SCUOLA NORMALE SUPERIORE
PISA
and
THE INTERNATIONAL CENTRE for GENETIC ENGINEERING
and BIOTECHNOLOGY (ICGEB)
TRIESTE

The Cytoplasmic Tail of the Notch Ligand Jagged-1: Intrinsic Disorder, Induced Order and Molecular Interactions

Thesis submitted for the Degree of Doctor of Philosophy
Perfezionamento in Genetica Molecolare e Biotecnologie

Candidate: **Matija Popović**
Supervisor: **Sándor Pongor, Ph.D., D.Sc.**

Academic Year: 2009-2010

Abstract

Notch signaling plays a key role in cell differentiation in all metazoans. As both receptors and ligands are cell-surface proteins, Notch signaling is restricted to nearby interacting cells. Notch ligands are membrane-spanning proteins made of a large extracellular region, a transmembrane segment, and a 100–200 residue cytoplasmic tail. The sequence of the intracellular region of Jagged-1, one of the five ligands to Notch receptors in man, is very well conserved throughout evolution but does not encode any globular domain. The cytoplasmic tail of Jagged-1 mediates protein–protein interactions through the C-terminal PDZ binding motif, is involved in ligand endocytosis triggered by mono-ubiquitination, and, as a consequence of regulated intramembrane proteolysis, can be released into the cytosol as a signaling fragment. The intracellular region of Jagged-1 may then exist in at least two forms: as a membrane-tethered protein located at the interface between the membrane and the cytoplasm, and as a soluble nucleocytoplasmic protein. To investigate its structural properties, a recombinant protein corresponding to the human Jagged-1 intracellular region (J1_tmic) was expressed, purified, and characterized in different environments using various biophysical methods such as circular dichroism, tryptophan fluorescence, size-exclusion chromatography, and NMR.

In solution, J1_tmic behaves as an intrinsically disordered protein, but displays a significant helical propensity. In the presence of SDS micelles or negatively charged phospholipid micelles and vesicles, used to mimick the interface between the plasma membrane and the cytosol, J1_tmic gains partial helical structure. The partial folding and association of the intracellular region of Jagged-1 with the membrane is expected to reduce its “capture radius” towards target proteins and to make selected residues unavailable for post-translational modifications or binding.

Binding of Jagged-1 intracellular region to the PDZ domain of afadin, a protein located at cell-cell adherens junctions, couples Notch signaling with the adhesion system and the cytoskeleton. The interaction between the PDZ domain of afadin (AF6_PDZ) and a series of polypeptides comprising the Jagged-1 PDZ-binding motif (EYIV) was investigated using NMR chemical shift perturbation and surface plasmon resonance. It was shown that binding of Jagged-1 intracellular region to

AF6_PDZ is strictly local, involving only the last six residues of the binding motif and the PDZ binding groove, and that it does not trigger global folding of J1_tmic.

In the C-terminal region of Jagged-1 cytoplasmic tail, four potential phosphorylation sites can be identified, one of them (Y1216) located in the PDZ-binding motif. It was found that, while phosphorylation at any of these sites disrupts binding of the C-terminal peptides to lipid micelles, phosphorylation at (Y1216) also affects the interaction with AF6_PDZ, with a reduction in the binding affinity. Phosphorylation thus provides a potential way to modulate the interaction of Jagged-1 C-terminal region not only with the membrane but also with the partner PDZ. It was also shown that the R1213Q mutation in the PDZ binding motif associated with a congenital obstruction of the bile ducts, increases the affinity for AF6_PDZ.

In summary, this work presents the first biochemical and structural characterization of Jagged-1 cytoplasmic tail in solution and in environments that mimic the membrane/cytoplasm interface, and the first biophysical study on its interaction with the afadin PDZ domain.

List of publications

Publications related to this thesis:

Popovic M, Bella J, Zlatev V, Hodnik V, Anderluh G, Barlow PN, Pintar A, Pongor S. The interaction of Jagged-1 cytoplasmic tail with afadin PDZ domain is local, folding independent, and tuned by phosphorylation. *J. Mol. Recognit.* (2010), in press.

De Biasio A, Guarnaccia C, Popovic M, Uversky VN, Pintar A, Pongor S. Prevalence of intrinsic disorder in the intracellular region of single-pass transmembrane proteins: the case of the Notch Ligand Delta-4. *J. Proteome Res*. (2008) 7:2496–2506

Pintar A, De Biasio A, Popovic M, Ivanova N, Pongor S. The intracellular region of Notch ligands: does the tail make the difference? *Biol. Direct.* (2007) 2:19.

Popovic M, De Biasio A, Pintar A, Pongor S. The intracellular region of the Notch ligand Jagged-1 gains partial structure upon binding to synthetic membranes. *FEBS J.* (2007) 274(20):5325-36.

Popovic M, Coglievina M, Guarnaccia C, Verdone G, Esposito G, Pintar A, Pongor S. Gene synthesis, expression, purification, and characterization of human Jagged-1 intracellular region. *Protein Expr. Purif.* (2006) 47:398–404.

Acknowledgments

This thesis was prepared in the Protein Structure and Bioinformatics Group at the International Centre for Genetic Engineering and Biotechnology (ICGEB) in Trieste, Italy.

First, I am very grateful to my supervisor, Sándor Pongor, for giving me the opportunity to carry out my work in his lab. I want to thank him for his support and helpful tips, which were necessary for the completion of this thesis.

Special gratitude goes to Alessandro Pintar who supervises the studies on Notch signaling within the lab, for his continuous help and sincere patience in taking me throughout the entire project. Thanks to him, now I can say that I am much closer to becoming an independent researcher in the field of structural biology and protein NMR spectroscopy.

I am very grateful to all the people of the Protein Structure and Bioinformatics Group that contributed to this thesis: Maristella Coglievina, who trained me in the molecular biology techniques, Corrado Guarnaccia for continuous help with HPLC, mass spectrometry, peptide synthesis, and peptide and protein purification, Ventsislav Zlatev for the synthesis and purification of many peptides, and Sotir Zahariev for helpful advice in the peptide chemistry field.

I am also grateful to all external collaborators that contributed, directly or indirectly, to this thesis: Doriano Lamba (CNR-ELETTRA, Basovizza, Italy) for extensive use of the CD spectropolarimeter, Fabio Calogiuri (CERM, Sesto Fiorentino, Italy) and Nicola D'Amelio (CBM, Basovizza, Italy) for technical assistance with the acquisition of NMR spectra, Juraj Bella, Paul Barlow and Dusan Uhrin (Biomolecular NMR Unit, University of Edinburgh, UK) for technical assistance and support in the acquisition and processing of the NMR spectra acquired in Edinburgh, Vesna Hodnik and Gregor Anderluh (Biotechnical Faculty, University of Ljubljana, Slovenia) for technical assistance and support in the SPR experiments.

I also acknowledge the EU (European Network of Research Infrastructures for Providing Access and Technological Advancements in Bio-NMR) for providing access to the CERM NMR facility, EMBO (European Molecular Biology Organization) for a Short Term Fellowship, and ICGEB for support with a pre-doctoral fellowship.

I want to thank my friend and colleague Vjeko, with whom my work at the institute became more interesting and amusing.

My greetings go to the young Buddhist group of Trieste, with whom I felt at home and started to love this town.

I am taking this opportunity to express my deepest gratitude to the most important people in my life:

my parents,

my sister,

and Vanessa.

This thesis is dedicated to them and to Josei Florit.

Table of Contents

1. INTRODUCTION

1.1 Notch signaling

1.1.1 Molecular mechanisms of Notch signaling

Notch signaling mediates cross-talk between two contacting cell. The main components in this signaling pathway are Notch receptors and their corresponding ligands, which are both membrane-bound proteins [1], [2] (**Figure 1.1**). Notch receptors, four members of which have been identified in man (Notch-1, -2, -3 and -4), are membrane-spanning glycoproteins assembled in a non-covalent heterodimeric complex. After their expression, Notch receptors are proteolytically processed (at the S1 cleavage site) by a furin-like convertase in the trans-Golgi to give an extracellular (N_ec) and a transmembrane subunit (N_tm) [3]. The N_ec contains an array of up to 36 EGF tandem repeats followed by three LIN-12 repeats that maintain Notch in a resting conformation before ligand-induced activation [4]. The EGF 11-12 couple in Notch receptor constitutes the binding site for the ligand DSL domain [5]. The N_tm has a short extracellular region, involved in the dimerization with the N_ec, and a ~900 residue long intracellular region (N_ic). The second one consists of a one RAM domain followed by seven ankyrin repeats with two lateral nuclear localization signals, a transactivation domain (TAD), and a C-terminal PEST region. Upon ligand binding, an extracellular ADAM type metalloproteinase cleaves the N_tm at the S2 site; 12 residues before the membrane-spanning segment. This removes the N_ec from the membrane, rendering the N_tm subunit susceptible to the cleavage at the S3 site, close to the inner leaflet of the plasma membrane. This proteolytic processing is mediated by the presenilin/γ-secretase complex and releases the intracellular domain (N_ic) from the membrane into the cytoplasm. This mechanism is known as "regulated intramembrane proteolysis" (RIP) [6] and is also present in other signal transduction pathways, such as those mediated by CD44, LRP, ErbB-4 and others [7]. Once released into the cytoplasm, N_ic translocates to the nucleus and associates with the CSL family of DNA-binding proteins (CBF-1/suppressor of Hairless/Lag-1) [8].

This association, along with the recruitment of Mastermind and other co-activators, converts the CSL-containing complex from a transcriptional repressor to a transcriptional activator of target genes. The best known Notch target genes are the

basic helix-loop helix (bHLH) transcriptional repressors of the HES and HEY families [8], [9]

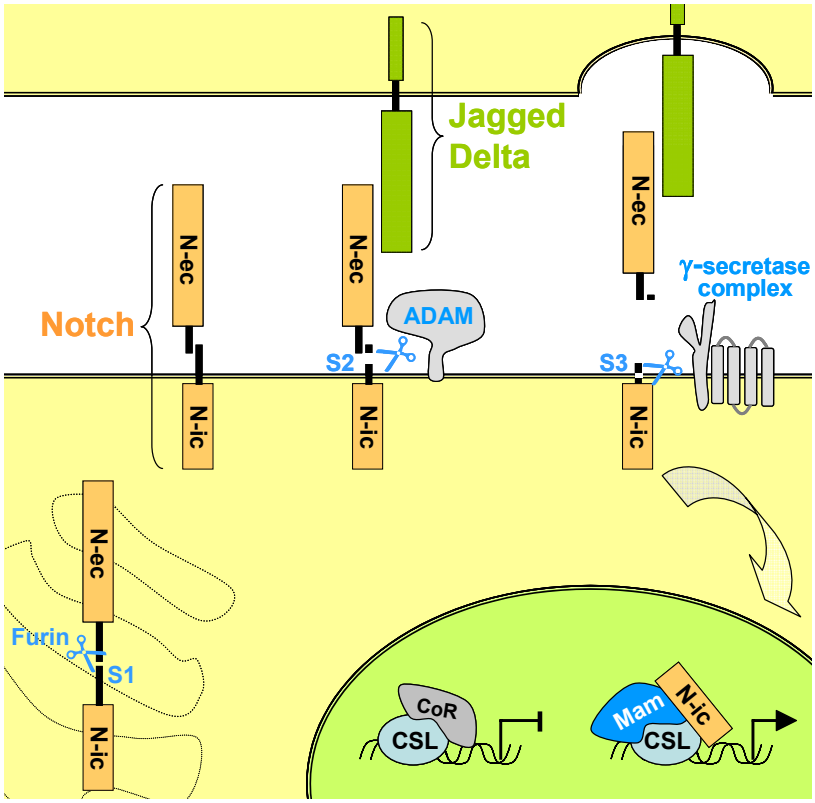


Figure 1.1. The Notch signaling pathway. Notch receptors are cleaved at the S1 site by furin, to be presented as non-covalent heterodimers at the cell surface (N-ec and N-ic, extra-and intracellular region of the receptor, respectively). Upon ligand binding, the receptor is cleaved by ADAM and γ-secretase at the S2 and S3 sites, respectively. In the nucleus, N-ic displaces the co-repressor (CoR) to form a transcriptional activator shown as a complex of N-ic, mastermind (MAM) and CSL bound to DNA.

These are basic helix-loop helix (bHLH) DNA binding proteins that, acting within a complex and strongly context dependent transcriptional network, interact with other bHLH proteins, bind to specific gene promoter regions, and, through the recruitment of transcriptional corepressors, inhibits transcription of the target genes. [10],[11]. Several DNA target sequences have been identified so far, including *HES1* own promoter, the promoter regions of the transcription factor Achaete Scute Homolog 1 (*ASH1*) [12], and the promoters of enzymes involved in cell cycle, like the cyclin-dependent kinase inhibitors *p21*WAF-1/CIP1 and *p27*kip1 [13].

In man, there are two families of Notch ligands, homologues of *Drosophila* Serrate (Jagged-1 and -2) and homologues of *Drosophila* Delta (Delta-like-1, -3 and -4) [14] (**Figure 1.2**). All five ligands share the same architecture and belong to the so called DSL family (Delta/Serrate/Lag-2, the latter being the *C. elegans* ligand). They are type I membrane spanning proteins consisting of a long extracellular region, a transmembrane segment and a relatively short cytoplasmic tail (100-150 amino acids). The ligand extracellular region has an N-terminal DSL domain followed by a variable number of EGF-like repeats. Ligands of the Jagged group have an additional cysteine rich region close to the transmembrane segment, which is not present in the Delta group ligands.

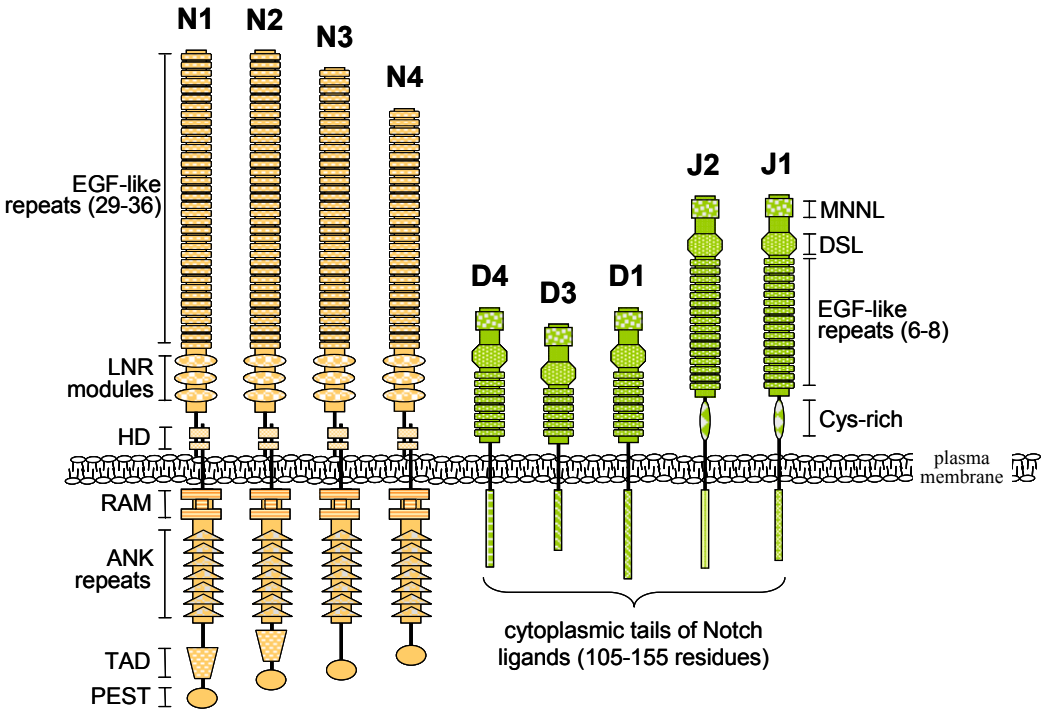


Figure 1.2. Domain organization of Notch receptors and ligands in man. The four Notch receptors (Notch-1 (N1), Notch-2 (N2), Notch-3 (N3) and Notch-4 (N4)) and five Notch ligands (Delta-like-4 (D4), Delta-like-3 (D3), Delta-like-1 (D1), Jagged-2 (J2) and Jagged-1 (J1)) are schematically shown. EGF, Epidermal Growth Factor repeat; LNR, Lin-Notch repeat; HD, hetero-dimerization region; RAM, RBP-Jkappa-associated module; ANK, Ankirin repeat; TAD, transcription activation domain; PEST, region containing P, E, S, T residues; MNNL, domain Mainly found in the N-terminal region of Notch Ligands; DSL, Delta-Serrate-Lag2 domain; Cys-rich, cysteine rich region.

1.1.2 Notch and cell-fate decisions

Development of a multicellular organism requires cell-fate decisions that are intimately associated with the determination of a cell-specific gene expression pattern. Notch signaling, having a direct effect on gene expression, acts in various development decisions either through lateral-inhibition or induction [15] (**Figure 1.3**). Lateral inhibition occurs between two equipotent neighbouring cells, which can signal one to another, as ligands and receptors are expressed in both of them. Due to fluctuations in the steady-state expression levels, the concentrations of these proteins between neighbouring cells start to differ. In cells with the higher receptor activation, the Notch feedback loop slightly inhibits the ligand expression. In this way, the initial small differences are amplified, which drives neighbouring cells into different developmental fates. Inductive signaling occurs when two non-equivalent groups of cells confront one another. Keeping the two populations distinct, Notch signaling is involved in establishing boundaries at the interface between them. In addition to this inductive role, Notch signaling also plays an important role in lineage decision between two daughter cells. A typical example is the development of sensory bristles in *Drosophila,* where sensory organ precursors (SOP) are initially selected from equivalent pro-neural cells by the Notch lateral inhibition [16]. SOP cells undergo multiple asymmetric cell divisions in which the Notch antagonist, Numb, is asymmetrically distributed. In each division step, only one daughter cell holds the Notch activity, eventually giving rise to a glial cell on one hand and cells that form the bristle on the other.

Notch signaling in development. Notch receptors and ligands are widely expressed during organogenesis in mammalian embryos, where they play a key role in establishing cell-lineage decisions in tissues derived from all the three primary germ layers: the endoderm (for ex. the pancreas), the mesoderm (skeleton, mammary gland, the vascular system and hematopoietic cells), and the ectoderm (neuronal cell lines) [17]. In the pancreas, where different cell types appear with different timing, yet stemming from the same early cells, Notch-1 appears to delay both endocrine and exocrine development, trapping progenitor cells in an undifferentiated state. In the presomitic mesoderm that will differentiate into the axial skeleton, muscles, tendons and dermis, Notch signaling plays a role as a molecular clock that controls regular segmentation of the mesoderm. Notch is also required in the later steps of vascular development, which includes proliferation and branching of the newly formed vessels

[18]. In the hematopoietic system, enforced activation of Notch-1 suppresses the differentiation of stem cells into myeloid, erythroid, or lymphoid lineages, and plays a role at a number of stages of lymphocyte development in the bone marrow and thymus. One of the essential functions of Notch-1 is the suppression of B cell development in the thymus. In the nervous system, Notch activation is required for the self-renewal of neural stem-cells, although it is not necessary for their generation [19], [20]. Furthermore, Notch signaling controls the differentiation of glial cells and the length and organization of dendritic extensions from neurons (neurite arborization).

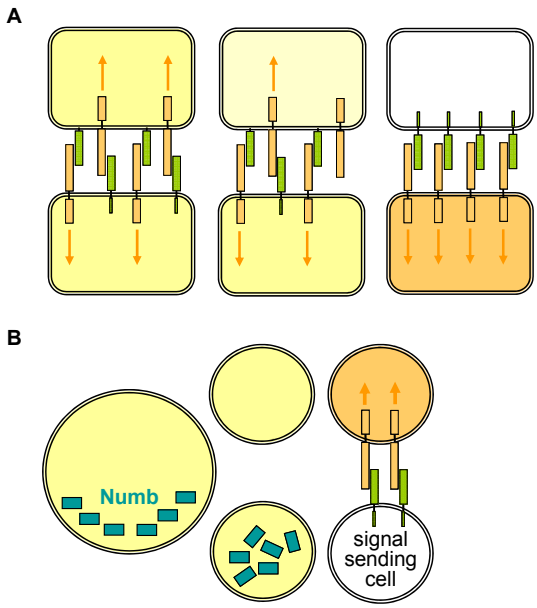


Figure 1.3. Notch signaling in cell-fate decisions. Receptors are colored in orange, ligands in green, and orange arrows represent the Notch signaling direction in (**A**) lateral inhibition and (**B**) line lineage decision between two daughter cells in the development of sensory bristles in *Drosophila*. The intensity of the orange color is proportional to the receptor activation.

Notch signaling in cancer. At least two direct links between alterations in Notch signaling and human cancer have been established to date. A rare form of T cell acute lymphoblastic leukemia (T-ALL) is associated with a translocation that fuses the intracellular portion of Notch-1 with the promoter/enhancer region of the T-cell receptor beta locus, leading to constitutive activation of Notch-1 signaling [21]. The majority of T-ALL cases have been recently associated with activating mutations

in Notch-1 [22]. Another chromosomal translocation, which is altering the function of Mastermind, a nuclear regulatory protein in the Notch signaling pathway, has been linked to mucoepidermoid carcinoma, a common type of malignant salivary gland tumor [23]. High levels of Delta-1 have been observed in neuroblastoma cell lines. High expression levels of Notch have also been reported in some breast cancers and in human colon adenocarcinomas. Intriguingly, Notch can behave both as an oncogene or a tumor suppressor, depending on the cellular context and on the interactions with other signaling pathways [14].

Notch signaling in genetic disorders. The importance of the Notch pathway in cell fate control and development is further confirmed by the association of several diseases with mutations in genes involved in this complex signaling network [24]. Alagille syndrome [25] (AGS, MIM #118450) is a rare autosomal dominant disorder characterized by a variety of clinical abnormalities, including a reduction in the number of bile ducts eventually leading to the obstruction of biliary flow, and cardiac, musculoskeletal, ocular, facial defects. Although no clear genotype-phenotype correlation has been defined, AGS is caused by mutations in *JAG1* [26], [27]. While the majority of the mutations causing AGS are related to the generation of stop codons leading to unstable mRNA or truncated proteins, many missense point mutations either introduce or delete cysteine residues that are critical for proper folding of the mature protein. Most of these mutations are located in the DSL domain and in the EGF tandem repeats. Similarly to AGS, but with a different pathogenesis, extrahepatic biliary atresia (EHBA) is related to mutations in *JAG1* and recognized as abnormalities in the development of the bile duct [28]. In the EHBA patients, besides mutations in the extracellular region, one missense mutation was also detected in the cytoplasmic tail of Jagged-1 (R1213Q). Familial tetralogy of Fallot (TOF, MIM #187500) is the most common form of complex congenital heart disease (~1/3000 births). It is characterized by ventricular septal defects, obstruction to right ventricular outflow, aortic dextroposition and right ventricular hypertrophy. A familial form of TOF was found to be associated with a missense G274D mutation [29], [26] occurring in the second EGF repeat of Jagged-1.

Spondylocostal dysostosis (SD, MIM #277300) is a vertebral malsegmentation syndrome characterized by multiple hemivertebrae, rib fusions and deletions. Mutations correlated with autosomal recessive SD have been identified in *DLL3* [30]. Two of these mutations are expected to lead to truncated forms of the protein, while

the third is a missense mutation in one of the EGF tandem repeats, G385D. Interestingly, this is the same kind of mutation observed in *JAG1* and for which the genotype has been correlated to the TOF phenotype.

Cerebral autosomal dominant arteriopathy, with subcortical infarcts and leukoencephalopathy [31] (CADASIL, MIM #125310) is associated with strokes and dementia. It is caused by mutations in the Notch-3 member of the Notch receptor family [23]. Most of the mutations involve the removal or insertion of cysteine residues in the EGF repeats and are likely to affect receptor folding, trafficking, maturation, or signaling.

1.1.3 Notch signaling regulation

Notch signaling controls cell differentiation, proliferation and survival and is used in a diversity of developmental processes. Despite the relatively simple core machinery of its signaling pathway, Notch has a very complex regulation system. The signal transduction between two neighbouring cells implies the necessity of regulation at the extra- and intra-cellular level. Notch is regulated at three major levels, glycosylation and cleavage control at the extracellular level, endosomal trafficking at the cytoplasmic level and transcriptional modulation in the nucleus (**Figure 1.4**).

Glycosylation. Glycosylation of the Notch extracellular domain is essential for its activity and can also regulate its binding affinity for different Notch ligands. Two different glycosylation modes were found in Notch receptors, O-linked fucosylation and O-linked glucosylation [32]. The O-linked fucosylation has a role in the modulation of Notch signaling. The O-linked glucosylation seems to be involved in the folding and the trafficking of Notch receptors, but its real function still has to be established. There are two enzymes involved in O-linked fucosylation of Notch receptors, *O*-FucT-1 and Fringe. The *O*-FucT-1 adds the first *O*-fucose on specific Serine/Threonine sites of EGF repeat, while Fringe uses fucose-*O*-EGF as a substrate for further glycosylation. *O*-fuct-1 is essential for Notch signaling as proved by mutation and knock-out experiments in *Drosophila* and mice (they all result in loss of Notch signaling) [33]. In *Drosophila*, Fringe regulates the Notch interaction with two different classes of ligands. It facilitates Delta and inhibits Serrate activation of Notch receptor. The *O*-fucosylation of Notch takes place in the Golgi and influences the receptor-ligand binding either by conformational changes of the N_ec or by creating recognition sites for ligands.

