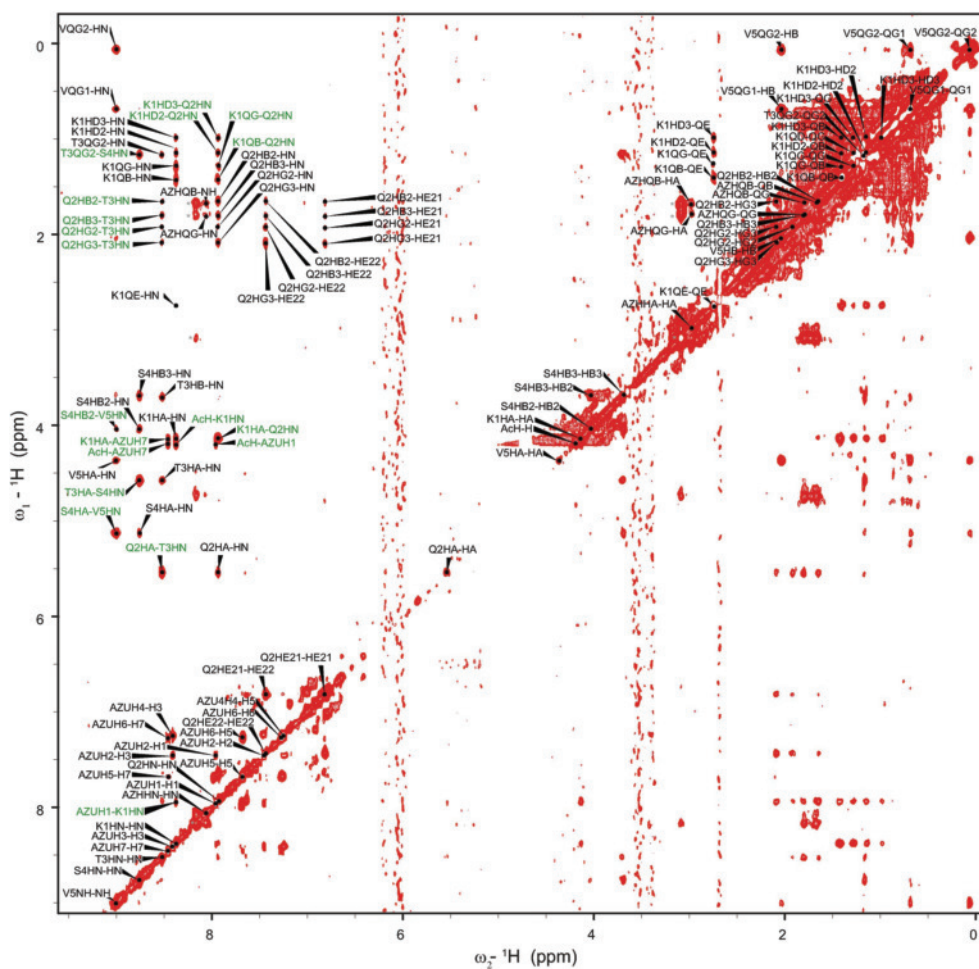


Figure 5. The 2D ^{15}N , ^{13}C doubly filtered NOESY spectrum of the 2:1 PDZ₃[³⁴¹AZH]-L_{5A} complex. Intraresidual NOEs are labelled in black, whereas inter-residual NOEs are labelled in green. The AZH labels describe AZH signals; AZU labels signals are derived from azulene. Signals from ligand L_{5A} and Ac describe the acetate group of the L_{5A} ligand. Unassigned signals are derived from non-isotope-labelled AZH in the ligand-free state and background. Signals marked with a grey star are derived from the AZH resonance of the protein in the ligand-free state. Data were acquired on a 700 MHz NMR spectrometer (^1H Larmor frequency).

to determine incorporation efficiency, assuming the difference in mass between wild type and mutant is sufficient. Circular dichroism (CD) can be used to determine the overall content of secondary structure elements and CD melting curves can be used to determine protein stability. Alternatively, protein stability can be measured by other methods, such as thermal shift assays, also known as differential scanning fluorimetry (DSF).^[35,36] 1D ^1H NMR spectroscopy can be used to determine the overall fold of the protein because the dispersion of signals report on the presence of folded and unfolded elements. From uniformly ^{15}N -labelled samples, ^1H - ^{15}N HSQC can also be used to report on the global structural fold of the protein. The congruency between wild-type and mutant ^1H - ^{15}N HSQC is a very sensitive indicator of the structural fold. Without major structural changes, the majority of signals should be unperturbed and only signals of residues close to the mutation site are affected. Global T_1 and T_2 relaxation rates can be measured easily through 1D NMR spectroscopy and an assignment is not necessary. These rates can be used to determine the rotational correlation time and, subsequently, MW.^[31] This will allow the