Notch ligands are also modified by *O*-fucose glycans, which is not surprising since they possess conserved *O*-fucosylation motifs in their EGF modules. However, there is no direct evidence for any functional relevance of these modifications. In vertebrates, the biological function of Fringe is more difficult to assess, as there are three different enzymes (Lunatic-, Manic- and Radical-Fringe) and several Notch proteins (four receptors and five ligands). The present experimental results suggest that the influence of Fringe on Notch signaling depends not only on the ligand, but also on the receptor. The expression of different Fringes is cell-type dependent and they glycosylate Notch receptors in a different way, indicating that the distribution of different Notch forms is determined by transcriptional regulation of fringe genes.

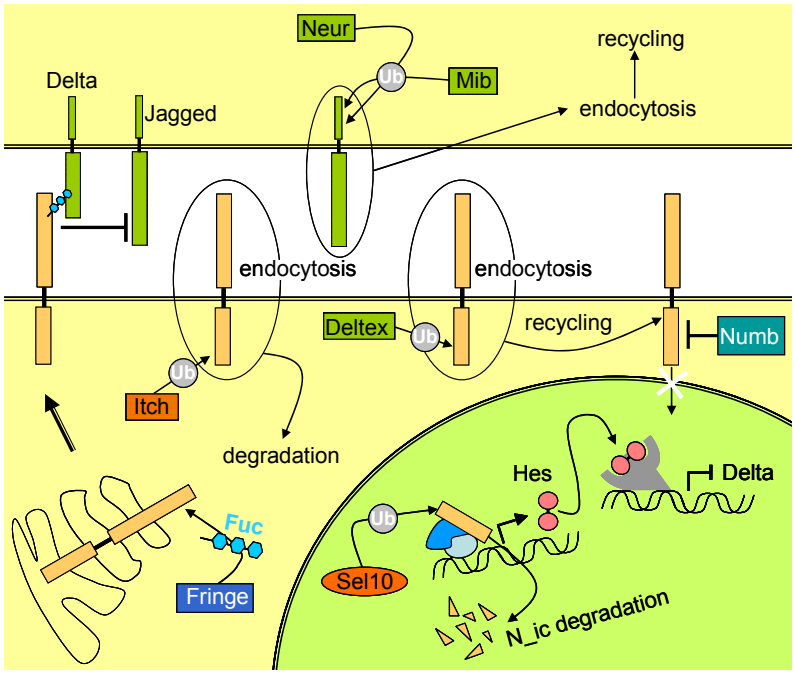


Figure 1.4. Regulation of Notch signaling. For simplicity, the signal sending and the signal receiving cells are shown as presenting only ligands (green) or only receptors (orange) at the cell surface, respectively. Ubiquitine ligases that up-regulate Notch signaling through recycling of receptors and ligands (Neur, Neuralized; Mib, Mind Bomb; Deltex) are in green. Ubiquitine ligases that down-regulate Notch signaling through degradation of receptors (Itch and Sel10) are in dark-orange. Ub, Ubiquitine (grey); Fuc, fucose (light blue). In the cell nucleus (green), CSL and Mastermind co-activators are colored in blue.

Endosomal trafficking. The endocytic regulation is essential for Notch signaling and both ligands and receptors are involved. It was shown that ligand endocytosis promotes activation of Notch signaling. The cytoplasmic tail of Notch ligands is ubiquitinated by E3 ligases (Mind Bomb, Neuralized 1 and 2 in mammals) and consequently internalized by endocytosis. The role of ligand endocytosis can be explained by three general models based on experimental data. Two of these models consider ligand recycling as the main role of endocytosis. In the first one, the ligand becomes activated by post-translational modifications during the endosomal trafficking and than recycled back to the cell surface. In the other model, recycling promotes ligand clustering, which was shown to be required to activate signaling [34]. In the third model, the internalization of the ligand bound to the receptor creates a pulling force on the Notch extracellular domain (N_ec) (**Figure 1.1**). Trans-endocytosis of N_ec induced by the ligand in *Drosophila* [35] and man [36] are in support of this theory. It was proposed that after the cleavage at the S2 site, the Notch heterodimer is mechanically dissociated allowing its S3 site to be unmasked and thus more accessible for the final cut. Later, it was found that the N_ec trans-endocytosis occurs before the proteolysis at S2, indicating that the separation of N_ec is a prerequisite for both proteolytic processes and thus necessary in Notch activation.

Endocytosis of Notch receptors can promote either negative or positive regulation of Notch signaling, depending on the context [37]. Unlike the ligands, endocytosis of the receptors generally down-regulates Notch signaling. Nedd4, Itch/AIP4 and Cbl are E3 ligases which target Notch to lysosomal degradation and thus regulate its basal activity on the cell surface. Deltex is another E3 ligase involved in Notch endosomal trafficking, and can behave as positive or negative Notch regulator, depending on the cell environment. In *Drosophila*, Deltex protects Notch receptors from entering the degradation pathway, possibly by promoting sorting of the receptor in endosomal compartments [38]. Endocytosis can regulate Notch activity also in an indirect way. One example is the interaction between two Notch E3 ligases, in which Nedd4 marks Deltex for lysosomal degradation, showing once more its antagonism towards Notch activity. The other case involves Numb, a membrane associated protein and a negative regulator of Notch signaling. Numb actually removes Sanpodo from the membrane, a protein which plays a yet unknown, but essential role in Notch signaling.

Transcriptional modulation. Once located in the nucleus, the Notch intracellular domain (N_ic) can switch on its target genes. N_ic binds to CSL, a DNA binding protein which actually confers specificity to the expression of Notch target genes. In the absence of N_ic, CSL recruits various co-repressors like SMRT, SHARP, CtBP, SKIP and CIR [39]. Moreover, to keep the chromatin in a silent mode, these co-repressors further associate with histone deacetylases. The binding of N_ic to CSL removes the co-repressors and recruits the co-activator Mastermind. This ternary complex recruits other co-activators such as p300, SKIP, the histone acetyl-transferase GCN5, the chromatin-remodelling enzyme Brahma and others. This activator complex has a short life because Mastermind and SKIP recruit the CDK8 kinase which specifically phosphorylates N_ic rendering it a substrate for the nuclear ubiquitin ligase SEL10. Once ubiquitinated, N_ic is targeted for degradation, which is a very common mechanism for the turnover of nuclear effectors. It was shown that N_ic is necessary in small concentrations and has a short half-life in the nucleus. This is to be expected considering the dynamics and constant activity of Notch signaling during the cell lifetime. The most commonly induced Notch target genes are proteins of the Hes and Hey families, which belong to the basic helix-loop-helix (bHLH) class of transcriptional repressors. These proteins regulate the maintenance of stem cells and binary cell-fate decision in the development of many organs. In *Drosophila*, one of HES repressor's targets is the Delta locus, which explains the mechanism of the negative feed-back loop in the above mentioned Notch lateral inhibition. Beside transcriptional repressors, N_ic/CLS can induce genes of cyclin D1 [40] and p21 [41], indicating the direct impact of Notch on the cell-cycle progression. Many of the above mentioned transcriptional modulators are tissue-specific and can have cross-talk with other signaling systems.

1.2 Jagged-1

1.2.1 Jagged-1 and other Notch ligands

Five different Notch ligands, divided into two families, and four different Notch receptors have been identified in man. *In vivo* experiments showed that Notch receptors and ligands can perform distinct, but also overlapping functions. This may occur in a different, but also in the same developmental context. For example, Notch-1 and Notch-4 can play partially redundant functions during embryonic vascular

development, while Notch-3 is unable to compensate for the loss of Notch-1 during T cell development [14]. Delta-1 is involved in T cell development, whereas Jagged-1 has influence on peripheral T cell differentiation [42]. On the other hand, the above mentioned ligands seem to have identical signaling pathways, and thus redundant role, during the osteoprogenitor cells differentiation in bone regeneration [17]. Many ligand knock-out experiments showed developmental defects and embryonic lethality in mice, indicating a unique role of each Notch ligand [43], [44], [45].

The molecular basis for ligand specificity is still awaiting to be discovered and described. The two ligand families display some differences in the domain architecture of the extracellular region. Yet, it is not known how these differences may affect the binding specificity for the receptors, nor is the binding mechanism well determined. In *Drosophila,* the DSL domain is required for the binding to Notch, but it was shown that the whole ligand sequence is required for the transmission of a specific signal by a given ligand [46].

Interestingly, the amino acid sequence of ligand intracellular regions is evolutionary well conserved within the same ligand type and, at the same time, different ligand types display different sequences [47], [48]. This fact provides new opportunities to study specificities of ligand-receptor interactions, but also opens up a new view on the Notch signaling in general.

1.2.2 The intracellular region of Notch ligands

While the role of the intracellular region has been carefully investigated in Notch receptors, we are still at the beginning in the understanding of the function of the ligands cytoplasmic tail. The cytoplasmic tail of Notch ligands is involved in ligand endocytosis, acts like a signaling fragment in the ligand bearing cell and is a mediator in the cross-talk between Notch and other signaling pathways (**Figure 1.5**). The cytoplasmic tail of all Notch ligands, with the exception of Delta-3, contains multiple lysine residues that can serve as potential sites for the ubiquitination by E3 ligases. Indeed, these ligands are ubiquitinated by Neuralized (Neur) and Mind bomb (Mib) E3 ligases, and then internalized by endocytosis. The different ubiquitination states of ligands mediated by these structurally distinct E3 ligases can explain their different roles in Notch signaling. It was demonstrated that monoubiquitination of Jagged-1 by Neur1 leads to the ligand degradation and thus attenuation of Notch signaling induced by Jagged-1 [49]. On the other hand, the ubiquitination of Jagged-2

by Mib2 seems to be associated with activation of Notch signaling [50]. It can be speculated that the differences in the cytoplasmic tails of Notch ligands may have an impact on the specificity of the regulation mechanism.

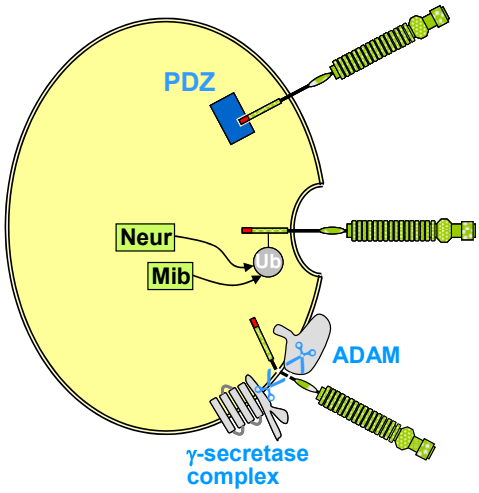


Figure 1.5. Role of the ligand intracellular region. From top: interaction with PDZ domain proteins (PDZ binding motif in red); endocytosis upon ubiquitination by Neuralized and Mind bomb; regulated intramembrane proteolysis and release of the ligand intracellular region in the cytoplasm.

Notch ligands undergo a proteolytic processing with the consequent release of their intracellular regions from the cell-membrane. The cleavage of rat Jagged-1 by metalloprotease ADAM 17 was demonstrated by *in vitro* experiments in cultured cells, and further proven by *in vivo* experiments [51]. Rat Jagged-1 is also subjected to intramembrane proteolysis carried out by the presenilin/γ-secretase complex. The released intracellular region (rat_J1_ic) was found in the cytoplasm and in the nucleus. The intramembrane cleavage site, even if not precisely determined, was proposed to be placed at the first valine residue close to the cytoplasmic region, in analogy with the position of the cleavage site found in Notch receptors [51]. Cytoplasmic tails of Notch ligands contain positively charged residues that could function as nuclear localization signals (NLS) [52]. In each of the two ligand families, this NLS sequence is conserved from flies to man. In the Jagged family, the KRRK amino acid sequence is located in the N-terminal region of the cytoplasmic tail. It was

shown that rat J1_ic cannot enter into the nucleus without this localization motif. In co-transfection studies, it was demonstrated that rat J1_ic specifically binds the transcription factor AP1 (Activator Protein 1, p39 *jun*) and thereby activates transcription of the reporter gene in COS, CHO, and HEK cells [51]. Evidences of the above described ligand processing were found also for human Jagged-2 [53] and mouse Delta-1 [54]. When released from the cell membrane, the intracellular region of Delta-1 (D1_ic) translocates to the nucleus, where it specifically binds to the Smad transcription factors. This binding enhances the transcription of genes involved in neuronal differentiation of mouse neuronal stem cells (NSC) [55]. Moreover, the development of neurons from NSC was enhanced by co-culture with Notch1-expressing cells, and depends on presenilin activity. These results indicate that the Delta-1 processing is enhanced through the interaction of the neighbouring cell with Notch-1, implying the existence of a bi-directional signaling pathway. Notch-related signal transduction pathways are thus active not only in the receptor bearing cell, but also in the ligand bearing one. The molecular mechanism of the latter, however, remains largely uncharacterized, and its role in Notch signaling feed-back and cell differentiation is still unknown.

Sequence alignments of the Notch ligand cytoplasmic tails show evident lack of homology between different ligands, but high evolutionary conservation within the same ligand type [48]. The Jagged-1 cytoplasmic tail has a remarkably well conserved C-terminal region (last 25 residues) while the same is not true for Jagged-2. The C-terminal region of Delta-1 and Delta-4 is evolutionary more conserved than in Delta-3. The cytoplasmic tails of Jagged-1, Delta-1 and Delta-4 contain a PDZ binding motif at their C-terminus. Jagged-1, with its C-terminal EYIV sequence, was shown to interact with the PDZ domain of the protein AF6 in a PDZ-dependent manner [56], [57]. Delta ligands share the same PDZ recognition motif (ATEV) at the C-terminus, and were shown to bind PDZ domains of Dlg1 [58] and MAGI proteins [59]. Jagged-1, on the contrary, does not bind to the PDZ-domain of MAGI-1 and Dlg1 [58], [60]. AF6, together with E-cadherin/catenin belongs to an adhesion system that plays a role in the organization of cell-cell junctions [61]. Dlg1 is a membrane-associated guanylate kinase involved in the maintenance of cell adhesion, cell polarity, growth control and cell invasion, and is essential for the assembly of multiprotein complexes at cell-cell junctions. The presence of PDZ binding motifs, together with the experimentally confirmed interactions with PDZ containing proteins, suggest that

Notch ligands are involved in a cell-autonomous, Notch-independent signal transduction pathway or that Notch signaling is coupled to other signaling networks [57]. Delta-1 reduces the motility of 3T3 cells and was shown to co-localize with Dlg1 at cell-cell contacts [58]. The Delta-1 mutant lacking its PDZ-recognition motif does not show the above characteristics, but still maintains its ability to activate the Notch signaling pathway. These results point to a PDZ-dependent activity of Delta-1 in the regulation of cell-adhesion.

1.3 Jagged-1 and AF6

The first evidence of the interaction between the C-terminal hexapeptide (RMEYIV) of Jagged-1 intracellular region (J1_ic) and the PDZ domain of AF6 was found using a yeast-two-hybrid assay [56]. The interaction was then confirmed in pull down experiments [57]. Expression of human Jagged-1 in RKE cells (rat kidney epithelial cells) mediates their neoplastic transformation in a dose-dependent manner. Interestingly, deletion of the Jagged-1 C-terminal PDZ-recognition motif results in a loss of its transformation activity, yet does not impair its ability to induce Notch receptors in the neighbouring cells (a coculture luciferase assay). Additionally, human Jagged-1 up-regulates gene transcriptions of Jagged-1 and Delta-1, as it was seen by the increase of the endogenous rat mRNA of these two ligands in RKE transformed cells. The induction of the Jagged-1 mRNA is due to the transcriptional activation of its gene, as it was demonstrated by luciferase assays using *JAG1* promoter reporter constructs. Jagged-1 with a deleted PDZ-binding motif does not increase the endogenous mRNA level of Jagged-1 and Delta-1 neither activates *JAG1* promoter. Moreover, the blockage of the PDZ-binding motif by the myc-tag on its C-terminus abrogates the promoter activation as well, indicating that signaling downstream Jagged-1 involves PDZ-domain proteins. Additionally, the Jagged-1 promoter could not be activated by the overexpression of the Notch receptor intracellular region. These results indicate that the above cellular transformation is a consequence of, at least in part, a PDZ-dependent signaling pathway intrinsic to Jagged-1 and probably independent of Notch activation. The main question is how this interaction between Notch ligands and specific PDZ proteins affects gene expression. Even if interactions with PDZ proteins mainly occur at the interface between the membrane and the cytosol, there are some PDZ-containing proteins acting like transcriptional activators

(CASK, Bridge-1 or GRIPtau) [62], [63], [64]. However, Notch ligands could affect the expression of the target genes indirectly, via signal transducers, without being released from the cell surface. One example is the binding of Delta-1 to Acvrinp1. Acvrinp1 is a PDZ containing protein that binds to Smad3, inhibiting in that way the Smad3-dependent transcription [65]. Moreover, AF6 interacts with RAS [66], [57] which further activates signaling to the nucleus promoting changes in gene expression. This suggests that Jagged-1 can indirectly transduce the signal to the nucleus. How the RIP mechanism of proteolytic cleavage occurring in Jagged and Delta proteins can affect their interaction with the partner PDZ proteins remains unknown, as well as the role of Notch receptors in these interactions.

1.3.1 PDZ domains

PDZ domains, one of the most abundant protein interaction modules in metazoans, play an essential role in the control of cell signaling. The PDZ acronym comes from the initials of the proteins where this domain type was originally identified; PSD-95 (proteins postsynaptic density 95), Dlg (disc large) and ZO-1 (zonula occludens). These ~90 residue long globular domains are composed of six β–strands (βA–βF) forming two opposing antiparallel sheets, and two flanking α–helices (αA and αB), with the N- and C-termini close in space (**Figure 1.6**).

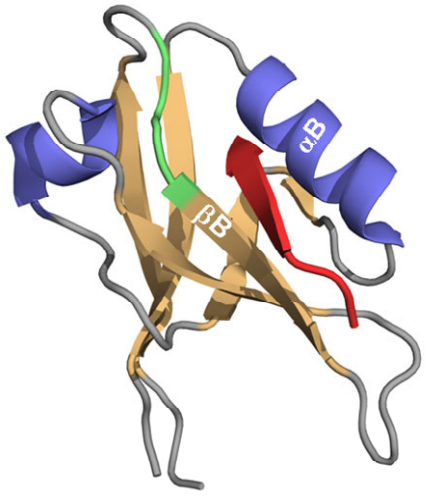


Figure 1.6. Structure of a PDZ domain bound to a peptide. Ribbon representation of the solution structure of AF6 PDZ domain complexed with the LFSTEV peptide (PDB code 2AIN) (Zhou 2005). The peptide (red) forms an additional antiparallel β-strand with the PDZ βB strand (orange) and interacts with the PDZ αB helix (blue) and the carboxylate-binding loop (green).

PDZ domains bind a short C-terminal peptide sequence of their partner proteins, usually transmembrane receptors and channel proteins. Binding of the peptide ligand takes place in an extended groove between the βB-strand and the αB-helix and does not significantly affect the overall PDZ structure. By extensive hydrogen bonding with the βB-strand, the peptide forms an additional β-strand, a mechanism known as β-strand addition. The peptide carboxylate group binds to the carboxylate-binding loop which precedes the βB and contains the conserved (R/K)xxxGψGψ motif (where ψ is a hydrophobic residue). The interactions between the peptide and the βB-strand are sequence-independent, involving only backbone bonds. On the other side of the binding groove, the side chains of the βA-helix residues have specific interaction with the side chains of the peptide residues. Generally, the last four C-terminal peptide residues are necessary and sufficient for the binding. Mutational experiments showed that residues in positions 0 and -2 are crucial for the binding specificity (where 0 is referred to the C-terminal residue, while preceding residues towards the N terminus are numbered as -1, -2 etc., hereafter referred as P-0, P-1, P-2, etc.). Structural studies confirmed the specific binding of these two residues with the PDZ and revealed that also residues at P–1 and P–3 and those at further upstream positions can contribute to the binding [67], [68], [69], [70]. PDZ domains are classified on the basis of their preferred peptide ligand sequence, or more precisely, on the peptide residues at P-0 and P-2. Class-I and class-II PDZ domains recognize S/T-x-Φ-$_{COOH}$ and Φ-x-Φ-$_{COOH}$ motifs, respectively (where ϕ is a hydrophobic amino acid and x is any amino acid). Ligand peptides were classified in the same way as PDZ domains with which they bind. In Class I, the side chain hydroxyl group of S/T at P-2 forms a hydrogen bond with a highly conserved histidine at position 1 of the αB helix of the PDZ (αB1). Class II PDZ domains have a hydrophobic residue (methionine or leucine) at the αB1 position, which interacts with the peptide hydrophobic P-2 residue [71]. In both PDZ classes, the hydrophobic C-terminal residue (usually valine, but also isoleucine or leucine) enters in a hydrophobic pocket which is separated from the pocket hosting residue P-2. Additionally, class III PDZ domains bear a tyrosine residue in the αB1 position, whose hydroxyl group forms an hydrogen bond with the side chain carboxylate of the peptide P-2 residue (aspartic or glutamic acid) [72], [73]. There are also PDZ domains that recognize peptides with a x-x-C-$_{COOH}$ binding motif, which are some time

considered as class III PDZ [74]. However, the above classification for the specificity should be taken with caution, because there is an increasing number of binding promiscuity cases whereby the same PDZ domain can recognize the first as well as the second class of ligands [75], [76], [77], [78], [79], [80].

The class I peptides have a Serine or Threonine residue at P-2, which can be phosphorylated by kinases. It was shown that the interaction between the PDZ domain of PSD-95 and its binding partners, the inward rectifier K^+ channel Kir2.3 and the NMDA receptor, is abolished upon serine phosphorylation of the ligand at P-2 [81], [82]. Phosphorylation can also occur at other positions of the ligand peptide. For example, the serine at P-3 of the AMPA receptor subunit GluR2 is phosphorylated by protein kinase C, disrupting in that way its binding to the PDZ domain protein GRIP [83]. Besides the negative regulation of the PDZ interaction, phosphorylation of the peptide ligand residues can also have the opposite effect. An example is given by serine phosphorylation at P-3 in the MRP2 C-terminus, which increases the binding affinity to its binding partner, the PDZ domain of EBP50 protein [84]. Even more intriguingly, phosphorylation at the same position in a certain ligand does not necessarily have the same effect on all its PDZ binding partners, but can rather serve as a partner switching regulator [85].

The repertoire of the PDZ binding motifs is not limited only to the C-terminal regions, but encompasses also specific internal sequences which can structurally mimic the free C-terminus. This binding mode was observed in the case of the nNOS PDZ which can heterodimerize either with the second PDZ domain of PSD-95 or with the syntrophin PDZ domain. Both PDZ domains can recognize the β-finger structure formed in the C-terminal extension of the nNOS PDZ. The β-finger enters in the peptide binding groove (αB/βB groove) and interacts in a similar way as a C-terminal peptide would do [86]. PDZ domain can dimerize using also other interfaces. GRIP1, a protein with 7 consecutive PDZ domains, can dimerize by its sixth PDZ domain, using the βA-strand and αA-βD loop from each domain. In this way, the peptide binding grooves are oriented in the opposite direction, which gives the possibility of simultaneous interactions with several ligand [87]. PDZ domains can also interact with lipids, in particular with phosphatidylinositol. It was shown that the PDZ domains of syntenin and CASK can bind to phosphatidylinositol phosphate (PIP) with a binding constant (10-50 μM) comparable to that of other known lipid-binding

domains (PH and C2 domains), suggesting a possible biological significance of this interaction [88]. Structural analysis showed that the PDZ-PIP interaction is mainly hydrophobic, but involves also some positively charged PDZ residues that are in contact with the negative PIP phosphate group. The PDZ residues involved in this interaction were found in the βB strand, αB helix and the βD-βE loop, which indicates a partial overlap between peptide and lipid binding regions [89].

PDZ-containing proteins can be divided into three groups. The first one comprises proteins containing only PDZ domains, from a single PDZ in PICK1 up to 13 PDZ domains in MUPP1. The second group includes MAGUK (membrane-associated guanylate kinase) proteins which contain between one and three PDZ domains, one SH3 domain and a guanylate kinase homology (GuK) domain [90]. There are 24 MAGUK proteins in the human genome and several of them works in the organization of signaling complexes (PSD-95, Dlg, ZO-1 and others). The third group is formed by proteins (AF6, Scribble and others) having single/multiple PDZ domain combined with other protein-interacting domains. Proteins containing PDZ domains are generally considered scaffold proteins; proteins that bind with multiple members of a signaling pathway, tethering them into complexes. PDZ scaffold proteins also play an important role in the establishment of cell polarity. One of the best studied examples of PDZ scaffolds is the MAGUK protein PSD-95, which organizes glutamate receptors and their associated signaling proteins to determine the size and strength of synapses [91].

1.3.2 Afadin/AF6

Afadin/AF6 is a multidomain protein that plays a scaffolding role in the organization of cell-cell junctions [61], [92]. In epithelial cells there are three types of cell-cell junctions: tight junctions, adherens junctions and desmosomes (**Figure 1.7**). In these junctions, neighbouring cells are linked by transmembrane adhesion molecules which are further connected to the actin cytoskeleton by cytoplasmic scaffolding proteins. Adherens junctions (AJ) regulate tissue formation during development and maintain the solid tissue in the adult organism [93]. There are two major adhesive protein complexes at adherent junctions, cadherin-catenin and nectin-afadin. Cadherins are transmembrane glycoproteins that mediate Ca^{2+}-dependent intercellular adhesion through homophilic interactions. Cadherins are linked to the cytoskeleton through catenins, cytoplasmic proteins associated with their cytoplasmic

tail and required for the adhesive function. Nectins are adhesion proteins that can form homo- and heterodimers in a Ca^{2+}-independent manner. There are four members in the nectin family, most of which are transmembrane Ig-like proteins and have a C-terminal PDZ binding motif at their cytoplasmic tail. Nectins couple adherent junctions to the cytoskeleton through the interaction with the PDZ domain of afadin, a protein that binds actin filaments.

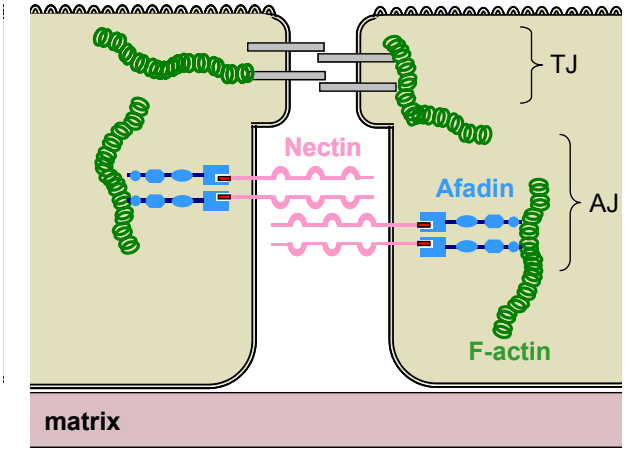


Figure 1.7. The Afadin-Nectin interaction in adherent cell junctions. In adherent cell junctions (AJ), the PDZ binding motif of Nectin (red) interacts with the PDZ domain of afadin (blue), coupling Nectin to F-actin (green).

There are two main splicing variants of this protein, a longer one, l-afadin, and a shorter one, s-afadin (**Figure 1.8**). L-afadin, simply called afadin, consists of two N-terminal Ras-association domains, one forkhead association domain (FHA), one dilute domain (DIL), one PDZ domain, three proline rich segments and one F-actin-binding domain. The smaller variant, which lacks the third proline-rich region and the F-actin–binding domain, is named AF6, like the afadin's gene AF-6 (**A**LL-1 **f**usion partner on chromosome **6**), which was found to be fused to the *ALL-1* gene in acute myeloid leukemias caused by chromosomal translocation (t 6;11) [94]. Afadin is ubiquitously expressed in epithelial and endothelial cells, whereas AF6 is mainly expressed in neural tissue [61], [95]. Afadin has also other binding partners involved in cell-cell junctions. For example, it was shown that afadin can interact directly and indirectly with catenins, clustering cadherin and nectin complexes and thus making the adhesion in adherent junctions stronger [96]. AF6, although unable to bind the

actin cytoskeleton, is however involved in cell-cell junctions. AF6, through its Ras-binding domain, can bind Z0-1, a scaffold protein implicated in clustering of several adhesion proteins in tight junctions. Using the same binding domain, AF6 interacts with several members of the Ras family, small GTPases involved in the regulation of cell proliferation and differentiation. Activated Ras inhibits the interaction between AF6 and Z0-1, which further results in the perturbation of cell–cell contacts. As a multidomain scaffold protein, AF6/afadin recruits different protein complexes near the plasma membrane cell-cell junctions. Through its single PDZ domain, AF6/afadin was shown to bind, beside nectins, some other transmembrane proteins like Bcr kinase (negative regulator of Ras signaling pathway [97], Neurexin (neuronal cell surface protein, [56]), JAM (junctional adhesion molecule in TJs, [98]), SPA-1 (GTPase-activating protein, [99]), Eph-related receptor tyrosine kinase ([56], [100]) and Jagged-1([56]).

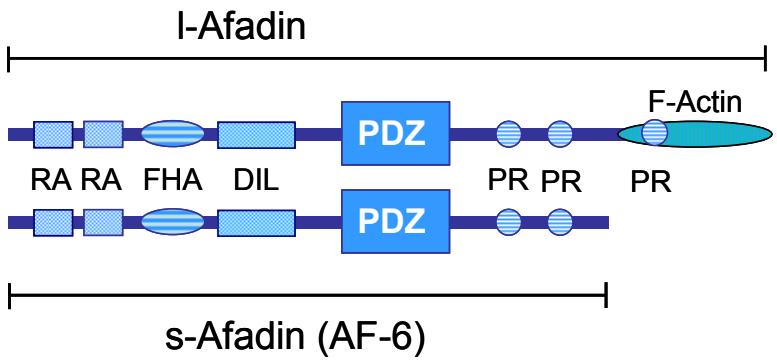


Figure 1.8. Domain architecture and splicing variants of afadin. The long (l-Afadin) and short (s-Afadin) variants are shown. RA, Ras-association domain; FHA, Forkhead associated domain; DIL, Dilute domain; PDZ, domain present in PSD-95, Dlg, and ZO-1/2; PR, proline-rich region; F-actin, F-actin binding domain

1.3.3 AF6_PDZ domain

The NMR 3D structure of AF6_PDZ is similar to X-ray structures of other PDZ domains [75]. Its structure is very close to the third PDZ domain of PSD-95 protein (RMSD 2.1 Å; sequence identity = 30.4%; sequence similarity = 50%). Even if sorted in the class-II by oriented peptide library technique, AF6-PDZ can also interact with class-I ligands. The amino acid sequence shows that the crucial residue for the peptide binding specificity, the first residue of the αB helix (αB1), is

glutamine. Generally, the αB1 position is occupied by a hydrophobic amino acid or by histidine, in Class II and Class I PDZ domains, respectively [68], [101]. Glutamine in this position is unique and is not found even in Class III PDZ domains, where αB1 residue is tyrosine. The side chain of the glutamine (αB1) points towards the peptide binding groove in a similar way the histidine (αB1) in the PSD-95 PDZ3 does, which is one typical Class I PDZ domain. Even if AF6-PDZ and PSD-95 PDZ3 share similar βB/αB grooves, there are length differences in the βB-strand flanking loops, βA/βB and βB/βC. Another feature that distinguishes the AF6 PDZ domain is its carboxyl-binding loop with a GMGL sequence, which is slightly different from the standard GLGF motif. However, the structure analysis shows that the hydrophobic residues, methione and leucine, have their side chains pointed toward the interior of the protein making a hydrophobic pocket for the ligand carboxyl group [75].

NMR chemical shift mapping of the AF6_PDZ titrated with C-terminal peptides from Bcr and Neurexin, ligands of Class I and II, respectively, showed in both cases binding via the βB/αB groove, with slightly differences in the binding mode. Regardless of its structural characteristics and binding promiscuity, AF6 PDZ has a greater tendency to bind Class II ligands [68]. Additionally, mutational analysis by two-hybrid assays showed that a minimum of six sequence specific C-terminal residues from the EphB2 and EphB3 ligands are required for binding to AF6-PDZ [56]. All together, AF6_PDZ is a very flexible interaction domain, which makes interesting to study how binding to the ligands is regulated.

1.4 Notch structural biology

Current structural analyses in Notch signaling are mainly focused on the receptor-ligand recognition, the activation of the receptor proteolysis and the nuclear complex assembly. Analysis of the receptor-ligand binding interface, carried out in *Drosophila* by cell-aggregation experiments, determined the pair of EGF-like repeats 11-12 in Notch and the DSL domain in Delta as a minimal region for the their association [5].

The first structural insights were the NMR and X-ray structures of the human Notch-1 tandem EGF-like repeats 11^{th} -13^{th} ($N1\_EGF_{11-13}$) and the X-ray structure of the human Jagged-1 region corresponding to the DSL domain and the first three EGF-like repeats ($J1\_DSL\_EGF_{1-3}$) [102], [103] (**Figure 1.9 A, B**). It was shown that each

EGF-like repeat has its characteristic fold consisting of a two-stranded antiparallel β-sheet and three disulfide bonds. The X-ray structure of N1_EGF_{11-13} showed that a coordinated Ca^{2+} ion fixes the orientation between adjacent repeats, creating an extended rod-like structure. In the NMR analyses of N1_EGF_{11-13}, the dynamic behavior of the EGF_{13} is distinctly different from that of the other two repeats, suggesting that the position of the EGF_{13} is not rigidly fixed with respect to the EGF_{11-12} repeats.

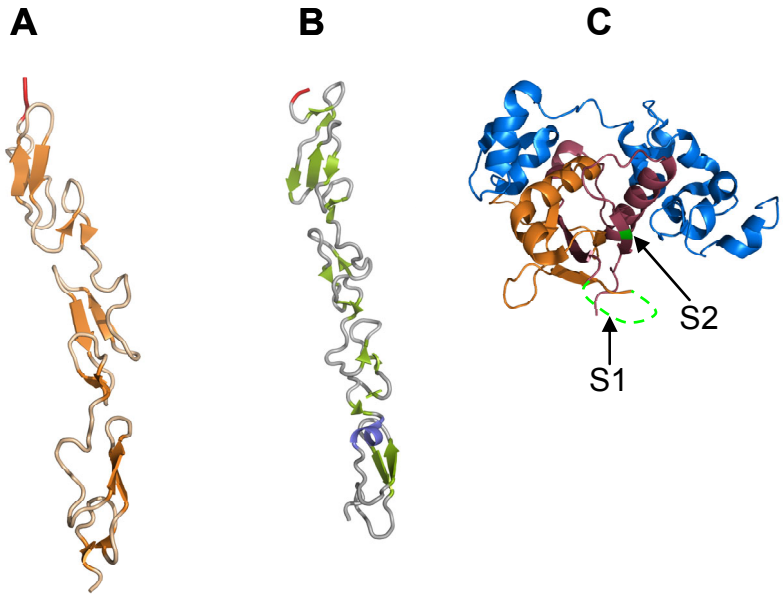


Figure 1.9. Structures of functionally relevant regions in Notch receptors and ligands. Cartoon representation of (**A**) the tandem EGF-like repeats 11th -13th of human Notch-1 (PDB code 2VJ3), (**B**) the region comprising the DSL domain and the first three EGF-like repeats from human Jagged-1 (PDB code 2VJ2), and (**C**) the negative regulatory region (NRR) of human Notch-2 (PDB code 3I08). The NRR region consists of three Lin/Notch repeats (LNR, blue) and of a heterodimerization domain (HD, N-terminal region in orange, C-terminal region in raspberry red) that is cleaved at S1. Cleavage sites S1 and S2 are in light green.

From the structure of these three tandem repeats, two distinct models for the structure of the Notch-1 extracellular region comprising all 36 EGF-like repeats were proposed. In the first model, the region has an extended conformation, mainly rigid along its length, but with flexible parts between some EGF-like repeats that do not bind Ca^{2+} ions. In the second model, the extracellular region folds onto itself to form a triple-stranded structure upon ligand binding. Folding is allowed by points of

flexibility in the linkers that connect non-Ca^{2+}- binding EGF-like repeats. Using the second model, we can speculate how the extracellular region of Notch receptors, having a different number of EGF-like repeats, can adopt different structures and thus can have different affinity for different Notch ligands. The X-ray structure of the Jagged-1 polypeptide corresponding to the DSL domain and the first three EGF-like repeats (J1_DSL_EGF_{1-3}) revealed an extended conformation that resembles the structure of the N1_EGF_{11-13} [103]. The DSL domain shows similarities with an EGF module, but has a recognizably different disulfide bond pattern. EGF_1 and EGF_2 exhibit a truncated version of the EGF-fold, with no canonical secondary structure and a more distant structural homology to other EGF-like domains while EGF_3 has a classical EGF-like fold. Recently, it was shown that exons 5 and 6 of the human *JAG1* gene, which together encode EGF_1 and EGF_2 repeats, are out of phase with respect to the domain boundaries [104]. Exon 5 encodes a truncated EGF with only four half-cysteines and exon 6 encodes the C-terminal half of EGF_1 and the entire EGF_2. It has been proposed that the DSL domain may have evolved from the truncation of tandemly connected, short EGF repeats [103]. Interestingly, the region encoded by exon 6, J1ex6, can be viewed in itself as two truncated tandem EGF repeats and has an amino acid sequence and a disulfide pattern reminiscent to those of a DSL domain.

Two models for the Notch-1/Jagged-1 binding interface were proposed on the ground of in-silico docking of these two structures, confirming the DSL domain as the minimal and indispensable unit for the binding. Moreover, the first two EGF-like repeats of Jagged-1 also participate in the formation of the interface. This is in agreement with findings obtained from deletion mutants, where it was demonstrated that the EGF-like repeats of mouse Jagged-1 can modulate the affinity of the binding to mNotch-2 [105].

Besides the pair of EGF-like repeats 11-12, the most important part of the Notch extracellular region is the so-called negative regulatory region (NRR). It is composed of three LIN-12 repeats and the heterodimerization domain (HD), which holds together the extracellular and transmembrane subunits of the receptor (**Figure 1.9 C**). The HD domain contains the S2 cleavage site which becomes exposed to the ADAM-metalloprotease upon ligand-receptor binding, but the mechanism is still not clearly described. The crystal structure of the NRR showed for the first time the autoinhibited conformation of the human Notch-2 [106]. Extensive interdomain interactions between the LIN-12 and HD domains bury the S2 site, suggesting that a

drastic conformational change in this region is required to expose this site upon activation triggered by the ligand. It has been proposed that ligand binding relieves this autoinhibition either by inducing an allosteric conformational change in the receptor or by the pulling force caused by the ligand endocytosis, but further experiments are necessary for the exact description of this activation switch.

The crystal structure of the Notch transcriptional complex highlighted how the Notch intracellular region (N_ic) binds CSL and then both recruit Mastermind to provide stability for the complex [107], [108] (**Figure 1.10**). It was shown that the Ankyrin repeats (ANK) of N_ic are indispensable for the interaction with Mastermind [109] while for the binding with CSL N_ic mainly uses its RAM region. ANK domains of Notch consists of seven ANK repeats, each of which is about 33 residues long and typically folds into a pair of antiparallel helices followed by a β-hairpin that connects to the next repeat [110]. RAM domains of Notch are approximately 100 residues long and defined as a region that starts from the γ–secretase cleavage site and ends at the first ankyrin repeat. RAM has very little or no secondary structure but becomes ordered upon complex formation [109], [111]. CSL consists of an N-terminal Rel-homology region (RHR-N), a central β-trefoil domain (BTD) and a C-terminal Rel-homology region (RHR-C). The DNA is bound by the BTD and the RHR-N and makes no contacts with the distal RHR-C domain. Mastermind proteins are ~1000 residues long proteins predicted to be of low structural complexity, suggesting that they function as scaffolds for further recruitment of additional co-activators and/or the transcription machinery [109]. Only a short N-terminal fragment (55 residues) of Mastermind is required for the interaction with CSL and N_ic [112]. The C-terminal portion of Mastermind is important for the interaction with CBP/p300 and to activate transcription [113]. Mastermind adopts a bent helical structure when complexed with CSL and N_ic. The N-terminal helical region interacts with RHR-C of CSL and ANK of N_ic, while its C-terminal helix binds RHR-N of CSL [107], [108]. In the CSL-DNA complex, when not bound to N_ic and Mastermind, the β-hairpin loop of RHR-N is closed. In order to accommodate binding of the C-terminal Mastermind helix, the β-hairpin loop needs to adopt an open conformation. It was demonstrated [114] that RAM binding to the BTD of CSL produces an allosteric change in the RHR-N, opening the β-hairpin loop and in that way creating one-half of the docking site required for Mastermind to bind to the complex. The CSL-ANK

interface, created upon the interaction between RAM to BTD, creates the complete docking site for Mastermind. Despite the structural and biochemical information about the RAM and Ankyrin repeats, less is known about the overall structure of the entire N_ic, especially the less conserved C-terminal region of N_ic. The future challenge is obtaining structural insights into the cooperativity between the Notch transcription complex and other transcription factors.

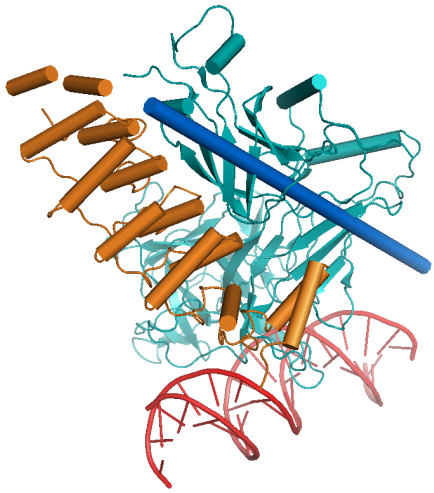


Figure 1.10. Crystal structure of the Notch transcriptional complex. Ankyrin repeats of N_ic in orange, the N-terminal region of Mastermind is in blue, CSL is in cyan and dsDNA is in red.

The cytoplasmic tails of the Notch ligands still need to be characterized biophysically. The cytoplasmic tail of Jagged-1 is 125 residues long region and, based on its amino acid composition, is predicted to have a high degree of intrinsic disorder. J1_ic has a positive mean net charge and is rich of lysines, arginines, glutamines and prolines, residues that predominates in disordered proteins. Moreover, J1_ic is significantly depleted of residues that usually form the hydrophobic core of globular proteins (Ile, Leu, Val, Trp, Tyr and Phe) [115]. Many intrinsically disordered proteins and protein domains have been identified, and regions predicted to be intrinsically disordered are quite common, especially in eukaryota [116]. Although these regions were considered in the past of little importance by structural biologists, they often play an important role in protein-protein interactions, especially in signaling networks [117]. The role of intrinsic disorder in these contexts is well

explained in terms of the plasticity that disorder confers to proteins, which allows proteins to bind different partners through disorder to order transitions [118], [119]. Intrinsically disordered regions can work as docking stations hosting several different protein interactions motifs on the same polypeptide chain. In protein/protein interactions, they potentially offer a very large interaction surface/volume ratio compared to globular proteins, which makes them capable of binding their target molecules with high specificity but low affinity. Detailed biophysical studies are crucial to clarify the relationship between biological function and structural characteristics of intrinsically disordered proteins. The application of increasingly sophisticated methods has revealed that disordered proteins are far from being homogeneous statistical random coil polymers. Intrinsically disordered proteins exhibit a rich diversity of local and even long-range structural preferences as well as dynamics, that are likely to be of functional significance.

1.5 Aim of the thesis

The signal transduction cascade induced in the receptor bearing cell by the interaction between Notch receptors and their ligands has been studied extensively and, although the structural description of the process is still patchy, the network of molecular interactions has been delineated. On the contrary, very little is known on the molecular processes occurring in the ligand bearing cell. Recent work has reshaped our view of Notch signaling, drawing it as a bi-directional process, and raised many issues on the role of Notch ligands cytoplasmic tails in ubiquitination-mediated endocytosis, interaction with PDZ containing proteins involved in the organization of cell-cell junctions, and in the release of signaling fragments upon proteolytic cleavage of the ligand. In fact, the intracellular region of membrane spanning receptors (and ligands), although generally predicted to be disordered, is in most cases the actual effector of signaling within the cell. Intrinsically disordered regions are indeed very common, evolutionary persistent, and functionally important, as their high structural plasticity make them very adaptable binding partners. The cytoplasmic tail of Jagged-1 was shown to interact with the PDZ domain of afadin, a protein located at cell-cell adherens junctions, and to undergo regulated intramembrane proteolysis with the release of signaling fragments. Jagged-1 cytoplasmic tail may then exist in at least two forms: as a membrane-tethered protein located at the membrane/cytosol interface and as a soluble nucleocytoplasmic protein. Understanding the role of Jagged-1 intracellular region in Notch signaling requires a thorough understanding, at the structural level, of its conformational properties and of its interaction with the inner side of the plasma membrane on one side, with the binding partners on the other. This work presents the first biochemical and structural characterization of Jagged-1 cytoplasmic tail in solution and in environments that mimic the membrane/cytoplasm interface, and the first structural investigation on its interaction with the afadin PDZ domain.

2. RESULTS

2.1 Structural characterization of J1_tmic

2.1.1 Expression and purification of the recombinant Jagged-1 intracellular region

Gene synthesis. The recombinant protein corresponding to Jagged-1 intracellular region (J1_ic, residues 1094-1218 of JAG1_human, 125 amino acids) was expressed in *E. coli* from a synthetic gene designed to optimize codon usage for heterologous expression. Codon optimization tailored on the expression host has proved to be an effective way to increase heterologous protein expression, especially in *E. coli* [120]. In the past decades, total gene synthesis was not considered of general use because of the cost of the oligonucleotides and the errors introduced both in the oligonucleotide synthesis and in the gene assembly. Today, synthetic genes can be easily purchased at competitive costs. However, gene synthesis is most attractive when the target protein or protein domain is relatively small and high expression levels are needed, as is the case of J1_ic. The synthetic gene encoding J1_ic and optimized for *E. coli* codon usage [121], [122] was assembled from 16 oligonucleotides designed by DNAWorks [123]. As a result of codon optimization, no codon with an abundance <20% was used, and ~80% of the protein residues were encoded by >50% abundant codons, according to Class II genes codon usage [121] (**Figure 2.1**).

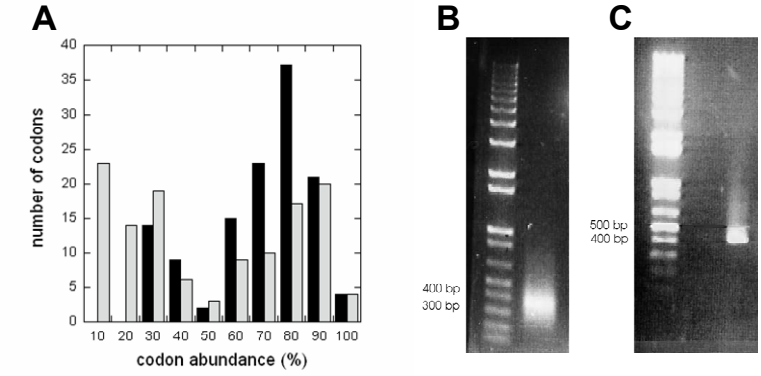


Figure 2.1. Gene synthesis of J1_ic. (A): number of codons used in the human cDNA (grey) and in the synthetic gene (black) of J1_ic partitioned according to their relative abundance in *E. coli* Class II genes. Agarose gel (1%) of the oligonucleotide assembly (**B**) and of the gene amplification (**C**) PCR mixtures.

Protein expression and purification. High expression levels were obtained, as measured by SDS-PAGE of total bacterial cell lysates (**Figure 2.2**). The protein was recovered from the soluble fraction, and the use of detergents increased the recovery yield, suggesting that the target protein can interact with other cellular components.

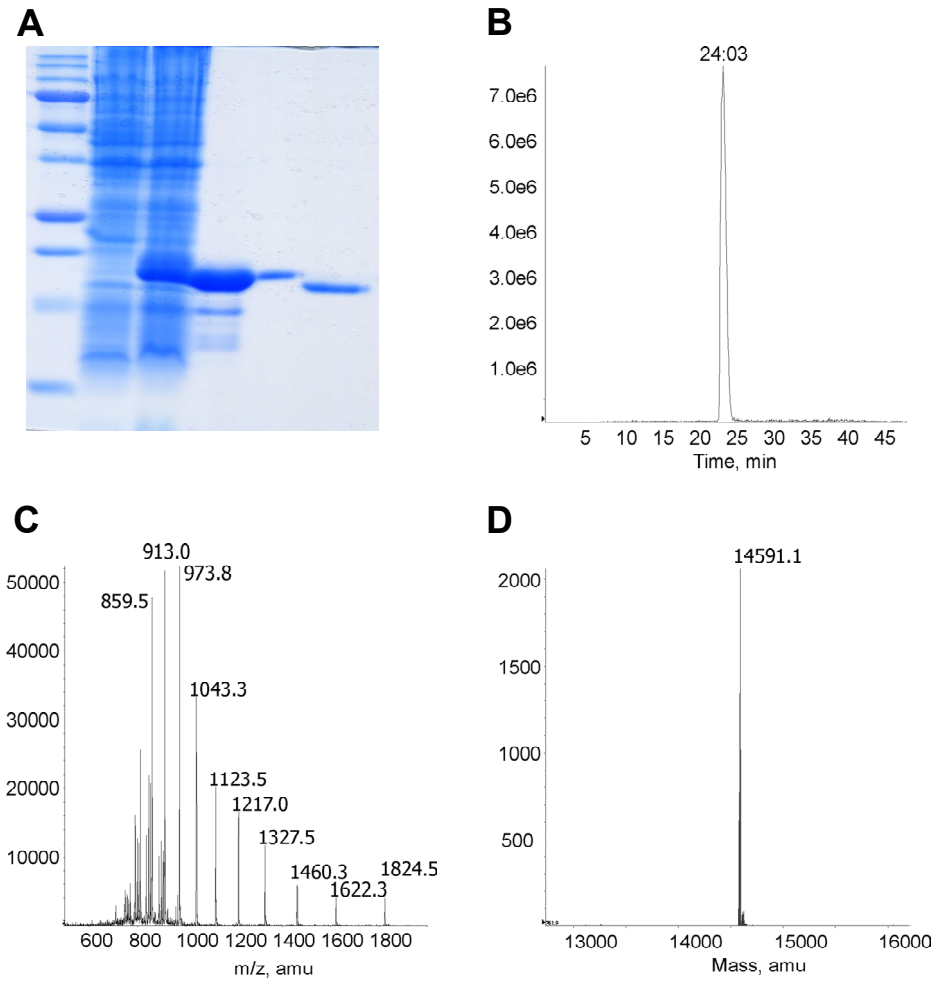


Figure 2.2. Protein expression and purification of J1_ic. (**A**): SDS-PAGE at different purification steps: lane 1, LMW markers; lane 2, total cell lysate from not-induced culture; lane 3, total cell lysate after 3 h induction; lane 4, J1_ic_his$_6$ after IMAC purification; lane 5: J1_ic_his$_6$ after IMAC and RP-HPLC purification; lane 6: J1_ic after his-tag cleavage and purification by IMAC and RP-HPLC. (**B**): LC chromatogram and (**D**): mass spectrum of J1_ic.