determination of the molecular assembly. Alternatively, DOSY can be utilised to investigate the molecular assembly of natural abundance samples.^[37] If ligand binding is an important part of the function of the investigated system, determining the binding affinity is crucial. Several techniques can be utilised, including isothermal calorimetry (ITC),^[38] microscale thermophoresis (MST),^[39,40] surface plasmon resonance (SPR)^[41,42] or fluorescence resonance energy transfer (FRET).^[43,44] Other protein functionality assays should be performed, depending on the function of the investigated system.

Advanced set: The advanced set is comprised of the experiments detailed in the basic set of experiments, plus additional experiments. The first step is to acquire a set of experiments to achieve backbone assignments on a uniformly ^{13}C , ^{15}N -labelled sample. Typically, a set of 3D HNCACB, 3D HNCO, and 3D HN(CO)CACB samples, among others, are used that yield the ^1H , ^{15}N , ^{13}CO , $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ chemical shifts for each residue.^[22,45,46] The use of these chemical shifts allows the determination of residue-specific secondary structure elements by means of, for example, TALOS or CSI.^[47,48] Furthermore, the as-

EFFORT (MONEY + TIME)	APO		HOLO (interactions)
	WT vs. mutant	WT vs. mutant	WT vs. mutant
BASIC	Expression (natural abundance)	2° structure	-
	Expression (uniformly ¹⁵ N)	protein stability	-
	Expression (uniformly ¹³ C/ ¹⁵ N)	protein stability	-
	Circular dichroism (CD)	incorporation efficiency	-
	Thermal-shift assay	structural fold	-
	Mass spectrometry	structural fold	-
	1D ¹ H NMR	CSPs	-
	¹ H, ¹⁵ N HSQC	incorporation efficiency	-
	T ₁ /T ₂ (global)	τ _c , molecular weight	-
	DOSY	hydrodynamic radius	-
ITC, MST, SPR, FRET, etc.	-	ligand affinity	
ADVANCED	Backbone assignment	2° structure	-
	T ₁ /T ₂ /hetNOE	backbone dynamics	-
	¹ H, ¹⁵ N HSQC	CSPs	-
	¹ H, ¹⁵ N HSQC (titration)	-	ligand affinity
EXTENSIVE	Side-chain assignment	-	-
	Structural restraints	-	-
	Edited/filtered NMR	ligand assignment	ligand/protein interface
	NMR structure calculation	NMR apo structure	NMR holo structure
	X-ray crystallization	-	-
X-ray structure calculation	X-ray apo structure	X-ray holo structure	

Figure 6. A flexible protocol for tracking the influence of NAA incorporation into proteins. The basic set provides global information with little effort to give a first impression of the NAA influence on protein folding and function. The advanced set includes site-specific information, along with increasing effort, to allow the investigation of residue-specific effects of NAA incorporation. The extensive set includes the determination of high-resolution structures to address the effects of NAA incorporation.

signment allows a residue-specific analysis of CSPs from ¹H-¹⁵N HSQCs of the wild type and mutant. Additionally, the measurement of T₁ and T₂ relaxation rates and the hetNOE are used to determine the residue-specific pico-/nanosecond backbone dynamics.^[49–51] If ligand binding is an important part of the investigated system, ligand titration can be used to determine ligand affinity from ¹H-¹⁵N HSQC experiments.

Extensive set: Experiments included in the extensive set aim to arrive at a full structural characterisation of the system. Typically, this is only necessary under two circumstances: 1) results obtained with the basic and advanced sets are not conclusive enough to exclude major effects of NAA incorporation, and 2) no structural data is available for the system. The first step is to acquire a set of experiments to achieve side-chain assignments on a uniformly ¹³C, ¹⁵N-labelled sample. Typically, these are 3D HCCH-TOCSY and 3D HCCCONH, among others.^[26,52–55] The side-chain proton assignment allows for the analysis of 2D and 3D NOESY spectra to determine structural restraints.^[25,27,56,57] Additional restraints can be gained from residual dipolar coupling, if necessary.^[50,51,58,59] If ligand binding is an important part of the investigated system, edited and filtered NMR spectroscopy experiments will allow assignment of the resonances of the unlabelled species (small molecule, peptide ligand, protein, etc.), as well as structural parameters for the protein–ligand interface.^[28,30,60–63] Crystallisation and X-ray structure determination can be useful under certain circumstances, if crystallisation of the sample is feasible. For PDZ3, the proton-less azide moiety of AZH did not permit the conclusive determination of side-chain orientation. Furthermore, the structure determined by NMR spectroscopy shows some minor

differences from already published wild-type structures of PDZ3. In such cases, it might be prudent to perform an additional X-ray structure analysis. In our case, the characterisation of ligand binding was not possible by X-ray because co-crystallisation and soaking of ligand L_{5A} did not yield any crystals, or no crystals of sufficient quality.