The protein purification strategy for the recombinant J1_ic, expressed with a cleavable N-terminal His-tag, was based on immobilized metal ion affinity chromatography (IMAC) followed by removal of the His$_6$-tag. The recombinant

dipeptidyl aminopeptidase I (DAPase) was used for the tag removal. DAPase cleaves two amino acid residues at a time, starting from the N-terminus of the polypeptide chain, and stops when it finds a KX, RX, XP, or XXP sequence (X=any residue). The cytoplasmic tail of Jagged-1 starts with the RK dipeptide, which is one of the stop sequences for the above enzyme. Thus, after the tag-cleavage, the native sequence is obtained. An additional purification step was necessary between the IMAC and the His_6-tag removal because of partial proteolytic degradation of the target protein (**Figure 2.2**), despite the use of protease inhibitors.

Removal of partially digested products, cleaved His_2 dipeptides and of the DAPase enzyme, which contains a C-terminal His-tag by itself, was achieved through an additional IMAC step. After a final RP-HPLC purification step, J1_ic was analyzed and identified by LC–MS (*M*r calculated: 14588; found: 14591). The final yield was ~8 mg of purified material from 1L culture.

An alternative purification protocol, based on IMAC, desalting by gel filtration, and ion-exchange chromatography on a strong cation exchanger resin was also developed in order to exclude the possibility that the protein biophysical properties may be affected by the denaturing conditions used in the standard purification protocol. The purity of the His_6-tagged protein recovered using this purification scheme was ~85%.

2.1.2 J1_ic is mainly disordered in aqueous solution

The first method applied for the J1_ic structural characterization was far-UV circular dichroism (190-250 nm), a spectroscopical method used for the determination of protein secondary structure content. Far-UV CD spectra of J1_ic in acqueous buffer show a strong minimum at ~198 nm, which is typical of random coil polypeptides or unfolded proteins (**Figure 2.3**). A quantitative estimation of the secondary structure content was achieved through deconvolution of the CD spectrum in the range 190–240 nm using several methods [124]. The best fit between the experimental and calculated spectrum was obtained from CDSSTR using a set of reference spectra that include unfolded proteins [125], [126] (normalized RMSD = 0.014). Results suggest a very low content of secondary structure: 5% helix, 16% strand, 10% turn and 67% coil and similar values were found also by other methods [126] (SELCON3, CONTINLL) (**Table 1**). J1_ic purified in native conditions (alternative purification

protocol) showed very similar far-UV CD spectra (**Figure 2.3**), thus confirming that the purification process did not significantly affect its intrinsic conformation.

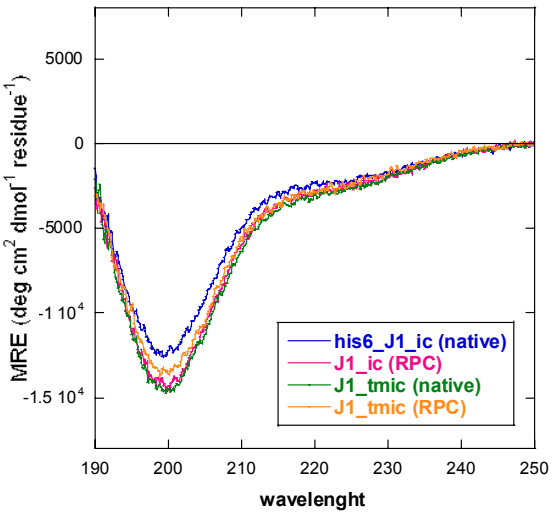


Figure 2.3. Circular dichroism. Far-UV CD spectra of his6_J1_ic (7 μM), J1_ic (7.5 μM) and J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer, pH 7.4. The purification conditions (RPC, reverse-phase chromatography, or native) are also specified.

The mainly disordered nature of J1_ic in solution was confirmed by the near-UV CD measurements (250-350 nm). Spectra showed no detectable optical activity of the aromatic residues indicating the lack of any tertiary structure in J1_ic.

The single Trp residue in the sequence of J1_ic (W1196) was used as a probe in fluorescence spectroscopy studies (**Figure 2.4**). The position of the maximum in the emission spectra (λ_{max} = 356 nm) (**Figure 2.5**) and the very low anisotropy value (R = 0.06) are consistent with a solvent-exposed displaying a high segmental mobility of W1196. These results support the structural disorder of J1_ic in solution and agree with those obtained by CD.

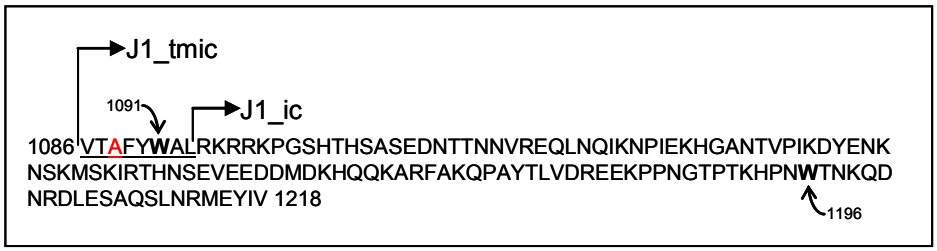


Figure 2.4. Amino acid sequences of J1_ic and J1_tmic. Residues of the hydrophobic intramembrane fragment of J1_tmic are underlined, tryptophan residues are in bold; C1092 was mutated to Alanine (in red).

J1_tmic

	H_{222} (%)	H (%)	E (%)	T (%)	C (%)	RMSD
Buffer, pH 7.4	7	4	19	12	65	0.013
Buffer, pH 6	7	3	14	9	73	0.011
Buffer, pH 7.4 (return)	7	4	22	14	60	0.017
5 % TFE	12	7	11	10	72	0.014
10 % TFE	14	8	14	11	66	0.016
25 % TFE	26	38	11	16	35	0.014
50 % TFE	35	50	12	27	27	0.012
SDS (3 mM), pH 7.4	14	17	18	14	50	0.013
SDS (3 mM), pH 6	21	33	15	19	33	0.010
SDS (3 mM), pH 7.4 (return)	14	17	18	15	48	0.012
DMPG (1mM), pH 7.4	16	19	12	12	57	0.016
DMPG (1mM), pH 6	23	36	11	16	36	0.013
DMPG (1mM), pH 7.4 (return)	17	21	15	14	51	0.017
DMPS (1mM), pH 7.4	15	17	14	13	55	0.016
DMPC (1mM), pH 7.4	9	5	15	11	68	0.017
LMPG (1mM), pH 7.4	13	16	13	12	58	0.020
LMPG (1mM), pH 6	22	34	13	18	35	0.023
LMPC (1mM), pH 7.4	8	7	9	9	75	0.014
LMPC (1mM), pH 6	8	6	10	8	75	0.027

J1_ic

	H_{222} (%)	H (%)	E (%)	T (%)	C (%)	RMSD
Buffer, pH 7.4	9	5	16	10	67	0.016
Buffer, pH 6	8	4	12	9	74	0.010
Buffer, pH 7.4 (return)	9	5	15	9	70	0.014
50 % TFE	33	42	14	15	29	0.013
SDS (3 mM), pH 7.4	15	17	10	11	60	0.018
SDS (3 mM), pH 6	19	23	12	14	51	0.020
SDS (3 mM), pH 7.4 (return)	15	17	11	11	61	0.017
DMPG (1mM), pH 7.4	14	17	12	12	59	0.013
DMPG (1mM), pH 6	20	25	10	13	54	0.020
DMPG (1mM), pH 7.4 (return)	15	19	11	14	57	0.020

Table 1. Circular dichroism – the secondary structure content. Helical content calculated from the ellipticity at 222 nm (H_{222}), secondary structure content (H, helix, E, strand, C, disordered) calculated by CDSSTR through deconvolution of CD spectra, and normalized root mean squared deviation of the fit (NRMSD). CDSSTR was run through the DichroWeb interface using a reference dataset containing also unfolded proteins. Results are reported both for J1_tmic (top) and for J1_ic (bottom), and shaded in different colors according to the conditions used (grey TFE, orange SDS micelles, yellow negatively charged phospholipid vesicles and micelles, and blue zwitterionic phospholipid vesicles and micelles).

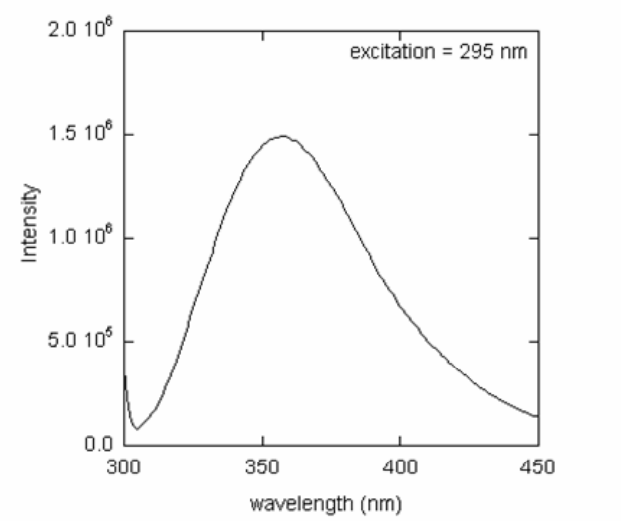


Figure 2.5. Fluorescence spectroscopy. Fluorescence emission spectrum of J1_ic (8.2 μM) in 5 mM MOPS. Excitation wavelength was set at 295 nm and the raw spectrum was corrected for the contribution of the buffer alone.

The 1D ^{1}H-NMR spectrum of J1_ic in H_2O (**Figure 2.6 A**) is characterized by rather sharp resonances, but low chemical shift dispersion. Very little dispersion in the chemical shift of both the backbone NH and the Hα protons, most of them being clustered in the 8.0-8.5 and 3.9-4.9 ppm range, respectively, could be observed in the TOCSY spectrum (**Figure 2.6 B**). The only exception is a spin system that was tentatively assigned to a valine, in which the NH is significantly shifted upfield (δ = 7.5 ppm). The corresponding Hα resonance is only slightly shifted (4.0 ppm) and the Hβ and Hγ resonances are within the random coil values. It is possible that this is the C-terminal valine, and that the shift in the NH chemical shift is due to charge effects. Methyl resonances from valine, leucine and isoleucine residues are all clustered in a small range (0.8-1.0 ppm) with no methyl resonance below 0.8 ppm. Aromatic signals were tentatively assigned and are all close to the random coil values, histidines: 7.9-8.0 ($H^{\varepsilon 1}$)/7.0-7.1 ($H^{\delta 2}$) ppm, phenylalanine: 7.3-7.4 ppm, tyrosines: 6.8 (H^{ε})/7.1 (H^{δ}) ppm, tryptophan: 10.2 ($H^{\varepsilon 1}$)/7.2 ($H^{\delta 1}$)/7.6 ($H^{\varepsilon 3}$)/7.4 ($H^{\zeta 2}$)/7.2 ($H^{\eta 2}$)/7.1 ($H^{\zeta 3}$) ppm (**Figure 2.6 C**). Overall, the lack of chemical shift dispersion in the NH, Hα, methyl groups and aromatics suggest that J1_ic is mainly disordered in solution, with little or no secondary and tertiary structure present.

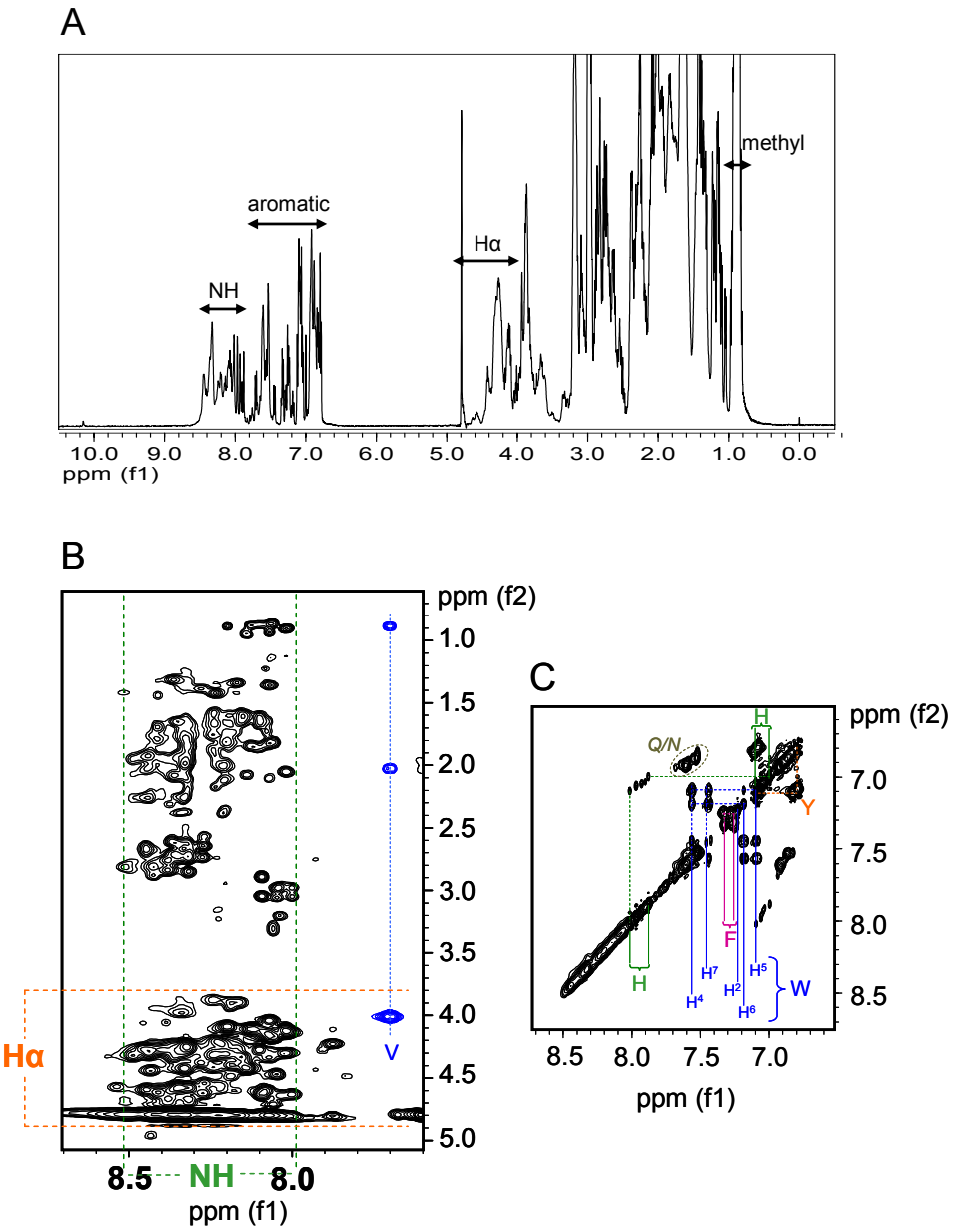


Figure 2.6. ^{1}H NMR spectroscopy. (**A**): 1D spectrum and (**B** and **C**): 2D TOCSY spectrum of J1_ic recorded in H_2O/D_2O (90/10 v/v), pH 6.6. In the NH/aliphatic region of the 2D TOCSY spectrum (**B**), the Hα and NH regions are labeled; cross peaks from the putative C-terminal Valine are in blue. In the aromatic region of the 2D TOCSY spectrum (**C**), cross peaks of the six histidines, one phenylalanine, two tyrosines and one tryptophan are marked.

To confirm the results obtained from these spectroscopic techniques, J1_ic was subjected to limited proteolysis using chymotrypsin. Chymotrypsin was selected because it has a higher specificity towards aromatic and bulky hydrophobic residues (F, Y, W, L, and M). These amino acid types are expected to be buried in the

hydrophobic core of the protein, hence less accessible to protease cleavage, when a globular structure is present. J1_ic, however, was very rapidly degraded by chymotrypsin. After 15 min, no native protein could be detected, and more than 15 different fragments were separated. Under the same conditions, small globular proteins such as myoglobin and ribonuclease A are not digested to an appreciable extent [127], [128]. The time course of the reaction and the identification of many of the fragments by LC–MS also show that no particular region in J1_ic is protected against proteolytic cleavage (**Figure 2.7**). These results thus confirm that J1_ic is essentially disordered in solution.

RKRRKPGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVPIKDYENKNSKMSKIRTHNSEVEEDDMD
KHQQKARFAKQPAYTLVDREEKPPNGTPTKHPNWTNKQDNRDLESAQSLNRMEYIV

pep.	res.	Mass,calc. (Da)	Mass,found (Da)	t(min)	sequence
1	1-26	3019.25	3019.9	17.8	<RKRRKPGSHTHSASEDNTTNNVREQL>N
2	1-46	5280.81	5281.6	23.3	<RKRRKPGSHTHSASEDNTTNNVREQLN…DY>E
3	1-53	6112.76	6114.4	22.6	<RKRRKPGSHTHSASEDNTTNNVREQLN…KM>S
4	27-53	3111.53	3112.2	20.9	L<NQIKNPIEKHGANTVPIKDYENKNSKM>S
5	47-71	3002.25	3001.9	17.8	Y<ENKNSKMSKIRTHNSEVEEDDMDKH>Q
6	47-83	4419.88	4418.9	22.1	Y<ENKNSKMSKIRTHNS…HQQKARFAKQPAY>T
7	54-71	2170.30	2170.2	17.3	M<SKIRTHNSEVEEDDMDKH>Q
8	54-77	2929.18	2930.0	18.3	M<SKIRTHNSEVEEDDMDKHQQKARF>A
9	54-83	3587.94	3587.4	19.5	M<SKIRTHNSEVEEDDMDKHQQKARFAKQPAY>T
10	60-77	2206.34	2207.1	19.4	H<NSEVEEDDMDKHQQKARF>A
11	69-103	4114.60	4114.3	24.5	M<DKHQQKARFAKQPAYTL…PPNGTPTKHPNW>T
12	78-83	676.77	676.5	17.0	F<AKQPAY>T
13	78-118	4676.10	4676.2	22.6	F<AKQPAYTLVDREE…WTNKQDNRDLESAQSL>N
14	84-103	2316.56	2316.8	21.1	Y<TLVDREEKPPNGTPTKHPNW>T
15	84-118	4017.34	4018.3	21.9	Y<TLVDREEKPPNGTP…TNKQDNRDLESAQSL>N
16	84-123	4711.12	4712.5	22.8	Y<TLVDREEKPPNGTP…NRDLESAQSLNRMEY>I
17	122-125	522.60	521.5	14.1	M<EYIV>

Figure 2.7. Limited proteolysis. Peptide fragment identification (fragment identification number, residue number, expected and found masses, RP-HPLC retention time, and sequence) in the chymotrypsin cleavage mixture after 30 min reaction. The amino acid sequence of J1_ic and the cleavage sites are also shown on top of the table.

2.1.3 The additional intramembrane residues do not promote folding of J1_ic

The intramembrane cleavage site in the Jagged-1 was proposed to be placed at the first valine residue close to the cytoplasmic region, between the L1085 and V1086 [51]. The part of Jagged-1 released from the membrane (J1_tmic), comprises an 8 residue long part of the transmembrane segment and the intracellular region (J1_ic) (**Figure 2.4**). To verify if the addition of that short hydrophobic fragment may change the structural properties of Jagged-1 cytoplasmic tail, J1_tmic was expressed using the

same construct of the J1_ic synthetic gene and applying similar expression and purification protocols used for J1_ic. J1_tmic was expressed without His-tag because it was noticed that J1_ic binds well to the IMAC columns also after the His-tag cleavage, possibly due to its 6 endogenous histidines (**Figure 2.10**). The yield of the purified J1_tmic was ~10 mg from 1L culture, which is slightly larger than the J1_ic yield. The difference in the yield is probably due to the reduced number of purification steps.

The far-UV CD spectrum of J1_tmic (**Figure 2.3**) in Tris buffer is very similar to that of J1_ic, showing a strong minimum at 198 nm, typical of disordered proteins. The estimated content of secondary structure obtained with CDSSTR (normalized RMSD=0.013) show a high content of unordered structure (65%) and a low residual presence of secondary structure (4% helix, 19% strand and 12% turns) (**Table 1**). Very similar results were obtained from the CD spectrum of J1_tmic purified in native conditions, confirming that the purification process did not affect the intrinsic conformation of J1_tmic. These results suggest that the strongly hydrophobic segment (VTAFYWAL), which is expected to be embedded in the membrane and to become exposed to the solvent upon cleavage of Jagged-1, is not sufficient to promote folding of J1_tmic in solution.

NMR spectra of the ^{15}N-labeled J1_tmic confirmed its properties in solution (**Figure 2.8 A**). When dissolved in water, 115 backbone NH cross-peaks (~92% of the total) of the 125 expected (133 residues of which 8 prolines) could be detected in the HSQC spectrum. Most of the peaks show extensive line broadening and resonates in a narrow region of between 7.7 and 8.4 p.p.m. The average value of HN chemical shifts is 8.08 p.p.m. with dispersion (r) of 0.22 p.p.m. For comparison, the random coil values for a protein of the same amino acid composition would have an average of 8.18 p.p.m and a dispersion of 0.16 p.p.m. (**Figure 2.17**). The presence of strong and sharp resonances accompanied by much weaker peaks in the ^{1}H-^{15}N HSQC spectrum, and few peaks that could be identified in the HN-Hα region of the ^{1}H-^{15}N HSQC-TOCSY spectrum (**Figure 2.18 A**) also point to the presence of conformational exchange processes. The lack of chemical shift dispersion in the HN region as well as in the methyl region (data not shown) is an indicator of the lack of globular structure, and of little, if any, secondary structure [129].

Additionally, in-cell NMR experiments ("***Biophysical methods***") were carried out on the ^{15}N-labeled J1_tmic overexpressed in *E.coli* cells. The slurry obtained by

resuspending the *E.coli* pellet in PBS gives a HSQC spectrum (**Figure 2.8 B**) that is very similar to that of the purified J1_tmic dissolved in PBS solution, i.e. in the same pH and ionic strength conditions (**Figure 2.8 C**).

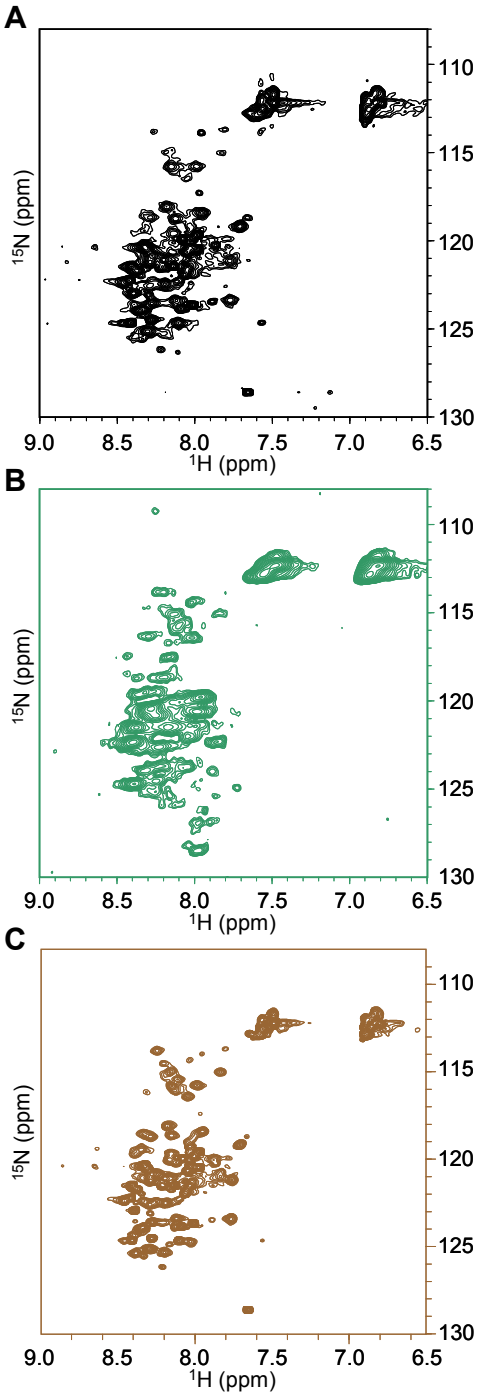


Figure 2.8. **In-cell NMR of J1_tmic.** ^{1}H-^{15}N HSQC spectra of (**A**) the purified J1_tmic protein (0.5 mM) in H_2O/D_2O (90/10, v/v), pH 7.2, (**B**) from in-cell experiments, pH 7.2 (**C**) of the purified protein (0.5 mM) in PBS buffer/D_2O (90/10, v/v), pH 7.2.