Conclusion

We performed a detailed analysis of the impact of AZH incorporation on the protein structure and dynamics of the PDZ3 domain and its ligand-binding characteristics, with respect to native-like ligand L₇ and ligand L_{5A} carrying azulene as a vibrational energy donor. Through high-resolution NMR spectroscopy and X-ray crystallography, we were able to solve the solution and crystal structure of PDZ3_{1341AZH} as the basis of our analysis. Structures of PDZ3_{1341AZH} determined by X-ray crystallography and NMR spectroscopy were highly conserved, with only minor differences resulting from internal conformational flexibility of the protein and crystal-packing effects that decreased the conformational space of the crystal structure, in comparison to that of the solution structure determined by NMR spectroscopy.

A comparison of the structures of PDZ3_{WT} and PDZ3_{1341AZH} clarified that AZH incorporation into PDZ3_{WT} was minimally invasive, that is, only local structural effects were observed. This was supported by CSP analysis, in which only moderate, localised perturbations appeared. However, order parameters indicated minor destabilisation of PDZ3_{WT} through the incorporation of AZH.

Binding of L₇ ligand to PDZ3_{WT} and PDZ3_{1341AZH} was mapped on the protein surface through CSP data to show that the wild-type and mutant proteins formed stable protein–ligand complexes, although minor CSP differences were observed. These deviations were caused by AZH and were also visible for the apo state of PDZ3_{1341AZH}, which implied that AZH did not affect ligand binding significantly. This was also valid for ligand L_{5A} because ligand binding of PDZ3_{1341AZH} to ligands L₇ or L_{5A} showed the same ligand-binding mode and only minor differences occurred that were localised to where the azulene moiety of ligand L_{5A} was close to the protein surface. A comparison of the order parameters of the wild-type and mutant complexes showed that the PDZ3_{1341AZH}·L_{5A} complex was slightly destabilised in comparison with that of the PDZ3_{WT}·L₇ complex; this was also found for a comparison of the apo proteins. We conclude that the azulene group of ligand L_{5A} has a minimally invasive character and show that AZH, in combination with azulene, is a valid silent spy system for IR spectroscopic investigations of VET pathways because the native properties of the system remain essentially unperturbed.

Our biophysical validation pipeline is generally applicable for the characterisation of the incorporation of NAAs and their structural and dynamic effects. This is highly recommended for experiments in which small changes are expected to unravel the effects resulting from the introduced internal spy to reassess the acquired data.

Experimental Section

Expression of PDZ3_{WT}, PDZ3_{I341M} and PDZ3_{I341AZH}: The amino acid sequences of PDZ3_{WT}, PDZ3_{I341M}, and PDZ3_{I341AZH} are shown in Figure S11. Uniformly ¹³C-, ¹⁵N- and ¹⁵N-labelled PDZ3_{I341AZH} was recombinantly expressed in methionine auxotrophic *E. coli* B834, whereas ¹³C-, ¹⁵N- and ¹⁵N-labelled PDZ3_{WT} and PDZ3_{I341M} were expressed in *E. coli* BL21(DE3). Cultures were grown at 37 °C on minimal medium supplemented with either ¹⁵NH₄Cl or both ¹⁵NH₄Cl and ¹³C₆-glucose. For PDZ3_{I341AZH}, 160 mg AZH was added to 1 L of the medium. Expression was conducted at 20 °C for about 18 h after induction with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1 mM. Cultures were harvested by means of centrifugation followed by cell lysis with a microfluidiser in buffer containing 50 mM Tris pH 8.0, 10 mM imidazole, 100 mM NaCl and ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor (Roche Applied Science). The insoluble fraction was removed by ultracentrifugation. His-tagged proteins were purified from the supernatant through immobilised metal-ion affinity chromatography (IMAC). The His tags were cleaved through TEV protease digestion during dialysis against buffer containing 50 mM Tris pH 8, 100 mM NaCl and 5 mM β-mercaptoethanol. After a second IMAC step, the flow-through fractions were collected and concentrated by using VivaSpin 20 spin concentrators (Sartorius, 3000 MWCO) and further purified by size exclusion chromatography (GE HiLoad 26/600 75pg) in buffer containing 50 mM sodium phosphate at pH 6.8. Proteins were subsequently concentrated by using VivaSpin 20 spin concentrators (Sartorius, 3000 MWCO).