In the in-cell HSQC spectrum the intensity of NH peaks is lower and peaks show more line-broadening. However, chemical shift dispersion in the NH region is comparable to that of the purified protein. The in-cell HSQC spectrum is also very similar to the spectrum of J1_tmic dissolved in water solution (**Figure 2.8 A**). Not induced *E. coli* cells were used as negative control (**Figure S1 A**). The almost complete vanishing of the NMR signal confirms that NH peaks in the figure come from amides of the overexpressed J1_tmic. Moreover, the HSQC spectrum of the supernatant (obtained during the bacterial slurry preparation (**Figure S1 C**)) shows a negligible signal, confirming that the J1_tmic was in the cell during the recording of the spectrum.

The lack of chemical shift dispersion in the HSQC spectrum obtained from in-cell NMR experiments is a further confirmation of the lack of globular structure, even in the molecular crowding conditions of a cell-like environment [130]. Moreover, this result suggests that the standard protein-harvesting (cell-lysis and sonication) applied to obtain the final purified protein, do not change the J1_tmic conformation.

To better characterize the conformation of J1_tmic in solution, its hydrodynamic properties were studied by size exclusion chromatography (S.E.C.). Gel filtration columns, generally used for S.E.C, separate proteins on the ground of their hydrodynamic size, measured by the Stokes radius (R_S). The apparent molecular weight (MW) of J1_tmic is determined from the protein elution volume which is compared to the elution volumes of a set of standard globular proteins, used for the column calibration (**Figure 2.9**). J1_tmic (15.5 kDa) was eluted as a peak corresponding to a 25.6 kDa globular protein. The observed mass is significantly larger, but lower than the double, of the calculated J1_tmic mass. The possibility of protein dimerization can be thus excluded. Moreover, the eluted peak has a very sharp and symmetrical form indicating the presence of a single and well-defined species. The calculated R_S for an apparent mass of 25.6 kDa is 23.57 ± 0.35 Å. This is slightly larger than the calculated value (R_SN = 19.6 ± 0.3 Å) for a globular protein with the same number of residues as J1_tmic, but considerably smaller than the expected value for a completely extended chain (R_SU = 36.4 ± 0.7 Å) as can be measured in denaturing conditions [131]. The structural data on J1_tmic collected by CD, size exclusion chromatography and NMR are consistent with a mainly disordered, but rather compact, state of the protein in solution, and the presence of very little or no secondary structure.

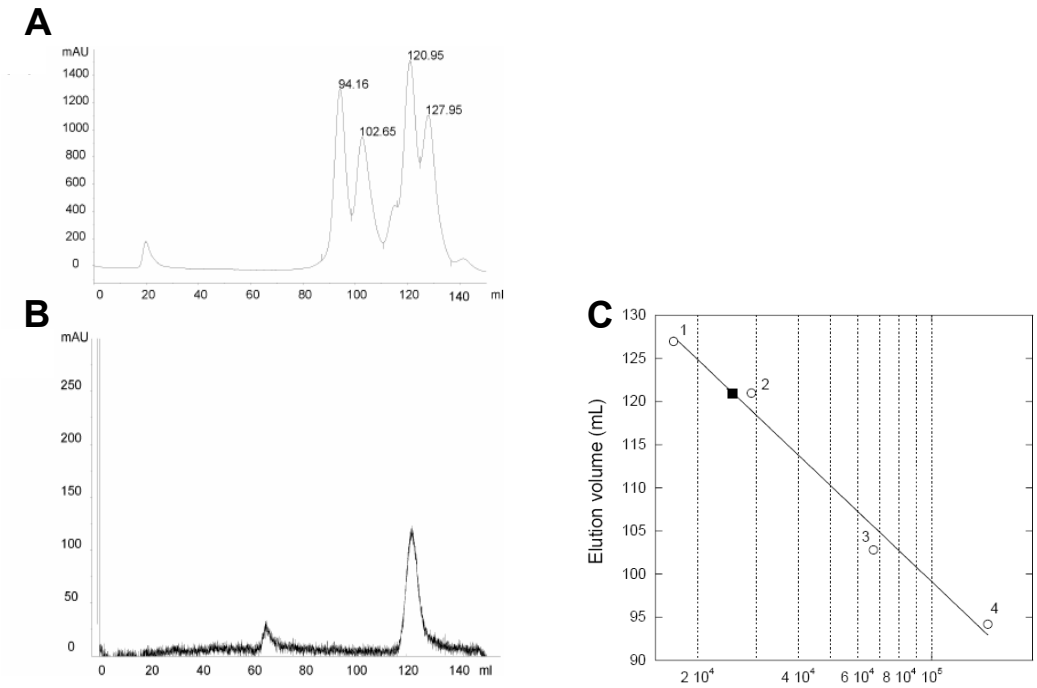


Figure 2.9. **Size exclusion chromatography.** Chromatograms (UV detection at 215 nm) of (**A**) the calibration standards and of (**B**) J1_tmic. In (**C**), calibration standards are marked as open circles (1, horse myoglobin (17 kDa); 2, carbonic anhydrase (29 kDa); 3, bovine serum albumin (67 kDa); 4, lactate dehydrogenase (147 kDa)), J1_tmic as a filled square (apparent MW = 25.6 kDa); the calibration curve is also shown.

2.1.4 J1_tmic exhibits intrinsic helical propensity

J1_tmic is predicted to adopt some secondary structure, as determined by subjecting the protein sequence to the analysis of different secondary structure predictors (PSIPRED [132], JNet [133], SSpro [134]) run from the PHYRE web server (http://www.sbg.bio.ic.ac.uk). From the consensus secondary structure prediction, four stretches of helix displaying a relatively high confidence can be identified (**Figure 2.10**). These predictions led us to speculate that the J1_tmic secondary structure might be stabilized in specific conditions. To test this possibility, the secondary structure of J1_tmic in the presence of different concentrations of trifluoroethanol (TFE) was analyzed by CD. TFE promotes secondary structure formation by reducing the protein backbone exposure to the aqueous solvent and favoring the formation of intramolecular hydrogen bonds [135]. Starting from a random-coil conformation in aqueous solution, a significant change in the secondary structure was observed upon addition of increasing amounts of TFE.

```
   VTAFYWALRKRRRKPGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVPIKDYENKNSKMSKIRTHNSE
a) hhhhhhhhhhhhccccccccccccchhhhhhhhhhhhhhhhhhhhccccccccccccccchhhhhhccccc
b) hhhhhhhhhhccccccccccccccccchhhhhhhhhhccccccccccccccccccccccceeeeeeccccc
c) chhhhhhhhhhccccccccccccccccchhhhhhhhhhhchhhhhccccccccccccccccheheeccccc
d) hhhhhhhhhhhccccccccccccccccchhhhhhhhhhhchhhhhccccccccccccccccheheehccccc
e) 688899988747888888888877765679999987547566647787665446688752434425788 8

   VEEDDMDKHQQKARFAKQPAYTLVDREEKPPNGTPTKHPNWTNKQDNRDLESAQSLNRMEYIV
a) cchhhhhhhhhhhccccccccccccccccccccccccccccccccccchhhhhhhhhhhhhhhcc
b) ccccchhhhhhhhhhccccceeeeeccccccccccccccccccccccccchhhhhhhccccceec
c) ccccchhhhhhhhhccccceeeeeccccccccccccccccccccccccchhhhhhhhcchhcc
d) ccccchhhhhhhhhccccceeeeeccccccccccccccccccccccccchhhhhhhhcchhcc
e) 875647888877655779864554466789999988899988877776567877875543337
```

Figure 2.10. **Secondary structure predictions.** Amino acid sequence of J1_tmic and secondary structure predictions (h, helix; e, β-strand; c, coil) obtained running (**A**) PSIPRED, (**B**) JNET and (**C**) SSpro from the PHYRE web server (http://www.sbg.bio.ic.ac.uk/). The consensus secondary structure (**D**) and the score (**E**) are also shown; segments with high score are highlighted in grey. Histidine residues are underlined and colored in green; the C-terminal PDZ binding motif is underlined and in bold. The N-terminal basic stretch and the central acidic stretch are in bold and colored in blue and red, respectively.

The CD spectra developed a strong ellipticity at 206 nm and a shoulder at 222 nm, characteristic of an α-helical structure, at the expense of the minimum at 198 nm, showing that TFE induces an α-helical conformation in J1_tmic (**Figure 2.11**).

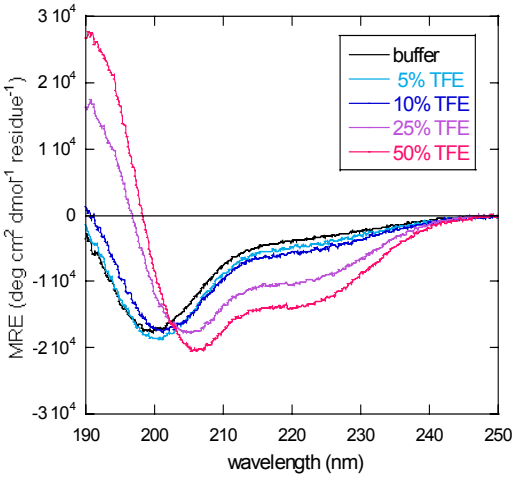


Figure 2.11. **Circular dichroism.** Far-UV CD spectra of J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer, pH 7.4, in the presence of different concentrations of 2,2,2-trifluoroethanol (TFE) (%, v⁄v).

The J1_tmic helical content increases from 4% to 50% upon TFE addition (0–50%, v/v) with a drastic change in ellipticity between 10 and 25% TFE (**Table 1**). These results confirm that J1_tmic possesses intrinsic helical propensity, and the measured α-helical content is consistent with the predicted one (23–35% for the consensus prediction, depending on the threshold set or the probability score).

2.1.5 J1_tmic binds to SDS micelles and phospholipid vesicles

The induction of the secondary structure by TFE provided a springboard for further investigation. The Jagged-1 intracellular region experiences neither the aqueous environment of the cytosol nor the hydrophobic environment of the lipid bilayer, but rather "lives" at the interface between the two. SDS micelles and lipid liposomes were used to mimic this environment.

Binding of J1_tmic to SDS micelles and phospholipid vesicles was monitored by tryptophan emission fluorescence spectroscopy and fluorescence anisotropy, taking advantage of the two tryptophans present in the sequence (**Figure 2.4**). At increasing SDS concentrations, an increase from 0.07 to 0.12 in anisotropy was observed (**Figure 2.12 A**).

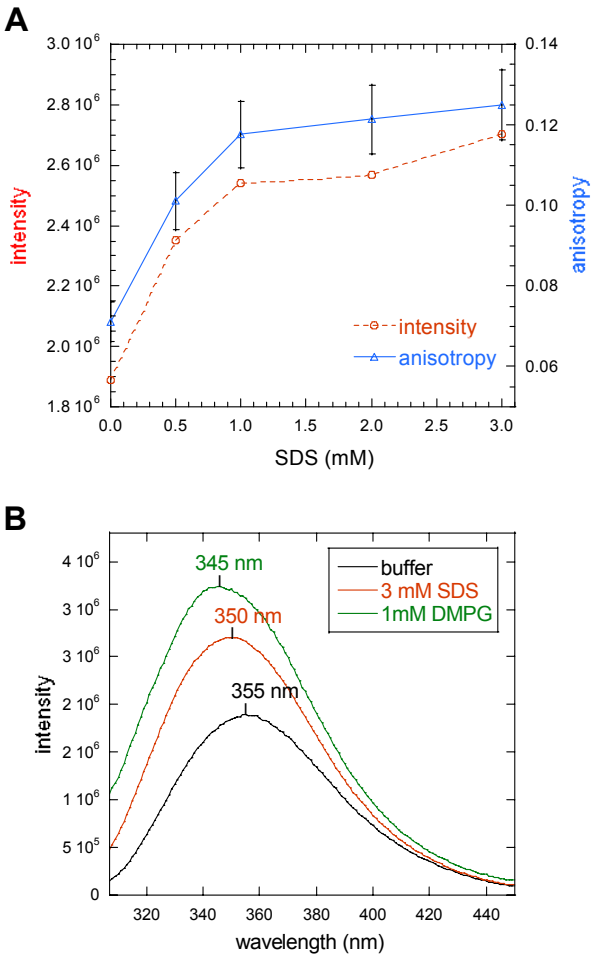


Figure 2.12. Fluorescence spectroscopy. (**A**): Tryptophan fluorescence anisotropy and emission intensity of J1_tmic (7.5 μM) in 5mM Tris⁄HCl buffer, pH 7.4, in the presence of increasing concentrations of SDS. (**B**): Tryptophan fluorescence emission spectra of J1_tmic (7.5 μM) in 5mM Tris⁄HCl buffer, pH 7.4, in the presence of SDS (3 mM), and in the presence of DMPG (1 mM) phospholipid vesicles; excitation wavelength was set to 295 nm.

At submillimolar concentrations (50–100 μM SDS) abnormally high anisotropy values were observed (data not shown), probably due to scattering associated with solution turbidity, which, however, disappeared at higher SDS concentrations (> 0.5 mM). The tryptophan anisotropy is mainly determined by the rotational motion of the fluorophore and proportional to the size of the rotating molecule. The anisotropy increase indicates the formation of the protein-micelle complex. Tryptophan fluorescence emission spectra showed an increase in intensity and a blue-shift of the maximum (λ_{MAX}) from 355 to 350 nm in the presence of SDS (**Figure 2.12 B**). In the presence of 1,2-dimyristoyl-sn-glycero-3-[phospho-*rac*-(1-

glycerol)] sodium salt (DMPG) phospholipid vesicles, changes were even more evident, with a marked increase in the emission intensity and a blue-shift from 355 to 345 nm. The increase of the tryptophan emission intensity and the blue-shift of its maximum is a consequence of the chromophore' s transfer from an aqueous to a non-polar environment.

Altogether, fluorescence data confirmed binding of J1_tmic to SDS micelles and DMPG phospholipid vesicles, with at least partial embedding of one or both the tryptophan residues in a more hydrophobic environment [136], [137]. As J1_tmic contains two tryptophans, W1091 in the N-terminal transmembrane region and W1196 in the C-terminal region, similar experiments were repeated on J1_ic, which lacks the transmembrane segment and thus contains only W1196 (**Figure 2.4**). Upon addition of SDS, a similar increase in the anisotropy (**Figure 2.13 A**), but a higher λ_{MAX} blue-shift (from 356 to 346 nm, at 3 mM SDS) (**Figure 2.13 B**) were observed.

However, unlike in J1_tmic, the emission spectra of J1_ic showed an intensity decrease in the presence of SDS micelles. In the presence of DMPG phospholipid vesicles, the J1_ic spectrum is characterized with the λ_{MAX} blue-shift (from 356 to 345 nm) and the increase in the emission intensity. However, the intensity increase is not evident as in the case of J1_tmic. The discrepancies between fluorescence spectra of J1_tmic and J1_ic can be explained by the different environments experienced by W1091 and W1196.

The fluorescence emission intensity is the sum of the contributions from all tryptophan residues. Comparing the spectra of the two proteins in buffer, the intensity of the J1_tmic spectrum is nearly two fold larger than that of J1_ic ($1.89*10^6$ and $0.97*10^6$, respectively). This suggests that W1091 and W1196 are equally exposed to the buffer and thus experience the same environment. If the two tryptophans were also equally exposed in DMPG vesicles, the same intensity ratio ($I^{J1\_tmic} : I^{J1\_ic} = 2 : 1$) would be expected. However, the emission intensity for J1_tmic and J1_ic is $3.24*10^6$ and $1.31*10^6$ respectively, suggesting that W1091 and W1196 do not interact with DMPG vesicles in the same way. It can be speculated that while the J1_tmic N-terminal hydrophobic segment (VTAFYWAL) is embedded in the hydrophobic environment of the lipid alkyl chains, W1196 most probably remains in the interfacial region of the bilayer (**Figure S2**).

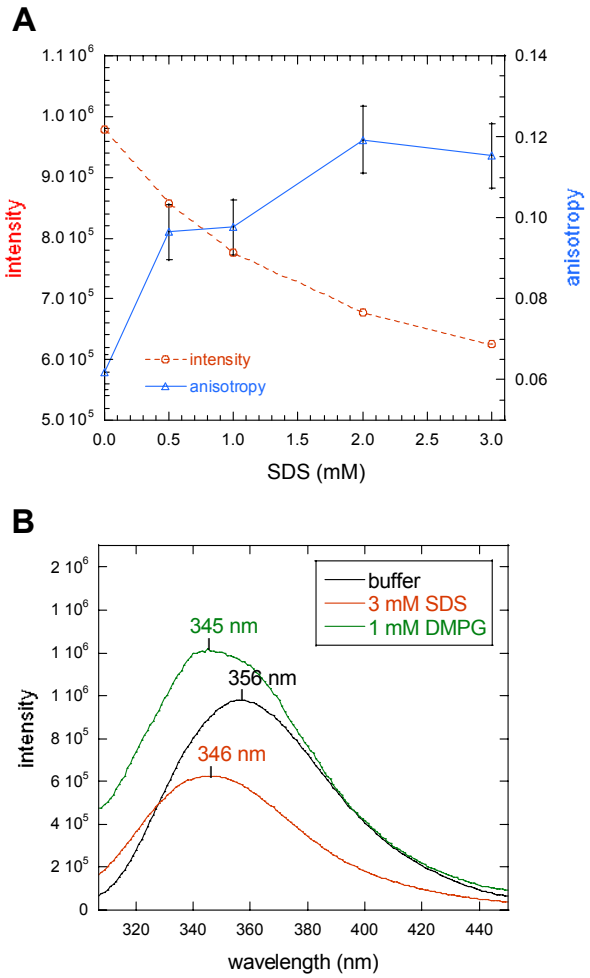


Figure 2.13. Fluorescence spectroscopy. (**A**): Tryptophan fluorescence anisotropy and emission intensity of J1_ic (7.0 µM) in 5 mM Tris/HCl buffer, pH 7.4, in the presence of increasing concentrations of SDS. (**B**): Tryptophan fluorescence emission spectra of J1_ic (7.0 µM) in 5 mM Tris/HCl buffer, pH 7.4, in the presence of SDS (3 mM), and in the presence of DMPG (1 mM) phospholipid vesicles; excitation wavelength was set to 295 nm.

In the presence of SDS micelle, the difference between the J1_tmic and J1_ic spectra is more evident. Binding of J1_ic to SDS micelles is associated with both a decrease in fluorescence intensity and a blue shift in its emission maximum. The presence of these two characteristics together is contradictory, since a blue shift indicates a transfer to a more hydrophobic environment, whereas a decrease in the intensity indicates a transfer to a more hydrophilic environment. A similar situation was however described for the tryptophan fluorescence of smooth muscle myosin upon binding with nucleotides [138]. At the same time, in the presence of SDS

micelles, the emission spectrum of J1_tmic shows a smaller blue-shift in λ_{MAX} (from 355 nm to 349 nm) compared to that in the presence of DMPG liposomes.

In conclusion, W1091 and W1196 interact with SDS micelle in a slightly different way than with DMPG liposomes. Moreover, there is an evident difference between these two tryptophans when they bind to SDS micelle. DMPG and SDS share hydrophobic chains of the same length and both possess negatively charged head groups. However, chemical (double chain hydrophobic tail and a larger polar head vs. single chain hydrophobic tail and a small polar head) and structural (low curvature bilayers vs. high curvature spherical particles) differences between SDS micelles and DMPG liposomes should be also taken into account, and can contribute to tryptophan fluorescence in a somewhat unpredictable way.

Although these results are not conclusive with respect to the determination of the precise environment of the two tryptophans, they show that both J1_tmic and J1_ic bind to SDS micelles and DMPG phospholipid vesicles, and thus that the transmembrane region of J1_tmic is not absolutely required for binding.

2.1.6 J1_tmic gains helical structure upon binding to SDS micelles

The secondary structure of J1_tmic in the presence of SDS micelles was analyzed by CD. As already seen with TFE, at increasing concentrations of SDS J1_tmic undergoes a significant conformational change towards an α-helical structure, reaching a maximum of ~17% of α-helix at 3 mM SDS (**Figure 2.14**), as estimated from CDSSTR (**Table 1**). The titration with SDS in the range below 0.5 mM does not show any ellipticity increase at 206 nm and 222 nm, yet the CD signal decreases in the whole spectrum (190-250 nm) (**Figure S3**). The most evident transition towards the final CD spectrum can be noticed from 0.5 mM to 1 mM SDS. On further increase in SDS concentration, CD spectra show only little change, reaching saturation at 3 mM SDS. The effect of the helical transition starts when the molar ratio surfactant/protein is 66:1 (0.5 mM SDS and 7.5 μM J1tmic). As one SDS micelle consists of ~60 molecules, in the above experiment the ratio of micelle and protein is equimolar. The saturation at 1-3 mM SDS agrees with the above described fluorescence anisotropy results (**Figure 2.12 A**), suggesting that binding to SDS micelle correlates with the induced helical formation.

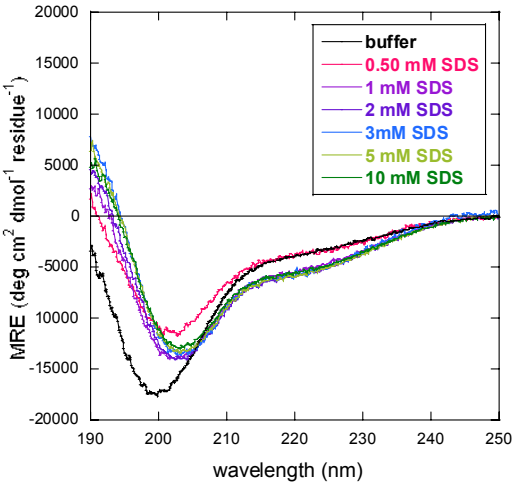


Figure 2.14. Circular dichroism in the presence of SDS. Far-UV CD spectra of J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer, pH 7.4 at different concentrations of SDS (from 0.5 mM to 10 mM).

At the same SDS concentration (3 mM), the α-helical content reversibly increases as the pH decreases (17% of α-helix at pH 7.4 versus 33% at pH 6), whereas the same pH change does not induce any significant change in the α-helical content in the protein in the absence of SDS (**Figure 2.15**), (**Figure S4**). The fluorescence emission spectra of J1_tmic in the presence of 3 mM SDS at pH 6 and pH 7, however, are not significantly different (**Figure S5**).

The conformation of J1_tmic in the presence of SDS was further analyzed by NMR spectroscopy. The ^{1}H-^{15}N HSQC spectrum of J1_tmic obtained in the presence of SDS micelles is somewhat different from that of the protein alone (**Figure 2.16**). Although several resonances are still missing, probably due to overlap, NH cross-peaks appear to be of similar intensity and slightly better dispersed. Most NH backbone resonances are still clustered in a relatively narrow region (7.5–8.5 p.p.m.), but the average value of NH chemical shifts (7.97 p.p.m.) is smaller and the dispersion slightly larger (σ = 0.25) compared to the values obtained for the protein alone (**Figure 2.17**). Most of the expected cross-peaks in the Hα region of the ^{1}H-^{15}N HSQC-TOCSY spectrum are still missing (**Figure 2.18**). The lack of significant chemical shift dispersion in the NH and Hα chemical shifts is an evidence of lack of tertiary structure. On the other hand, NMR spectra suggest that the conformation of J1_tmic is at least partially restrained in the presence of SDS micelles.

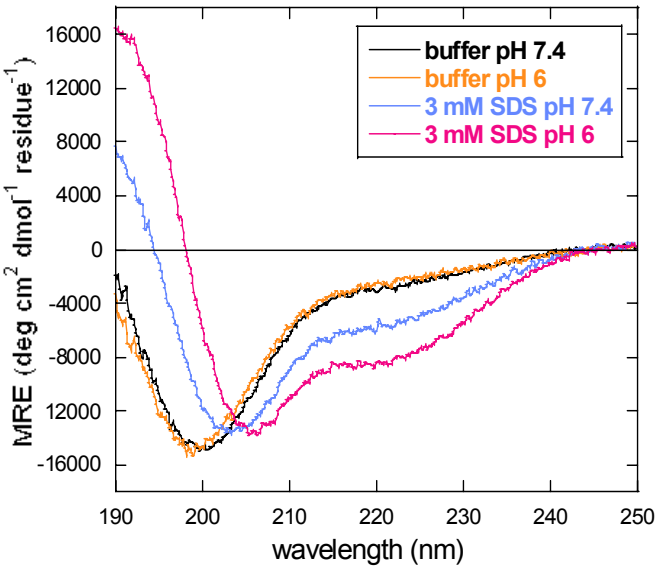


Figure 2.15. Circular dichroism in the presence of SDS – pH effect. Far-UV CD spectra of J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer, at different pH values (7.4 and 6.0) in buffer alone and in the presence of SDS (3 mM).

At a lower pH, the appearance of both the ^{1}H-^{15}N HSQC and the ^{1}H-^{15}N HSQC-TOCSY spectra is markedly different (**Figure 2.16 B**), (**Figure 2.18 C**). Most of the expected NH cross-peaks (93%) and of the Hα peaks could be detected, and lines are much narrower than at pH 7. The average chemical shift of backbone amides is 8.04 p.p.m. and the dispersion 0.23 p.p.m. Also in these conditions, however, the lack of chemical shift dispersion points to the absence of tertiary structure.