Backbone resonance assignment of PDZ3_{WT}, PDZ3_{I341M} and PDZ3_{I341AZH} in complex with ligand L₇: Spectra of PDZ3_{I341M} were acquired at 298 K on a 600 MHz spectrometer (Bruker Avance III HD) equipped with a Prodigy 5 mm TCI ¹H/¹⁹F,¹⁵N,¹³C Z-GRD cryoprobe, whereas all spectra of PDZ3_{WT} were acquired on a 600 MHz spectrometer (Bruker DRX) equipped with a 5 mm TXI ¹H,¹⁵N,¹³C XYZ-GRD probe. All samples contained 50 mM sodium phosphate buffer at pH 6.8. Amide backbone resonances of PDZ3_{I341M} were assigned through a 3D HNCACB experiment on a 2.6 mm uniformly ¹³C-, ¹⁵N-labelled sample.^[22] Amide backbone resonances of PDZ3_{WT} were assigned through a 3D HNCACB experiment on a 0.9 mm uniformly ¹³C-, ¹⁵N-labelled sample. The amide proton assignment of PDZ3_{WT} in complex with L₇ ligand was performed on a 0.9 mm sample of uniformly ¹³C-, ¹⁵N-labelled PDZ3_{WT} with ligand L₇ present at a concentration of 1.8 mM. Ligand L₇ was synthesised in house on an ABI 433A peptide synthesiser (Applied Biosystems, USA). All NMR spectra were processed by using Topspin 3.5 (Bruker BioSpin) and analysed by using SPARKY.^[64]

Resonance assignment of PDZ3_{I341AZH} and its ligand complexes: All spectra were acquired at 298 K on 600 (Bruker Avance III HD, Prodigy 5 mm TCI ¹H/¹⁹F,¹⁵N,¹³C Z-GRD cryoprobe), 700 (Bruker Avance III HD, 5 mm QCI ¹H,¹⁵N,¹³C,³¹P Z-GRD cryoprobe), 800 (Bruker Avance III HD, 5 mm TCI ¹H,¹⁵N,¹³C Z-GRD cryoprobe) and 950 MHz (Bruker Avance III, 5 mm TCI ¹H,¹⁵N,¹³C Z-GRD cryoprobe) spectrometers. Assignment of the backbone resonances of PDZ3_{I341AZH} was achieved by 3D triple-resonance NMR spectroscopy experiments (HNCACB, HNHA, HNCB and 3D ¹H-¹⁵N NOESY HSQC) on a 0.5 mm uniformly ¹³C-, ¹⁵N-labelled sample in 50 mM sodium phosphate buffer at pH 6.8.^[22–24,56] Aliphatic side chains were assigned through a 3D HCCCONH experiment and 3D ¹³C-edited ¹H NOESY-¹H,¹³C HSQC spectra recorded under the same sample conditions.^[27,54] The unnatural amino acid within the uniformly ¹³C-, ¹⁵N-labelled sample was assigned with 2D [¹³C,¹⁵N] doubly filtered ¹H,¹H NOESY and ¹H,¹H TOCSY experiments.^[28–30] Through-space NOE contacts were assigned through 2D ¹H,¹H NOESY, 3D