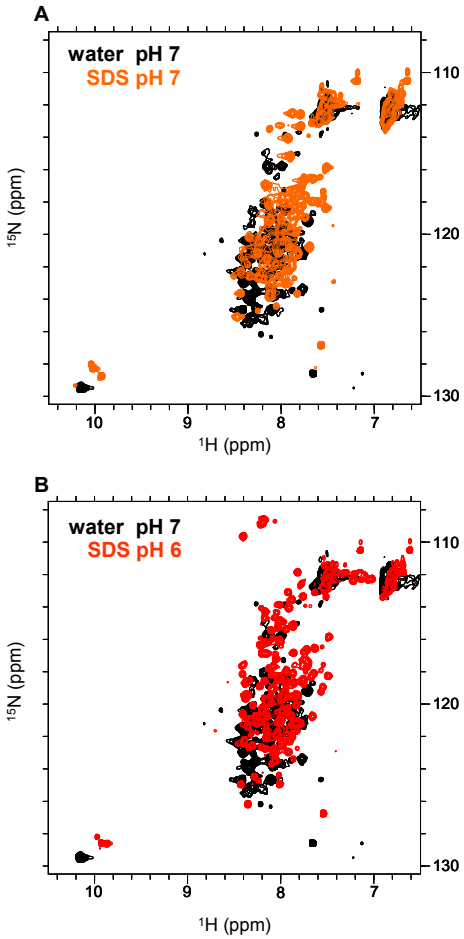


Figure 2.16. NMR spectroscopy. ^{1}H-^{15}N HSQC spectrum of J1_tmic (0.5 mM) in water, pH 7 (black) overlaid with the ^{1}H-^{15}N HSQC spectrum of J1_tmic (0.5 mM) in the presence (**A**) of SDS (50 mM), pH 7.0 (orange) and in the presence (**B**) of SDS (50 mM), pH 6.0 (red).

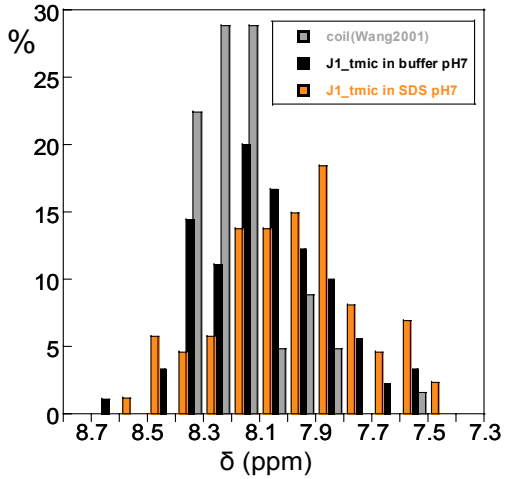


Figure 2.17. Chemical shift dispersion. ^{1}H chemical shift dispersion of detectable amides of J1_tmic in water (black bars), of J1_tmic in the presence of SDS (orange bars), and of random coil values for a protein of the same sequence (grey bars). Chemical shift values were estimated from ^{1}H-^{15}N HSQC spectra and referenced to DSS.

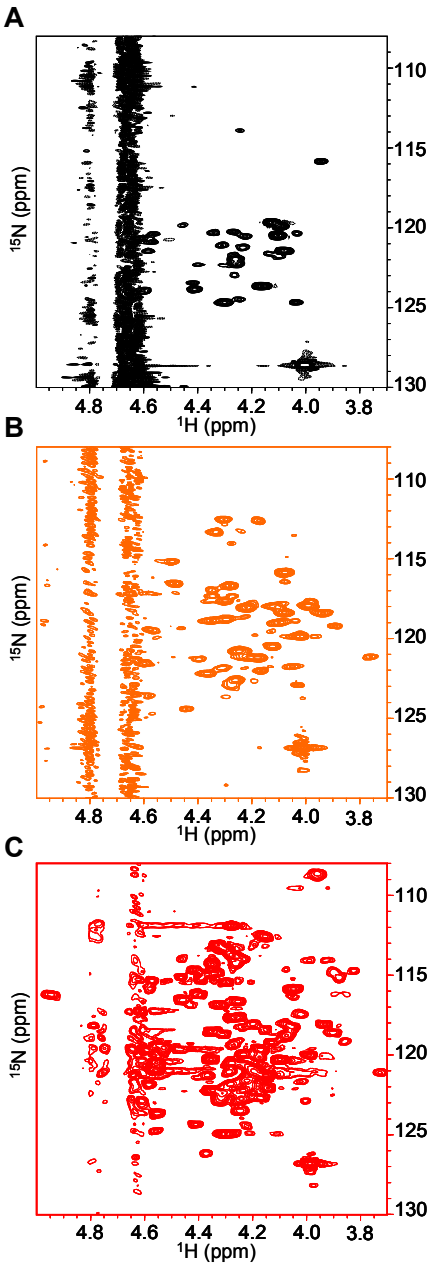


Figure 2.18. NMR spectroscopy. The Hα region of the ^{1}H-^{15}N HSQC-TOCSY spectra of J1_tmic (0.5 mM) in (**A**) water pH 7, (**B**) in the presence of SDS (50 mM), pH 7.0 and (**C**) in the presence of SDS (50 mM), pH 6.0.

2.17 J1_tmic gains helical structure upon binding to negatively charged phospholipid vesicles

The secondary structure of J1_tmic was analyzed in the presence of vesicles prepared from various phospholipids that are typical components of eukaryotic cell membranes. The far-UV CD of J1_tmic in the presence of vesicles prepared from the negatively charged phospholipids DMPG or dimyristoylphosphatidylserine (DMPS)

(**Figure 2.19**) showed spectral variations similar to those obtained in the presence of SDS micelles (**Figure 2.14**). The estimated α-helical content was 19% and 17% in the presence of DMPG and DMPS vesicles, respectively (lipid concentration: 1 mM; protein/lipid molar ratio = 1:130). On the contrary, no change could be detected in the presence of vesicles made of the zwitterionic phospholipid dimyristoylphosphatidylcholine (DMPC).

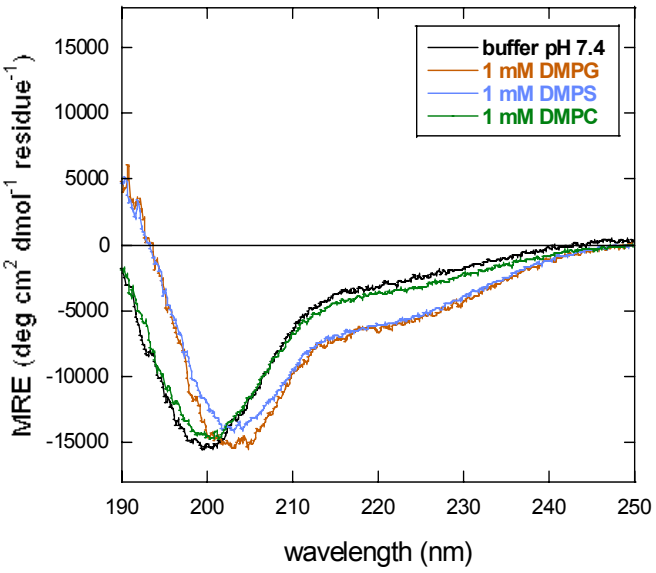


Figure 2.19. **Circular dichroism in the presence of phospholipid vesicles.** Far-UV CD spectra of J1_tmic (7.5 μM) in 5mM Tris/HCl buffer, pH 7.4, in the presence of DMPG, DMPS, or DMPC phospholipid vesicles.

DMPG, and not DMPS, were used in further CD analysis. DMPG liposomes have a lower phase transition temperature (23°C) compared to DMPS (35°C) and are thus more stable and much easier to use. Additionally, the head group of DMPG is more stable in an acidic environment, maintaining its negative charge at pH>3 (pKa of the phosphate group is 2.9), while DMPS becomes neutral at pH<5.5 because of the pKa=5.5 of its PS carboxyl group [139]. In the presence of DMPG phospholipid vesicles, a decrease in pH from 7.4 to 6.0 led to a reversible increase in the helical content of J1_tmic from 19 to 36% (**Figure 2.20**), (**Table 1**), (**Figure S6**). Far-UV CD measurements were carried out also on J1_ic. Upon addition of 3 mM SDS or 1 mM DMPG, J1_ic showed an increase of its helical structure very similar to that of J1_tmic (**Figure S7**), (**Figure S8**), (**Table 1**). When in the above membrane

mimicking environments the pH was decreased from 7.4 to 6.0, a remarkable difference in behavior between J1_tmic and J1_ic was observed. While in J1_tmic the pH decrease almost doubles the helical content, in J1_ic the helicity increase is clearly lower. The helical content increased from 17% to 23% in the presence of SDS micelles and from 17% to 25% in the presence of DMPG liposomes.

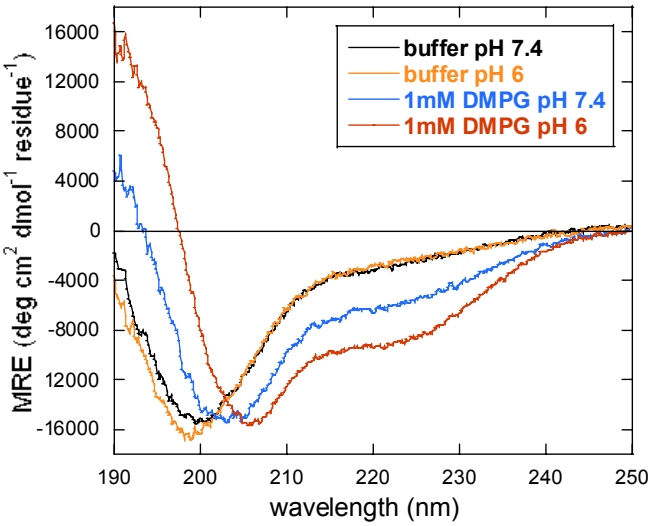


Figure 2.20. **Circular dichroism in the presence of DMPG**. Far-UV CD spectra of J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer and in the presence of DMPG (1 mM) phospholipid vesicles at pH 7.4 and pH 6.0.

2.1.8 J1_tmic in the presence of lipid micelles – NMR studies

To get a better insight into the structural and functional properties of J1_tmic, it would then be necessary (i) to determine, at the residue level, which regions become helical upon interaction with the inner leaflet of the plasma membrane, (ii) to map the interface between the protein and the membrane and (iii) to measure the relative orientation of the helices.

Preliminary NMR experiments on ^{15}N-labeled J1_tmic were carried out, using lysophospholipid micelles as a model of the interface between the plasma membrane and the cytosol (**Figure 2.21**). Like SDS micelles, lysophospholipid micelles are prepared in a straightforward manner and are relatively stable in time in the temperature and pH conditions used for acquisition of NMR data. Moreover, lysophospholipid micelles are biologically more relevant than SDS micelles, as their

head groups are identical to those found in natural phospholipids. Although liposomes are often considered as a better membrane mimicking system because of their low curvature, they are not suitable for high-resolution NMR techniques due to low reorientation rates that lead to significant line broadening.

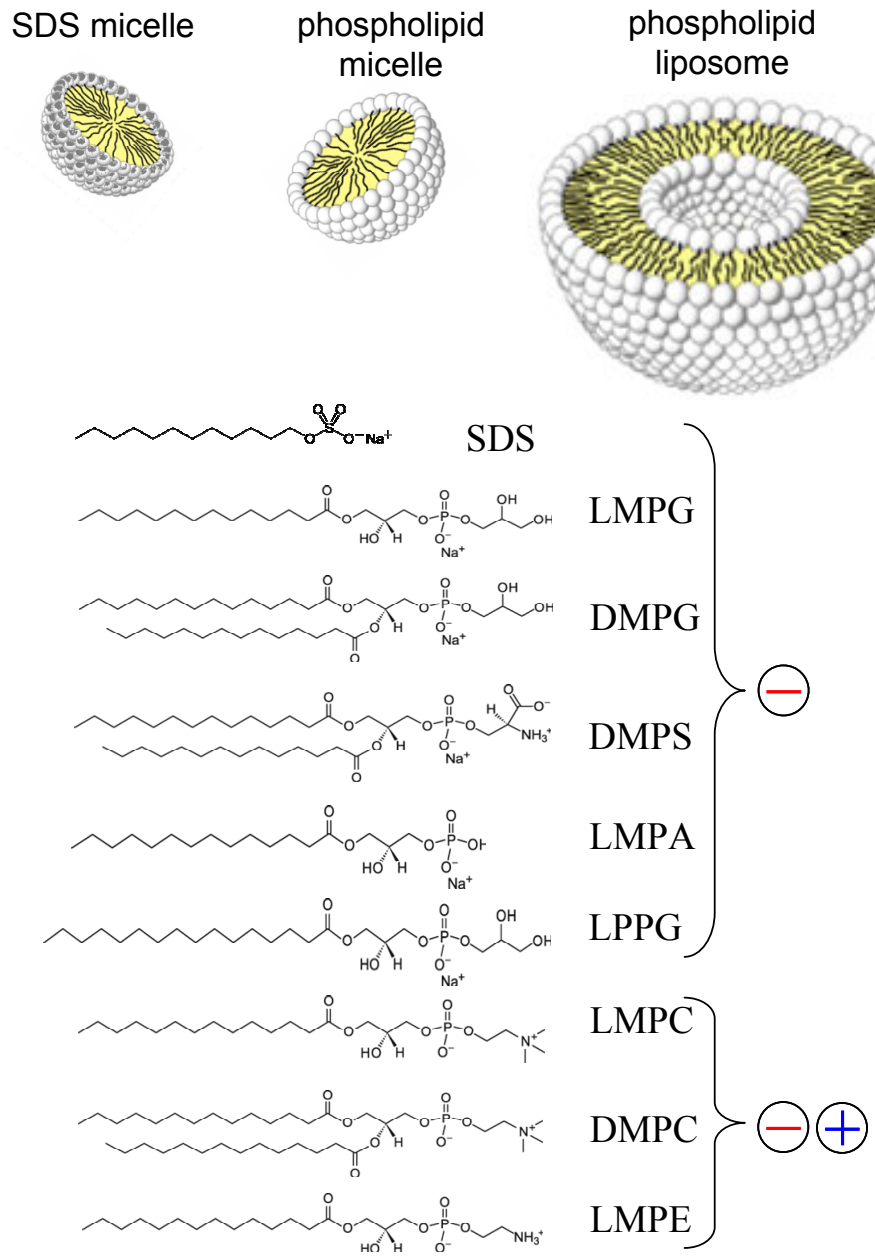


Figure 2.21. **Membrane models**. The top part of the figure shows cross sections of micelles and liposomes; the white beads represent the polar heads of detergents/lipids and the black tails represent the alkyl hydrophobic chains. The bottom part of the figure shows the chemical structures of phospholipids; phospholipids are grouped according to the charge of the polar head.

Preliminary analyses were performed using far-UV circular dichroism on J1_tmic in the presence of different lysophospholipid micelles (**Figure 2.22**). The spectra of J1_tmic in the presence of negatively charged micelles (LMPG, LPPG and LMPA), at pH 7 as well as at pH 6, are very similar to those in the presence of DMPG liposomes (**Figure S9**). The spectra in the presence of zwitterionic (LMPC and LMPE) micelles do not differ from spectra acquired in buffer, and are similar to spectra in the presence if DMPC liposomes.

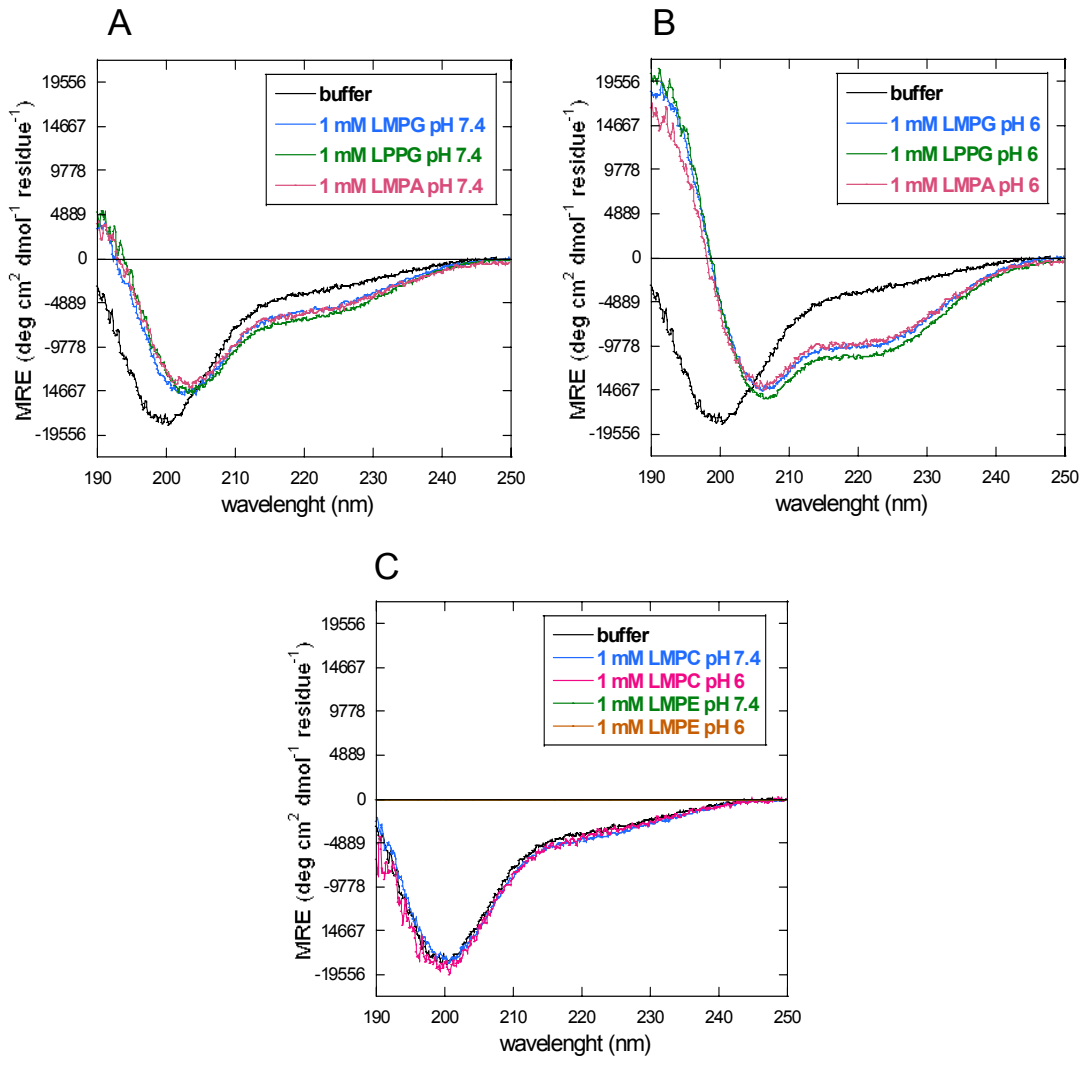


Figure 2.22. **Circular dichroism in the presence of phospholipid micelles**. Far-UV CD spectra of J1_tmic (7.5 µM) in the presence of (**A**) negatively charged lysophospholipid (LMPG, LPPG, or LMPA) micelles at pH 7.4, (**B**) negatively charged lysophospholipid (LMPG, LPPG, or LMPA) micelles at pH 6, and (**C**) zwitterionic lysophospholipid (LMPC or LMPE) micelles.

Differences in the length of the alkyl chain length (14C or 16C in LMPG or LPPG, respectively), or in the polar head group (phosphoglycerol in LMPG and LPPG, phosphate in LMPA), of the negatively charged lysophospholipids did not affect the interaction between the micelle and J1_tmic. The same applies for the zwitterionic lysophospholipids where LMPC and LMPE have a C14 long alkyl chain (of myristoyl phosphatidic acid), but different polar heads (phosphocholine in LMPC, phosphoethanolamine in LMPE). These results further confirmed that the negatively charged surface of micelles and liposomes is a prerequisite for the partial helical induction in J1_tmic. For NMR studies on J1_tmic, LMPG and LMPC lysophospholipids were chosen, as they share the same headgroup and alkyl chain length with the phospholipids used for liposomes (DMPG and DMPC, respectively).

HSQC spectra of the ^{15}N labeled J1_tmic in the presence of LMPC and LMPG micelles were acquired and compared to the HSQC acquired in buffer. In phosphate buffer (**Figure 2.23**), as in water (**Figure 2.8**), J1_tmic shows 115 backbone NH cross-peaks (~ 92% of the total), including those arising from the three glycines. The amide resonances display very limited chemical shift dispersion in the ^{1}H dimension (7.6-8.7 ppm), and several residues seem to have a major and a minor form. In the 7.2-7.3 ppm range in the ^{1}H dimension (~125 ppm in the ^{15}N dimension, probably folded), three peaks tentatively assigned to the $^{\varepsilon}$NH of the R side chains are detectable, two of which are broad and overlapped. The $^{\varepsilon 1}$NH peaks of the two W side chains appear as distinct resonances, although very close in chemical shift (10.09 and 10.11 ppm). The amide resonances of the 21 N/Q side chains are all clustered in a narrow region.

The HSQC spectrum of ^{15}N-J1_tmic in the presence of LMPC micelles at pH 6 is very similar to that in buffer at the same pH (**Figure 2.23**), with no or very little chemical shift variation for most of the resonances. The only significant differences are the doubling of W side chain $^{\varepsilon 1}$NHs with two new peaks at 10.30 and 10.35 ppm, suggesting the presence of a major and a minor form and the appearance of three peaks in the 7.5-7.7 ppm range (7.51/118.76; 7.58,/117.28; 7.71,/118.27). Altogether, this HSQC spectrum shows that in the presence of LMPC micelles J1_ tmic does not acquire globular structure neither any noticeable secondary structure elements, which is in accordance with above CD results. The appearance of a few new peaks and the doubling of W side chain $^{\varepsilon 1}$NH resonances suggest that, although remaining mainly

unstructured, J1_tmic could interact, through the hydrophobic residues in its N-terminal region, with the hydrophobic core of the lysophospholipid micelle. This interaction is likely to be local, as most of the other resonances are not affected. This is in agreement with the results obtained by circular dichroism (CD) where little or no difference was detected between the far-UV CD spectrum of J1_tmic in buffer and that of the protein in the presence of LMPC micelles.

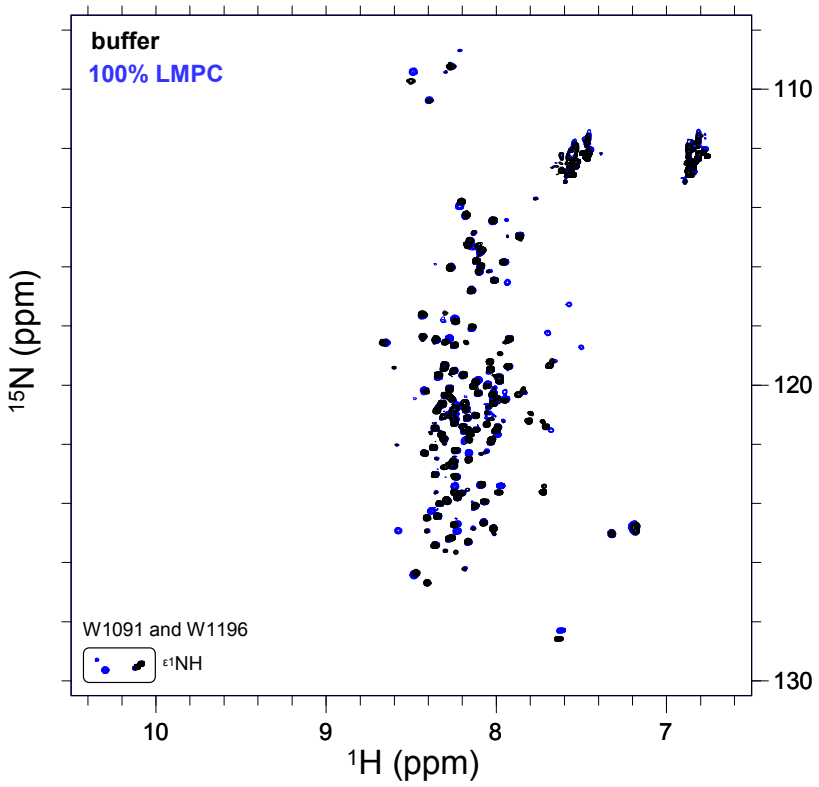


Figure 2.23. NMR spectroscopy. Overlay of ^{1}H-^{15}N HSQC spectra of J1_tmic (0.5 mM) in buffer (black) and in the presence of LMPC (80 mM) micelles.

On the contrary, the HSQC spectrum of ^{15}N-J1_tmic in the presence of LMPG micelles (at pH 6) is quite different (**Figure 2.24**), with 119 detectable backbone NH peaks, including those arising from the three glycines (~ 95% of the total), a new set of resonances appearing in the 7.5-7.8/116.3-118.5 (^{1}H/^{15}N) ppm region and a set of at least 6 distinct peaks possibly arising from the $^{\varepsilon}$NH of the arginine side chains. The amide resonances of the N/Q side chains show larger chemical shift dispersion and many of them appear as distinct pairs of peaks. Three distinct W side chain $^{\varepsilon 1}$NH

peaks can be identified, one sharper and more intense at 10.15 ppm, the other two at 10.25 and 10.31 ppm. The differences observed in the HSQC spectrum of J1_tmic in the presence of LMPG micelles show that a significant interaction is occurring between the protein and the micelle, and that possibly a conformational change is induced by the binding of J1_tmic to the LMPG micelle. These differences are in agreement with far-UV CD spectra, which show a clear coil-helix transition of J1_tmic in the presence of LMPG micelles.