¹H NOESY-¹H,¹⁵N HSQC, and 3D ¹H NOESY-¹H,¹³C HSQC experiments.^[25] Partial backbone assignment of PDZ3_{I341AZH} in complex with ligand L_{5A} was performed on a 0.6 mm uniformly ¹³C-, ¹⁵N-labelled sample in the presence of 1.2 mM ligand L_{5A} in buffer containing 50 mM sodium phosphate pH 6.8, 10% D₂O, 6% [D₆]DMSO through 3D HNCACB and 3D ¹H NOESY-¹H,¹⁵N HSQC experiments. Aliphatic side-chain assignment of the PDZ3_{I341AZH}-L_{5A} complex was achieved through 3D HCCCONH and 3D CCONH experiments under the same sample conditions.^[26] NOE contacts between PDZ3_{I341AZH} and ligand L_{5A} were investigated through a 3D [¹³C,¹⁵N] doubly half-filtered ¹⁵N-separated NOESY- HSQC experiment.^[29,30] Assignment of ligand L_{5A} bound to PDZ3_{I341AZH} was performed on a 0.9 mm uniformly ¹³C-, ¹⁵N-labelled sample in the presence of 0.45 μM ligand L_{5A} through 2D [¹³C,¹⁵N] doubly filtered ¹H,¹H NOESY and ¹H,¹H TOCSY experiments. Partial backbone assignment of PDZ3_{I341AZH} in complex with ligand L₇ was performed on 0.43 mm uniformly ¹³C-, ¹⁵N-labelled sample containing 0.86 mM ligand L₇ in buffer containing 50 mM sodium phosphate at pH 6.8, with 10% D₂O, through a 3D HNCACB experiment. Ligand L_{5A} was synthesised in house on an ABI 433A peptide synthesiser (Applied Biosystems, USA). AzuCa was purchased from Iris Biotech, Germany. All NMR spectra were processed by using Topspin 3.5 (Bruker BioSpin) and analysed by using SPARKY.^[64]

Structure calculations through NMR spectroscopy: Restraints were extracted for structure calculations of PDZ3_{I341AZH} from six different spectra containing structural information: 1) 2D ¹H,¹H NOESY, 2) 3D ¹H NOESY-¹H,¹³C HSQC, 3) 3D ¹H NOESY-¹H,¹⁵N HSQC, 4) 2D [¹³C,¹⁵N] doubly filtered ¹H,¹H NOESY, 5) 3D ¹⁵N-edited ¹H NOESY-¹H,¹⁵N HSQC, and 6) 3D ¹³C-edited ¹H NOESY-¹H,¹³C HSQC. From these contacts, a consensus list of unique restraints was generated (Table S3), containing each contact only once, even if observed in multiple spectra. An upper restraint of 6.5 Å was used. For structure calculations, CYANA 3.97 was used.^[65] The artificial amino acid AZH was defined in the cyana.lib library file (see Table S4). Additionally to the upper restraints, φψ dihedral angles derived from TALOS+ calculations were used as dihedral restraints.^[47] Furthermore, the CYANA functions *ramaaco* and *rotaaco* were used to generate additional angle restraints. The 400 conformers were generated with 40000 torsion angle dynamics steps for each conformer. The final structure bundle was comprised of 20 conformers with the lowest CYANA target function.

Relaxation of PDZ3_{WT} and PDZ3_{I341AZH} and their ligand complexes: The T₁, T₂ and hetNOE data for PDZ3_{I341AZH} were recorded on a 0.45 mm uniformly ¹⁵N-labelled sample in buffer containing 50 mM sodium phosphate at pH 6.8, with 10% D₂O, on a 600 MHz spectrometer (Bruker Avance III HD) equipped with a Prodigy 5 mm TCI ¹H/¹⁹F,¹⁵N,¹³C Z-GRD cryoprobe that used Bruker standard pulse sequences. For measurements on PDZ3_{I341AZH} in complex with ligand L_{5A}, a 0.45 mm uniformly ¹⁵N-labelled sample in buffer containing 50 mM sodium phosphate at pH 6.8, 10% D₂O, 6% [D₆]DMSO with 0.9 mM ligand L_{5A} was prepared. For T₁ times, 12 increments with delays ranging from 50 to 2000 ms, including two repeats, were used as a control. For T₂ times, 11 increments with relaxations delays between 16 and 359 ms, including three repeats were used. Relaxation measurements of PDZ3_{WT} were recorded on a 3.8 mm uniformly ¹⁵N-labelled sample in buffer containing 50 mM sodium phosphate pH 6.8, with 10% D₂O. Measurements of PDZ3_{WT} in complex with ligand L₇ was performed with a protein concentration of 3.5 mM and a ligand concentration of 7 mM under the same buffer conditions. For T₁ times, 12 increments with delays ranging from 100 to 1600 ms, including two repeats, were used as a control. For T₂ times, 12 increments with relaxations

delays between 17 and 340 ms, including two repeats, were used. Data analysis was performed by using Topspin 3.5 (Bruker BioSpin), SPARKY (SPARKY 3.114, University of California, San Francisco) and TENSOR2.^[66] Relaxation data for PDZ3_{WT}, PDZ3_{WT-L7}, PDZ3_{I341AZH}, PDZ3_{I341AZH-L5A}, and PDZ3_{I341M} can be found in Tables S5–S9, respectively.

Rotational correlation time and MW: The rotational correlation times (τ_c) of PDZ3_{WT} and PDZ3_{I341AZH} were estimated by using the previously published relationship [Eq. (1)].^[31,67]

$$\tau_c \approx \left(\sqrt{\frac{6T_1}{T_2}} - 7 \right) / 4\pi\nu_N \quad (1)$$

in which T_1 and T_2 are the longitudinal and transverse ¹⁵N relaxation times, respectively, and ν_N is the nuclear frequency of nitrogen. The average τ_c was calculated by using all T_1 and T_2 values representative of the core protein; residues exhibiting conformational exchange or faster-than-average motion were excluded. The nuclear frequency $\nu_N = 60$ MHz was used. MW was estimated from $\tau_c = 0.6111 \text{ MW}$ (τ_c : rotational correlation time in ns; MW in kDa). The relationship between τ_c and MW was derived from 16 monomeric proteins of known size and rotational correlation time.^[31]

ITC: ITC was performed on a VP-ITC microcalorimeter (Malvern Instruments). The final concentration of PDZ3 was between 150 and 180 μM and the final ligand concentration was 390 μM in buffer containing 50 mM sodium phosphate at pH 6.8. For experiments with ligand L_{5A}, 2% DMSO was added. The sample was degassed before measurements and the temperature was set to 25 °C. ITC data was analysed with the MicroCal ITC analysis software in Origin 7, by assuming a one site binding model.

CSPs: Chemical shifts of the backbone amide resonances were compared by calculating the CSPs from the wild type to the mutant structure with Equation (2):

$$\text{CSP} = \sqrt{(0.1\delta_{15\text{N}}^{\text{PDZ3X}} - 0.1\delta_{15\text{N}}^{\text{PDZ3Y}})^2 + (\delta_{\text{H}}^{\text{PDZ3X}} - \delta_{\text{H}}^{\text{PDZ3Y}})^2} \quad (2)$$

in which $\delta_{\text{H}/15\text{N}}$ describes the chemical shift in ppm (either ¹H or ¹⁵N, respectively). PDZ3X and PDZ3Y labels data are derived from the two data sets being compared.

Crystallisation of PDZ3_{I341AZH} and data collection: Purified protein fractions were concentrated to 15 mg mL⁻¹ in buffer containing 10 mM HEPES/NaOH at pH 7.2, 10 mM NaCl, 0.5 mM dithiothreitol (DTT). Crystals of PDZ3_{I341AZH} were grown at 291 K by vapour diffusion of sitting drops equilibrated against a 50 μL reservoir of 1.1 M sodium citrate and 0.1 M HEPES at pH 8.5. Drops (2 μL) were set up at a 1:1 ratio of reservoir to protein solution. After 2 weeks, crystals were harvested by using mounted cryoloops (Hampton Research), equilibrated for 5 min in mother liquor containing 20% glycerol (v/v) as cryo-protectant and flash-cooled in pucks immersed in liquid nitrogen for storage. Crystallisation of PDZ3_{I341AZH} in complex with ligand L_{A5} through co-crystallisation and soaking failed because crystals did not grow in the sitting droplets and crystals of the apo protein dissolved after the ligand was added.

X-ray structure determination and refinement: X-ray diffraction data was collected at 100 K on beamline X06DA at the Swiss Light Source (SLS) operated by the Paul Scherrer Institute, Villigen, Switzerland. Data sets were processed by using XDSAPP.^[68] The structure was determined by molecular replacement by using Phaser, which facilitated the crystal structure of rat PDZ3_{WT} domain as a search model.^[33,69] Model building was performed by using Coot

and the structure was refined by using PHENIX.^[70,71] Figures containing molecular graphics were prepared by using PyMOL (Molecular Graphics System, Version 1.7.6.3 Schrodinger, LLC). Structure solution and refinement statistics can be found in the Supporting Information.

Acknowledgements

We thank Andreas Marx and Silvia Eger (University of Konstanz, Germany) for providing us with plasmids for the expression of PDZ3_{WT} and PDZ_{I341AZH}. R.S. is funded by a DFG research fellowship (SI2105/1-1). The BMRZ is supported by the state of Hesse. We are thankful for support from the Horizon 2020 program of the European Commission (iNEXT). We acknowledge the Paul Scherrer Institute, Villigen, Switzerland, for provision of synchrotron radiation beam time at beamline X06DA of the SLS.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: amino acids · NMR spectroscopy · proteins · structure elucidation · X-ray diffraction