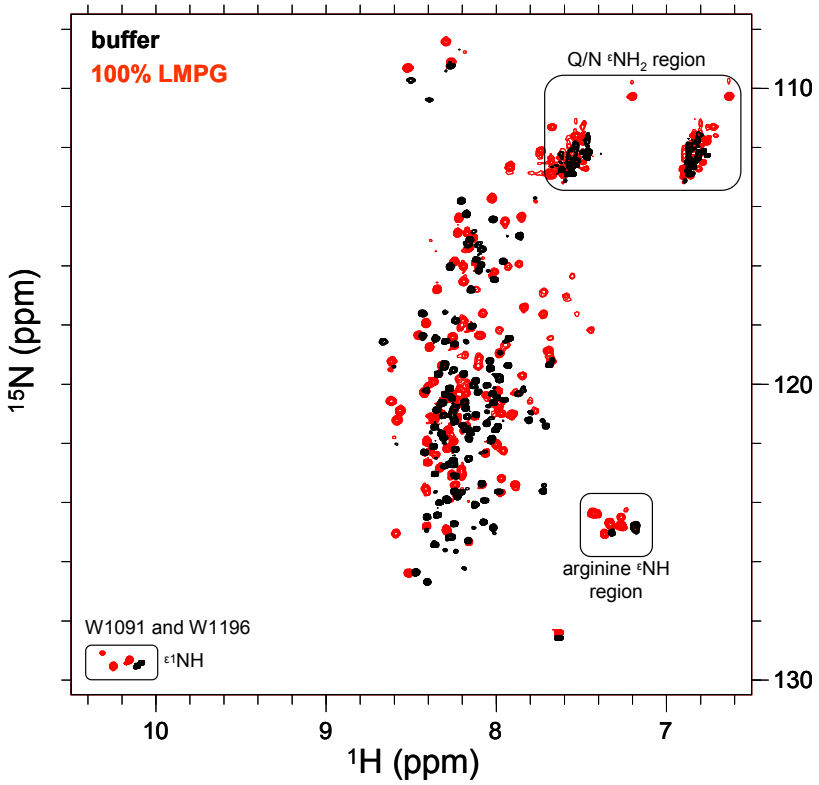


Figure 2.24. NMR spectroscopy. Overlay of ^{1}H-^{15}N HSQC spectra of J1_tmic (0.5 mM) in buffer (black) and in the presence of LMPG (80 mM) micelles.

In the attempt to get a better insight into this system, preliminary ^{15}N relaxation studies were carried out. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates were measured on the ^{15}N-J1_tmic sample in the presence of LMPG micelles (**Figure 2.25**). R_1 relaxation rates are in the range 0.83-1.93 s^{-1}, with an average of 1.48 s^{-1} and a standard deviation (σ) of 0.246; for 87 of the 119 peaks (73%) R_1 values are in the 1.30-1.70 s^{-1} range. R_2 relaxation rates display a much wider variability, varying between 2.46 and 25.9 s^{-1}, with an average of 10.2 s^{-1} and

σ=5.96. The distribution of R_2 values is quite complex, with 14 peaks (13%) showing very slow R_2 (<4.0 s^{-1}), a set of 42 over 106 peaks (40%) showing R_2 values in the 4.0-9.0 s^{-1} range, a second set of 18 peaks (17%) with R_2 in the 11.0-14.0 s^{-1} range, and a third set of 10 peaks (~9%) in the 17.0-20.0 s^{-1} range. For 13 peaks, R_2 values could not be determined due to vanishingly small or negative values of the peak intensity at longer delays. The ratio R_2/R_1 was also calculated as a rough estimation of the local correlation time (τ_c) assuming isotropic motion. Values are spread between 1.78 and 25.5, with a mean value of 7.28 and a large dispersion (σ = 5.39).

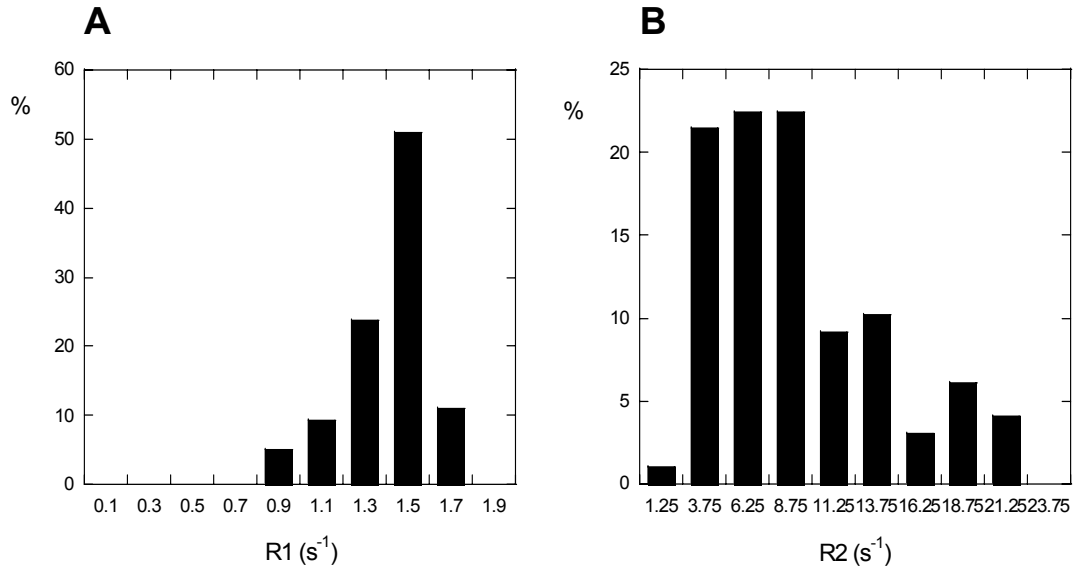


Figure 2.25. NMR spectroscopy. Distribution of relaxation rates for the amide peaks of J1_tmic in the presence of LMPG micelles: (**A**) R1 rates and (**B**) R2 rates.

While R_1 values showed a relatively narrow distribution, R_2 values are widely variable. This can be due very different local correlation times, to strong anisotropy, and to exchange processes and, given the complexity of this molecular system, all these mechanisms can contribute at the same time. A possible scenario consistent with the wide range of R_2 observed is given by the simultaneous presence of regions that are partially or totally embedded in the micelle, other regions that bind the surface of the micelle, probably in a dynamical way, and other regions that do not significantly interact with the micelle and freely tumble in solution.

2.1.9 J1_ic as a potential zinc-binding protein

As J1_tmic has 6 endogenous histidine residues, and was found to bind to Ni^{2+}-loaded IMAC columns, it was hypothesized that it might behave as a Zn^{2+}-binding protein. As preliminary tests to evaluate the potential metal binding properties, far-UV CD spectra of J1_ic in the presence of a molar excess of either EDTA or Zn^{2+} were recorded (**Figure 2.26**). No remarkable changes in the shape of the CD spectra were observed in either case. However, the addition of EDTA induced a slight decrease (more negative MRE values) in the band at 200 nm. This change is reversible, in that the original CD spectrum could be restored by a molar excess of Zn^{2+}. Addition of Zn^{2+} to the original sample did not affect the CD spectrum. J1_tmic, titrated with EDTA and Zn^{2+}, gave similar results (data not shown). These results suggest that the Jagged-1 cytoplasmic tail might behave as a zinc-binder, but binding does not induce significant conformational changes or folding of the protein. Further studies are required to understand its behavior in the presence of metal ions.

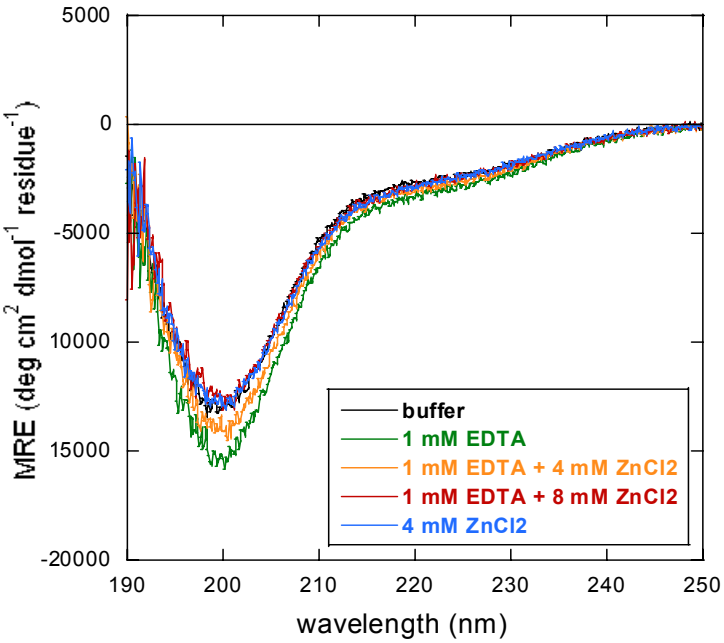


Figure 2.26. Circular dichroism. Far-UV CD spectra of J1_ic (8 μM) in 5 mM MOPS buffer, pH 7.4, in the presence of $ZnCl_2$ or EDTA.

2.2 The interaction between J1_tmic and AF6_PDZ

The goal of this study was to analyze the interaction between the Jagged-1 intracellular region and the PDZ domain of AF6 (AF6_PDZ). For this purpose, the recombinant AF6_PDZ protein and a series of peptides corresponding to the C-terminal region of Jagged-1 (the region which comprises the PDZ-binding motif and is proved to bind AF6_PDZ), were produced (**Figure 2.27 A**). Mapping of the Jagged-1/AF-6 PDZ interface was performed using NMR chemical shift perturbation experiments (***"Biophysical methods"***) and the binding constants were measured by surface plasmon resonance.

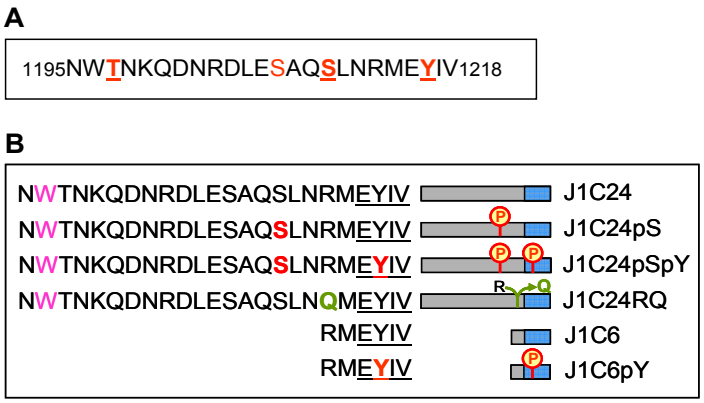


Figure 2.27. Jagged-1 C-terminal region. (**A**): potential (red) and predicted (redunderlined) phosphorylation sites. **(B):** peptides corresponding to the C-terminal region of Jagged-1; phosphorylated residues in red, mutated residues in green; the PDZ binding motif is underlined, W1196 is colored in pink.

2.2.1 Expression and purification of AF6_PDZ

The recombinant Afadin PDZ domain (AF6_PDZ, residues 987-1078 of AFAD_HUMAN, isoform 5), with a cleavable N-terminal His_6-tag was expressed in E. *coli* (**Figure 2.28 A and B**). For the preliminary NMR analysis, uniformly ^{15}N–labeled AF6_PDZ was produced in a minimum medium enriched with $^{15}NH_4Cl$. For the assignments and final NMR experiments, the ^{13}C,^{15}N–labeled AF6_PDZ was produced in a minimum medium enriched with $^{15}NH_4Cl$ and ^{13}C-glucose.

The yield of the purified unlabeled fusion protein (purity >97%) was ~10 mg/L; while in case of isotopically labeled proteins yields were ~6 mg/L. His-tag removal reduced the final protein yield by ~30%. Bearing in mind the relatively large amount of protein necessary for NMR studies and the considerable cost of ^{13}C-

glucose, only his-tagged AF6_PDZ was used. The correct fold of AF6_PDZ, with and without his-tag, was checked by far-UV CD (**Figure 2.28 C**). CD spectra are very similar. The only difference is a slight shift of the minimum towards 200 nm (his-tagged AF6_PDZ), which is probably due to the disorder in the 8 residue long N-terminal tag. Both spectra show the minimum between 205 and 220 nm, typical for an overlay of structural contributions from α-helices and anti-parallel β-strands, as present in the PDZ domain structure.

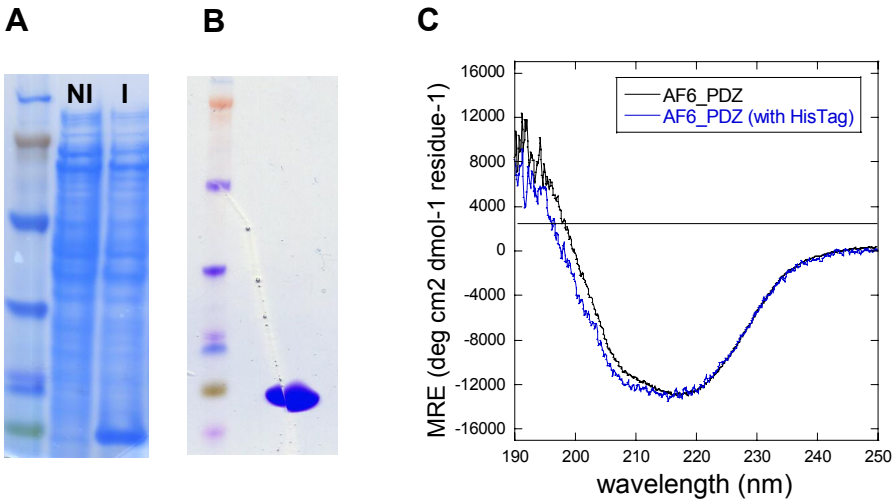


Figure 2.28. Expression and purification of AF6_PDZ. (A): SDS-PAGE of the total cell lysate from not-induced culture (NI) and after 3 h induction (I). (B): his6_AF6_PDZ after IMAC purification. (C): far-UV CD spectra of 8 μM AF6_PDZ in 5 mM Tris/HCl buffer, pH 7.4, with (blue) and without (black) N-terminal his-tag.

2.2.2 NMR experiments – AF6_PDZ resonance assignments

The preliminary ^{1}H–^{15}N HSQC spectrum of the uniformly ^{15}N–labeled AF6_PDZ shows a good dispersion in the ^{1}H chemical shifts, with the average value of 8.5 p.p.m and a chemical shift dispersion of 0.48 p.p.m., which confirms the folded state of the protein (**Figure 2.29**), (**Figure S10**). The fusion protein contains a total of 100 residues; 8 residues from the N-terminal tag (MKHHHHHH) and 92 residues from the PDZ domain (987-1078 of AFAD_HUMAN). In the HSQC spectrum, 111 NH cross-peaks should be detected, 99 corresponding to the backbone amides (proline residues do not give any cross-peak) and 12 corresponding to the sidechain amides of glutamines and asparagines. In AF6_PDZ there are 5 glutamines and 1 asparagine and each of them should show a pair of ^{15}N-^{1}H cross-peaks with identical ^{15}N-resonance.

In the $^{1}H–^{15}N$ HSQC, 108 separate peaks could be detected, among which all 12 sidechain amides. Additionally, $^{\varepsilon}NH$ peaks of 3 arginine residues are also visible.

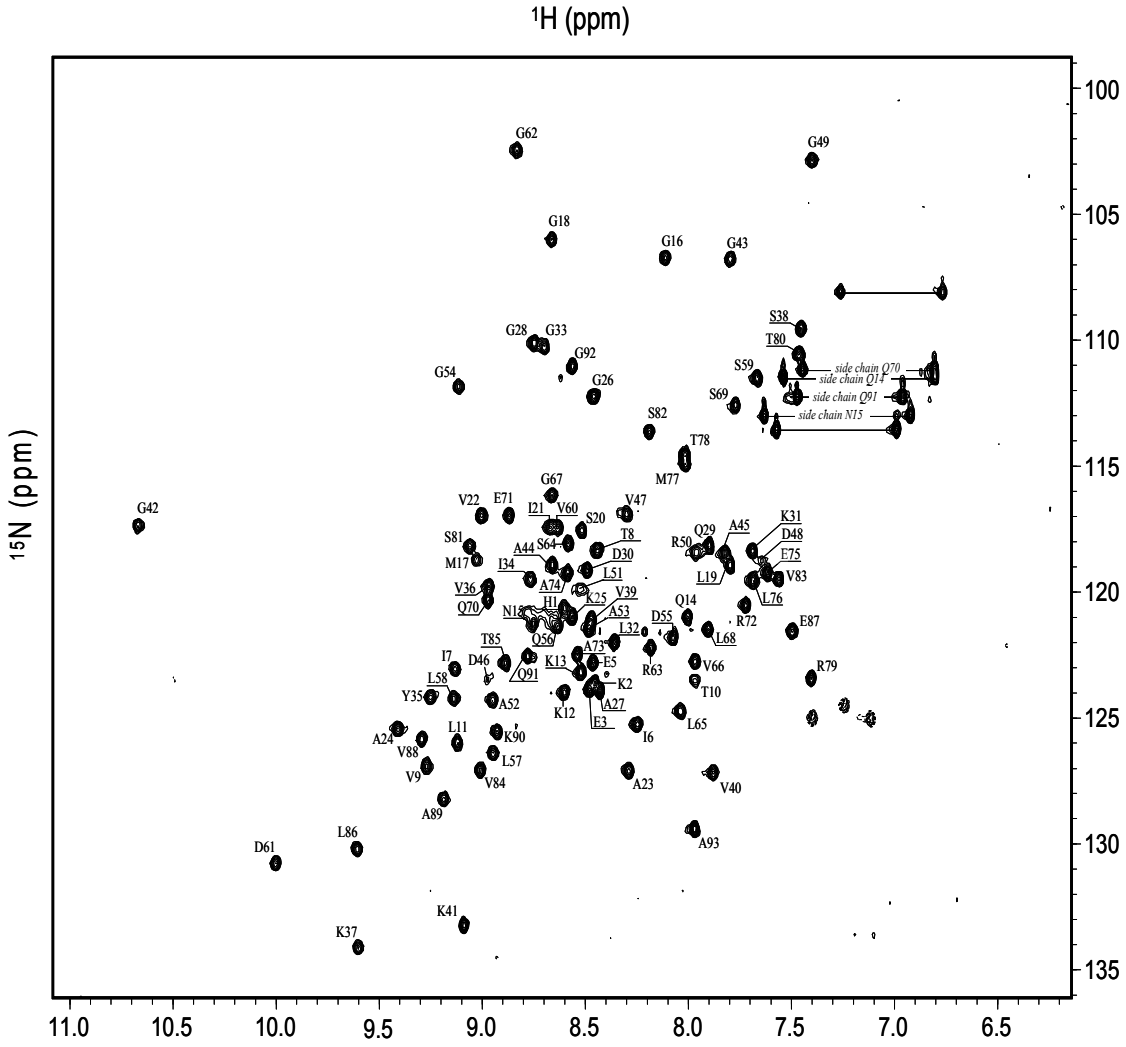


Figure 2.29. NMR of AF6_PDZ. ^{1}H-^{15}N HSQC spectrum of AF6_PDZ and sequential assignments of NH cross peaks.

To perform the chemical shift mapping analysis, it was first necessary to assign each NH cross-peak to the corresponding residue. Nearly complete backbone chemical shift assignments were obtained using the TOCSY-HSQC and NOESY-HSQC spectra acquired on the ^{15}N-labeled AF6_PDZ (**Figure 2.30**). Residues D46, D48 and L51, showing NH peaks of very low intensity, did not give any signal in the TOCSY-HSQC spectrum. The combination of CBCANH and CBCA(CO)NH spectra, acquired on the ^{13}C,^{15}N-labeled AF6_PDZ, were used to assign NH peaks. No one of the three above residues showed signals in the CBCANH spectrum, but all three gave

a signal in the CBCA(CO)NH spectrum, which was sufficient to determine their sequential position. Eventually, all backbone NHs (91 peaks) were assigned (**Table S1**). Using the CBCA(CO)NH spectrum it was also possible to assign the sidechain $^{15}N^{1}H_2$ cross-peaks of the 3 glutamines and of the single asparagine. Starting from the NH_2 sidechain resonance, asparagines show correlations with the Cβ and the Cα while glutamines with the Cγ and the Cβ resonances of the residue. This magnetization transfer is similar to the magnetization transfer from the backbone NH of i-residue to the Cα and Cβ of i-1 residue. For Q29 and Q56, because of the peak overlap of their Cβ resonances, it was not possible to assign side chain NH_2s. The his-tag resonances (MKHHHHHH) are clustered in the region of the ^{1}H–^{15}N HSQC spectrum at ^{1}H ~ 8.7±0.15 ppm and ^{15}N ~ 121±0.5 ppm, but could not be assigned with the exception of the eighth residue (H8), which was renumbered H1 and considered as the first residue of the protein for practical reasons.

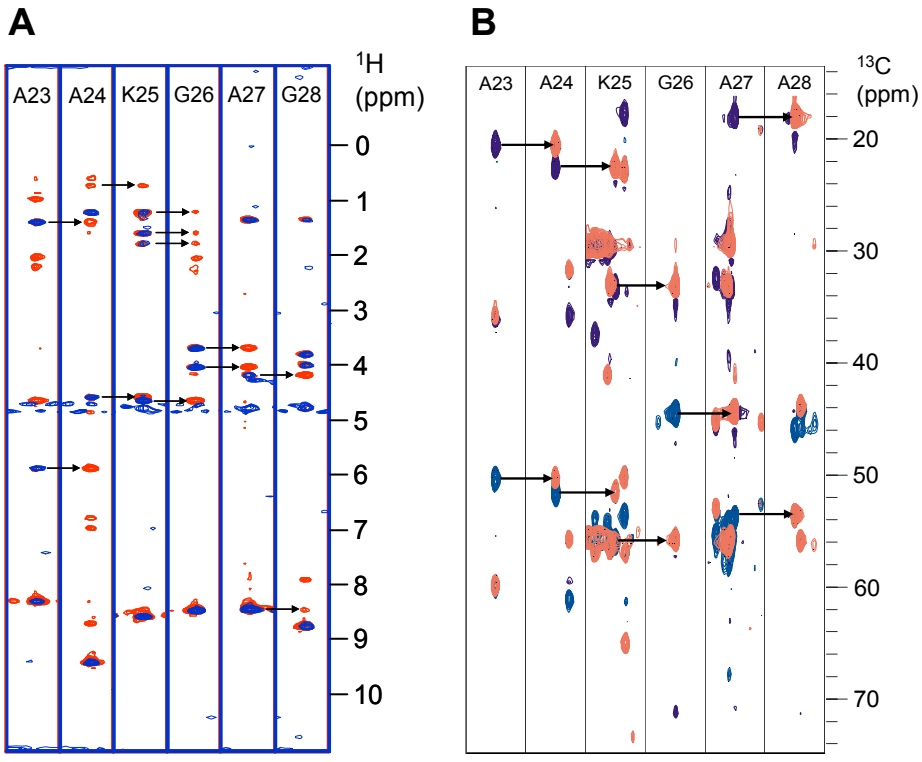


Figure 2.30. NMR: backbone assignments of AF6_PDZ. Sequential assignments of the segment corresponding to the end of the βB strand and the beginning of the βB/βC loop (sequence: ^{23}AAKGAG^{28}). (**A**) overlay of the HSQC-TOCSY (blue) and HSQC-NOESY (red) strips extracted from the corresponding 3D spectra and (**B**) overlay of the CACBNH (blue) and CBCA(CO)NH (red) strips extracted from the corresponding 3D triple resonance spectra.

Methyl group resonances were analyzed in the ^{1}H-^{13}C HSQC spectrum of the ^{13}C,^{15}N-labeled AF6_PDZ. As the amino acid sequence of AF6_PDZ contains 11 alanines, 5 threonines, 3 methionines, 11 valines, 10 leucines and 4 isoleucines, a total of 69 methyl resonances are expected in the ^{1}H-^{13}C HSQC spectrum. Cross-peaks of methyl groups in the ^{1}H-^{13}C HSQC spectrum cover the region of ~0.6-2.2 ppm in the ^{1}H dimension and ~9-28 ppm in ^{13}C dimension. In the methyl region of the AF6_PDZ's ^{1}H-^{13}C HSQC spectrum, 65 cross-peaks can be distinguished. Methyl resonances were assigned using the 3D HCCH-TOCSY spectrum. HCCH-TOCSY strips were resolved in the ^{1}H dimension for Cα and Cβ resonances of all 44 residues bearing methyl groups. For alanine residues, which have a methyl group at Cβ, the assignment was straightforward (**Figure 2.31**).