- [1] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413.
- [2] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764–806.
- [3] A. Dumas, L. Lercher, C. D. Spicer, B. G. Davis, *Chem. Sci.* **2015**, *6*, 50–69.
- [4] K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19–24.
- [5] J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.* **2005**, *1*, 13–21.
- [6] D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell, E. M. Schuman, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9482–7.
- [7] G. Zhang, H. Bowling, N. Hom, K. Kirshenbaum, E. Klann, M. V. Chao, T. A. Neubert, *J. Proteome Res.* **2014**, *13*, 5707–5714.
- [8] R. B. Deal, J. G. Henikoff, S. Henikoff, *Science* **2010**, *328*, 1161–4.
- [9] S. Calve, A. J. Witten, A. R. Ocken, T. L. Kinzer-Ursem, *Sci. Rep.* **2016**, *6*, 32377.
- [10] Z. Li, Y. Zhu, Y. Sun, K. Qin, W. Liu, W. Zhou, X. Chen, *ACS Chem. Biol.* **2016**, *11*, 3273–3277.
- [11] H. Taskent-Sezgin, J. Chung, P. S. Banerjee, S. Nagarajan, R. B. Dyer, I. Carrico, D. P. Raleigh, *Angew. Chem. Int. Ed.* **2010**, *49*, 7473–7475; *Angew. Chem.* **2010**, *122*, 7635–7637.
- [12] R. Bloem, K. Koziol, S. A. Waldauer, B. Buchli, R. Walsler, B. Samatanga, I. Jelesarov, P. Hamm, *J. Phys. Chem. B* **2012**, *116*, 13705–13712.
- [13] H. M. Müller-Werkmeister, J. Bredenbeck, *Phys. Chem. Chem. Phys.* **2014**, *16*, 3261–6.
- [14] V. Botan, E. H. G. Backus, R. Pfister, A. Moretto, M. Crisma, C. Toniolo, P. H. Nguyen, G. Stock, P. Hamm, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12749–54.
- [15] F. Agostini, J.-S. Völler, B. Kokschi, C. G. Acevedo-Rocha, V. Kubyshekin, N. Budisa, *Angew. Chem. Int. Ed.* **2017**, *56*, 9680–9703; *Angew. Chem.* **2017**, *129*, 9810–9835.
- [16] D. B. Cowie, G. N. Cohen, *Biochim. Biophys. Acta* **1957**, *26*, 252–261.
- [17] G. N. Cohen, D. B. Cowie, *C. R. Hebd. Séances Acad. Sci.* **1957**, *244*, 680–3.
- [18] W. A. Barton, D. Tzvetkova-Robev, H. Erdjument-Bromage, P. Tempst, D. B. Nikolov, *Protein Sci.* **2006**, *15*, 2008–13.
- [19] W. A. Hendrickson, *Science* **1991**, *254*, 51–58.
- [20] H. M. Katzen, J. M. Buchana, *J. Biol. Chem.* **1965**, *240*, 825–35.
- [21] F. W. Studier, *Protein Expression Purif.* **2005**, *41*, 207–234.
- [22] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *99*, 201–207.
- [23] G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 7772–7777.
- [24] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* **1990**, *89*, 496–514.

- [25] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546–4553.
- [26] S. Grzesiek, J. Anglister, A. Bax, *J. Magn. Reson. Ser. B* **1993**, *101*, 114–119.
- [27] A. L. Davis, J. Keeler, E. D. Laue, D. Moskau, *J. Magn. Reson.* **1992**, *98*, 207–216.
- [28] M. Ikura, A. Bax, *J. Am. Chem. Soc.* **1992**, *114*, 2433–2440.
- [29] K. Ogura, H. Terasawa, F. Inagaki, *J. Biomol. NMR* **1996**, *8*, 492–489.
- [30] A. L. Breeze, *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 323–372.
- [31] P. Rossi, G. V. T. Swapna, Y. J. Huang, J. M. Aramini, C. Anklin, K. Conover, K. Hamilton, R. Xiao, T. B. Acton, A. Ertekin, J. K. Everett, G. T. Montelione, *J. Biomol. NMR* **2010**, *46*, 11–22.
- [32] C. M. Petit, J. Zhang, P. J. Sapienza, E. J. Fuentes, A. L. Lee, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18249–54.
- [33] D. A. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, R. MacKinnon, *Cell* **1996**, *85*, 1067–1076.
- [34] H.-J. Lee, J. J. Zheng, *Cell Commun. Signaling* **2010**, *8*, 8.
- [35] F. H. Niesen, H. Berglund, M. Vedadi, *Nat. Protoc.* **2007**, *2*, 2212–21.
- [36] R. Silvers, H. Keller, H. Schwalbe, M. Hengesbach, *ChemBioChem* **2015**, *16*, 1109–1114.
- [37] C. S. Johnson, Jr., *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203–256.
- [38] Y. Liang, *Acta Biochim. Biophys. Sin.* **2008**, *40*, 565–576.
- [39] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, S. Duhr, *Assay Drug Dev. Technol.* **2011**, *9*, 342–353.
- [40] R. Silvers, K. Saxena, D. Kudlinzki, H. Schwalbe, *Protein Expression Purif.* **2012**, *85*, 142–147.
- [41] B. Liedberg, C. Nylander, I. Lunström, *Sens. Actuators* **1983**, *4*, 299–304.
- [42] S. G. Patching, *Biochim. Biophys. Acta Biomembr.* **2014**, *1838*, 43–55.
- [43] M. Walker, J. G. Kublin, J. R. Zunt, *Clin. Infect Dis.* **2006**, *42*, 115–125.
- [44] A. Periasamy, *J. Biomed. Opt.* **2001**, *6*, 287.
- [45] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432–440.
- [46] S. Grzesiek, A. Bax, *J. Am. Chem. Soc.* **1992**, *114*, 6291–6293.
- [47] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213–223.
- [48] N. E. Hafsa, D. Arndt, D. S. Wishart, *Nucleic Acids Res.* **2015**, *43*, W370–W377.
- [49] A. G. Palmer III, *Chem. Rev.* **2004**, *104*, 3623–3640.
- [50] R. Silvers, F. Sziegat, H. Tachibana, S. Segawa, S. Whittaker, U. L. Günther, F. Gabel, J. Huang, M. Blackledge, J. Wirmer-Bartoschek, H. Schwalbe, *J. Am. Chem. Soc.* **2012**, *134*, 6846–54.
- [51] F. Sziegat, R. Silvers, M. Hähnke, M. R. Jensen, M. Blackledge, J. Wirmer-Bartoschek, H. Schwalbe, *Biochemistry* **2012**, *51*, 3361–72.
- [52] E. T. Olejniczak, R. X. Xu, S. W. Fesik, *J. Biomol. NMR* **1992**, *2*, 655–659.
- [53] A. Bax, G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* **1990**, *88*, 425–431.
- [54] T. M. Logan, E. T. Olejniczak, R. X. Xu, S. W. Fesik, *J. Biomol. NMR* **1993**, *3*, 225–231.
- [55] M. Montelione, G. T. Lyons, B. A. Emerson, D. S. Tashiro, *J. Am. Chem. Soc.* **1992**, *114*, 10974–10975.
- [56] D. Marion, P. C. Driscoll, L. E. Kay, P. T. Wingfield, A. Bax, A. M. Gronenborn, G. M. Clore, *Biochemistry* **1989**, *28*, 6150–6156.
- [57] G. Lippens, C. Dhalluin, J. M. Wieruszkeski, *J. Biomol. NMR* **1995**, *5*, 327–331.
- [58] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279–9283.
- [59] A. Saupe, G. Englert, *Phys. Rev. Lett.* **1963**, *11*, 462–464.
- [60] G. Otting, H. Senn, G. Wagner, K. Wüthrich, *J. Magn. Reson.* **1986**, *70*, 500–505.
- [61] G. Otting, K. Wüthrich, *J. Magn. Reson.* **1989**, *85*, 586–594.
- [62] G. Otting, K. Wüthrich, *Q. Rev. Biophys.* **1990**, *23*, 39–96.
- [63] C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, L. E. Kay, *J. Am. Chem. Soc.* **1997**, *119*, 6711–6721.
- [64] T. D. Goddard, D. G. Kneller, *SPARKY 3*, University of California, San Francisco.
- [65] P. Güntert, C. Mumenthaler, K. Wüthrich, *J. Mol. Biol.* **1997**, *273*, 283–298.
- [66] P. Dossset, J. C. Hus, M. Blackledge, D. Marion, *J. Biomol. NMR* **2000**, *16*, 23–28.
- [67] L. E. Kay, D. A. Torchia, A. Bax, *Biochemistry* **1989**, *28*, 8972–8979.
- [68] K. M. Sparta, M. Krug, U. Heinemann, U. Mueller, M. S. Weiss, *J. Appl. Crystallogr.* **2016**, *49*, 1085–1092.
- [69] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- [70] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D* **2010**, *66*, 486–501.
- [71] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221.

Manuscript received: August 14, 2017

Accepted manuscript online: September 26, 2017

Version of record online: November 7, 2017