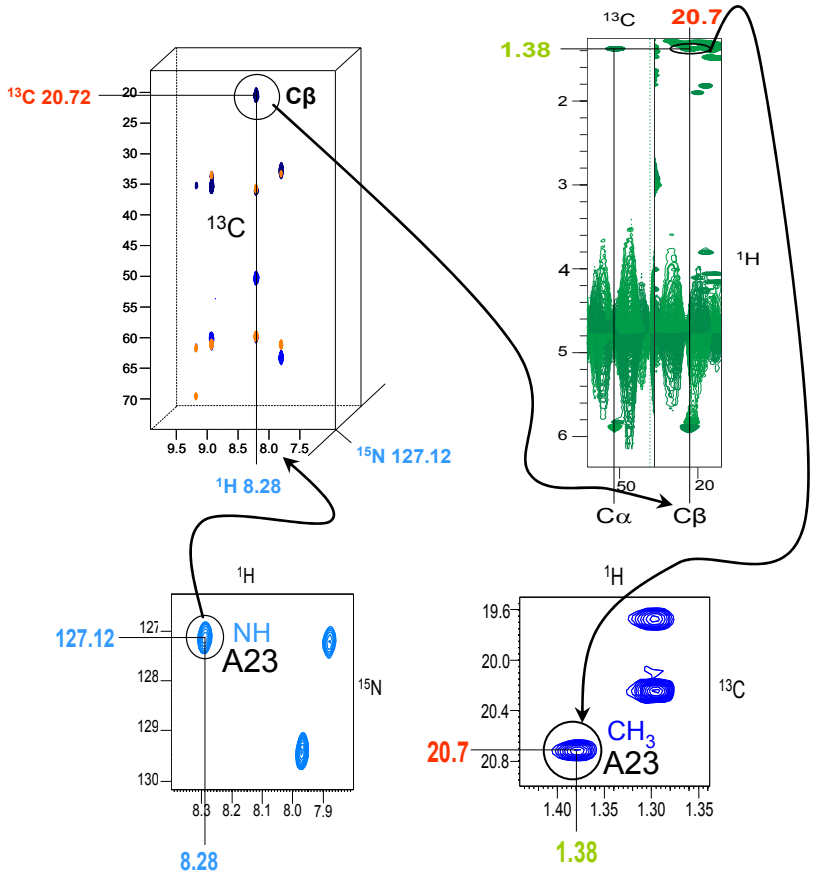


Figure 2.31. NMR: assignments of methyl groups. The residue specific assignment of A23 methyl group is shown as an example. Starting from the A23 backbone NH peak in the ^{1}H-^{15}N HSQC spectrum, the associated Cβ peak is found in the CACBNH spectrum, and correlated with the alanine CH_3 resonance in the HCCH-TOCSY spectrum. The assigned CH_3 resonance is shown in the ^{1}H-^{13}C HSQC spectrum.

For all other residues it was necessary to find strips for γ and δ carbons as well. Finally, 62 of 65 identified cross-peaks were assigned. The ^{1}H-^{13}C HSQC of the "peptide-bound" AF6_PDZ showed 68 cross-peaks, suggesting that some peak overlap is occurring in the spectrum of the "peptide-free" AF6_PDZ. Monitoring the PDZ peak chemical shift changes in the titration with the Jagged-1 peptides, three couples of completely overlapped cross-peaks were identified and assigned as methyls of V47/V66 ($^{1}H/^{13}C$: ~0.92/21.08 ppm), V9/V40 ($^{1}H/^{13}C$: ~0.83/21.02) and V9/V83 ($^{1}H/^{13}C$: ~0.72/21.57). Three peaks resonating at $\delta^{1}H > 2$ ppm could be assigned to the methione methyl groups. However, they could not be sequentially assigned as do not show any correlation with other carbons in the HCCH-TOCSY spectrum. The methyl of T85 was not assigned probably because overlap with other peaks (**Figure 2.32**), (**Table S2**), (**Figure S11**). Eventually, 65 methyl cross-peaks were assigned.

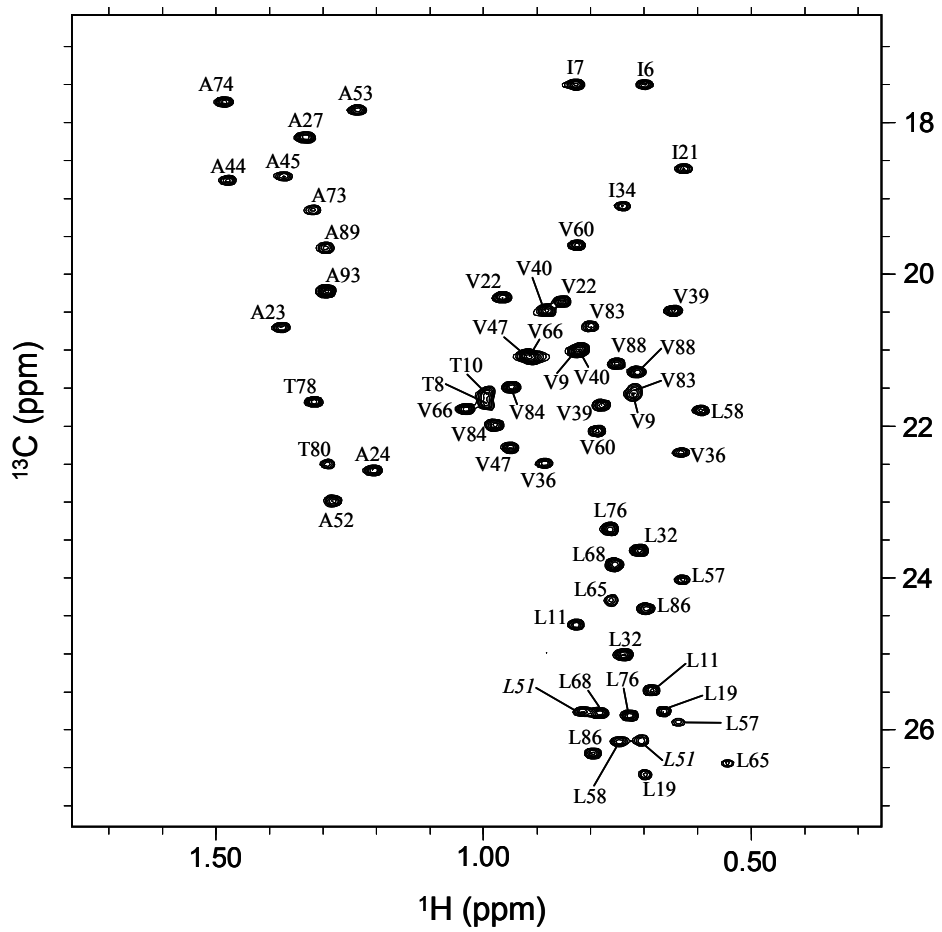


Figure 2.32. NMR of AF6_PDZ. ^{1}H-^{13}C HSQC spectrum of ^{13}C,^{15}N labeled AF6_PDZ (methyl region) and residue specific assignments of CH_3 cross peaks. The region of the $C^{\delta 1}(H^{\delta 1})_3$ isoleucine cross-peaks (9-16 ppm in the ^{13}C dimension) is not shown. The complete ^{1}H-^{13}C HSQC spectrum is shown in the **Figure S12**.

2.2.3 Jagged-1 C-terminus binds into the βB/αB groove of AF6_PDZ

The mode of the interaction of Jagged-1 intracellular region with the PDZ domain of afadin (AF6_PDZ) was investigated by heteronuclear NMR spectroscopy (***"Biophysical methods"***). The binding site on AF6_PDZ was initially mapped titrating the ^{15}N-labeled AF6_PDZ with J1C6, a short peptide comprising the C-terminal PDZ binding motif of Jagged-1 (RMEYIV-COOH) and recording ^{1}H-^{15}N HSQC spectra at each titration point. In these spectra, only NH peaks of AF6_PDZ are observable, since J1C6 is not isotopically labeled.

The NH cross-peaks of the protein backbone and of glutamine and asparagine side chains are sensitive local probes of ligand binding. With the NH cross-peak assignments it is possible to determine residues of AF6_PDZ involved in binding with J1C6 measuring the chemical shift perturbation (CSP) of peaks during the titration. For each backbone NH peak, CSP was calculated from the ^{1}H and ^{15}N chemical shifts differences between the PDZ in the free and bound form ($\Delta\delta_H$ and $\Delta\delta_N$, respectively) according to the formula [140] CSP = sqrt $[(\Delta\delta_H)^2 + \alpha_N(\Delta\delta_N)^2]$ using a $\Delta\delta_N$ scaling factor (α_N) of 0.102. As the chemical shift ranges (in ppm) in the ^{1}H and ^{15}N dimensions are very different (~3 ppm and 30 ppm respectively), $\Delta\delta_N$ values were scaled down using conversion factor corresponding to the gyromagnetic ratio of the two nuclei: $\alpha_N = |\gamma^{15}N| / |\gamma^{1}H| = 4.3156/42.576$, [141].

The titration of AF6_PDZ started with equimolar peptide:PDZ ratio and finished when no chemical shift change was further observed (peptide:PDZ 10:1). Th e majority of the AF6_PDZ NH peaks either do not change their chemical shift or display a progressive chemical shift change upon addition of J1C6, which is consistent with a fast exchange regime on the NMR time scale (**Figure 2.33**). However, the NH cross peaks of G18, L19 and A45 disappeared at low peptide:PDZ ratios to reappear towards saturation (G18 at peptide:PDZ 6:1, L19 and A45 at peptide:PDZ 4:1). Moreover, the NH peak of M17 vanished at the beginning of the titration and could not be tracked again. These results suggest an intermediate exchange time scale for these NHs. In the fast exchange regime, a single NMR peak is observed at a chemical shift that is population-averaged between the resonance position for the "free" and "bound" states. Upon ligand addition, the peak shifts towards a position of the "bound" state and its position reflects the free/bound ratio. In the slow exchange rate, the peak of the "bound" state grows while the original peak of

the “free” state, disappears. The exchange rate is said to be intermediate if it cannot be approximated by the above two cases. Monitoring chemical shift changes is generally straightforward for peaks in fast exchange, but may become problematic for peaks that are in intermediate exchange. Amide NHs of L19 and A45 are examples of such peaks. They are originally very close to each other, have a remarkable shift during the titration and do not appear in “halfway” positions. Other examples are peaks which overlap either in the free form, “halfway” or in the bound form. All these ambiguities were solved by the backbone assignment of the ^{13}C,^{15}N-labeled AF6_PDZ bound to J1C24pS peptide (described later) by using CBCANH and CBCA(CO)NH spectra.

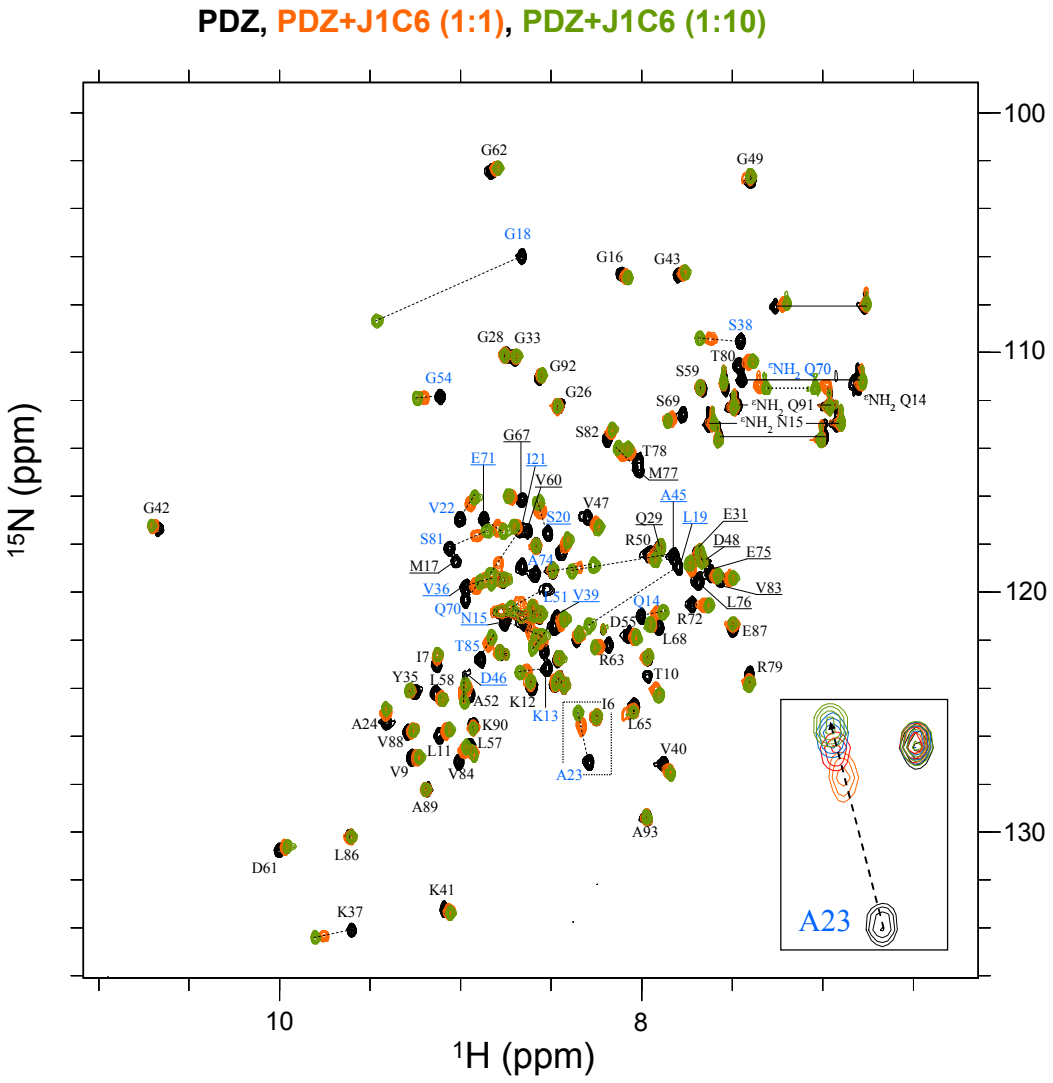


Figure 2.33. Chemical shift mapping. ^{1}H-^{15}N HSQC spectrum of AF6_PDZ (black) superimposed on the HSQC spectra obtained in the presence of J1C6 at protein:peptide molar ratios of 1:1 (orange), and 1:10 (green). Residues displaying a significant CSP are labeled in blue. Inset: amide cross peak of A23 at different titration points: 1:1 (orange), 1:2 (red), 1:4 (blue), and 1:10 (green).

Once calculated for all backbone NH peaks, CSPs were plotted against the amino acid sequence of AF6_PDZ (**Figure 2.34**). Most of the residues displaying significant CSP values are located on one side of the PDZ (βB, βC, αA, αB) whereas the other side (βA, βD, βE, βF) is hardly affected (**Figure 2.35**). From the three residues that experienced the largest change in chemical shift (G18, L19 and A45), the first two constitute the GMGL motif ($G_{16}M_{17}G_{18}L_{19}$) which is responsible for the binding of the ligand carboxylate group. The intermediate exchange rate of its NH peak suggests that the residue M17 could also be involved in the binding to J1C6. The first glycine of the motif (G16), however, did not show a significant chemical shift perturbation. In the βB strand (between L19 and A23) all residues are significantly perturbed and show an alternating CSP pattern which suggests the presence of hydrogen bonds between the backbone of the peptide and the βB strand (**Figure S12**).

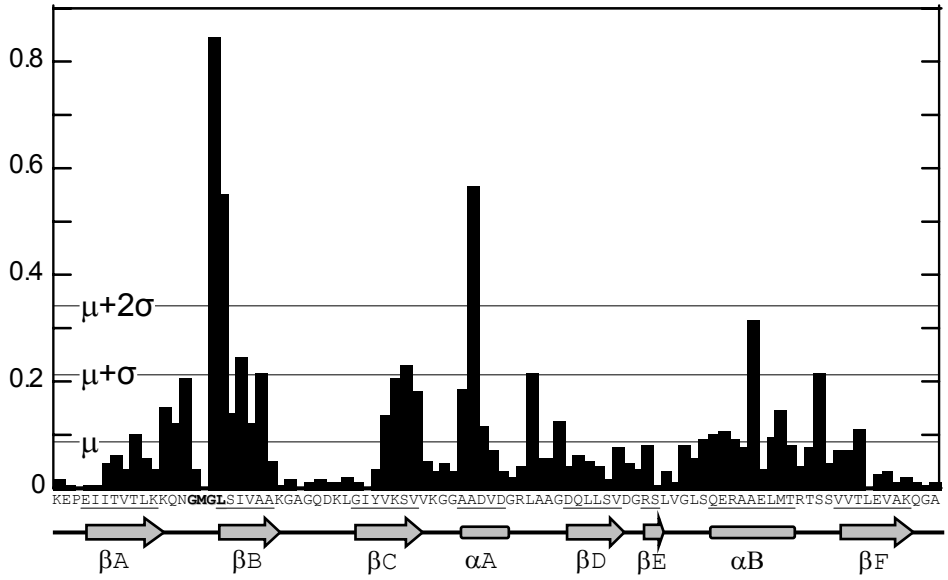


Figure 2.34. Chemical shift mapping. Combined chemical shift perturbation (CSP) values of AF6_PDZ backbone amides obtained from the titration with J1C6. For each backbone NH peak, CSP was calculated from the ^{1}H and ^{15}N chemical shift differences between the PDZ in the free and bound form at the saturation point. CSP values are plotted against amino acid sequence. CSP mean values (μ), CSP standard deviations (σ) and secondary structure elements are also shown.

The βC strand, which is antiparallel to the βB strand, also contains residues with significant CSP values (V36, K37, S38 and V39) while in the other four β-strands (βA, βD, βE and βF) residues remain unaffected. On the other hand, both α-

helices (αA and αB) are influenced by the interaction with J1C6. In the αA helix the perturbation can be mapped to the first three residues (A44, A45 and D46), in particular A45. In the αB helix four residues (Q70, E71, A74 and M77) are significantly perturbed, three of which are oriented towards the βB strand (Q70, A74 and M77). Outside regular secondary structure elements, three residues (N15, L51 and S81) show significant CSP values.

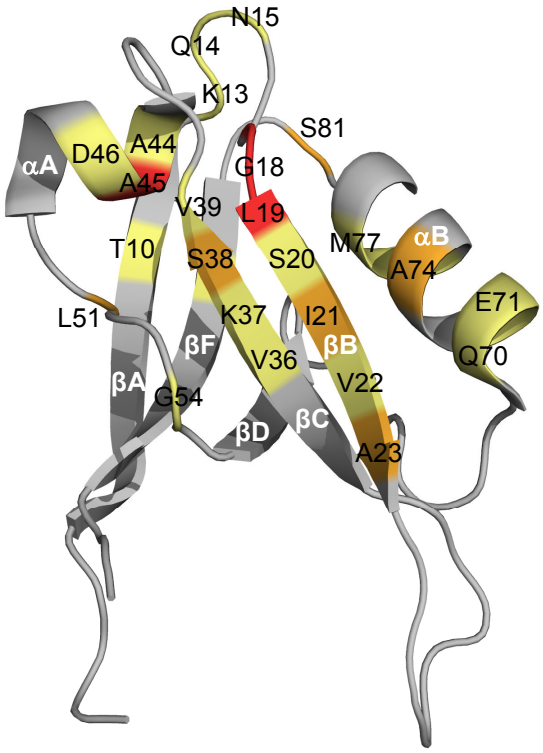


Figure 2.35. Chemical shift mapping. Cartoon representation of the AF6_PDZ structure (PDB: 2AIN) colored according to backbone amide CSP values (μ<CSP<μ+σ, yellow; μ+σ<CSP<μ+2σ; orange; CSP>μ+2σ, red) obtained from the titration with J1C6.

Overall, the CSP pattern observed is very similar to that reported for AF6_PDZ binding to the C-terminal peptide of neurexin (KKNKDKEYYV-COOH), which bears identical residues at P-0, P-2, and P-3 [75] (**Figure S13**). These CSP similarities are most evident in the perturbation of the GMGL motif, including the disappearance of M17 peak. The CSP profile in the βB strand (L19-A23) is also comparable, except the S20 which is remarkably more perturbed in the titration with neurexin peptide. Moreover, similarity between two CSP patterns is also observable in

the αB helix region (from Q70 to M77). A74 in particular was proposed to make hydrophobic contacts with the aromatic ring of the tyrosine at P-2 of neurexin, and J1C6 bears the same amino acid type in that position.

Q70, which is located at the entrance of the peptide binding groove and points its side chain towards the βB strand, is also significantly perturbed by J1C6. The glutamine in this position is unique and is not found in PDZ domains of Class II. As already specified in the introduction, the residue in this position (the first residue in the αB helix; αB1) should interact with the peptide's P-2 residue and should give the largest contribution to the PDZ-peptide binding specificity. Additionally, like in the interaction with neurexin, significant changes in the chemical shift of Q70 $^{\varepsilon}NH_2$ side chain were observed upon J1C6 binding [75] (**Figure 2.36**).

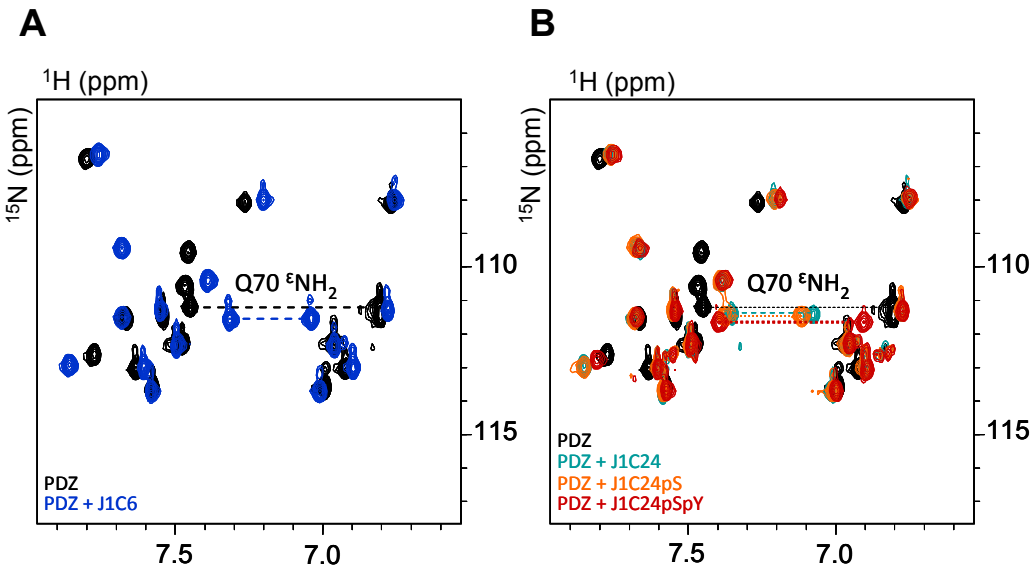


Figure 2.36. Chemical shift mapping: effects of peptide binding on Q70 $^{\varepsilon}NH_2$ side chain. ^{1}H-^{15}N HSQC spectrum of AF6_PDZ (black) superimposed on the HSQC spectra obtained with an excess of (**A**) J1C6 (blue), and **(B)** J1C24 (light blue), J1C24pS (orange) or J1C24pSpY (red). Only the region of Asn and Gln side chains is shown.

In the βC strand, J1C6 and the neurexin peptide overall perturb the same residues. In both cases, the most affected residue is S38, but in the binding with neurexin the perturbation is much more pronounced. Unlike J1C6, the neurexin peptide has another tyrosine residue in P-1 position, which most likely interacts with S38. The glutamic acid at P-3, present in both peptides, is proposed to form an ion-

pair with the side chain of K37 [75]. K37, however, is more perturbed by J1C6. Residues in the αA helix are not directly involved in peptide binding, and it has been suggested that the perturbation of their chemical shifts might arise from a cascade effect [142], [143]. The same could be also for N15, L51 and S81, residues which experienced a significant chemical shift perturbation upon addition of J1C6, but are distal from the peptide binding site.

Taken all together, the above similarities with the binding mode of the neurexin C-terminal peptide suggest that J1C6 binds specifically to AF6_PDZ. The peptide binding site can be located between the βB strand and the αB helix, and CSP values of AF6_PDZ in the presence of J1C6 are compatible with a canonical binding mode, with the carboxylate of V at P-0 pointing towards the loop between βA and βB and establishing hydrogen bonds with the NH of G18 and L19. Hydrogen bonds between the backbone of the peptide and the βB strand are consistent with the extended conformation of the bound peptide observed in PDZ/peptide complexes (**Figure S12**).

To estimate not only the binding mode but also the affinity, K_d values for selected residues in βB, βC, αA, and αB were calculated from the backbone NH CSP to yield a mean K_d of 17.2±2.3 μM (**Table 2**), (**Figure S14**). Because the chemical shift of backbone NHs may be influenced not only by ligand binding but also by subtle changes in secondary structure, hydrogen bonding, and exchange with the solvent, the titration with J1C6 was carried out also on the ^{13}C,^{15}N-labeled AF6_PDZ. ^{1}H-^{13}C HSQC spectra were recorded at each titration point and the CSP of side chain methyl groups (CH_3) were calculated (**Figure 2.37**). Calculation of K_d values for selected residues lead to a mean value of 16.4±4.0 μM in excellent agreement with the K_d calculated from backbone NH CSP values (**Table 3**), (**Figure S14**). The NH peak perturbations of AF6_PDZ residues are also in good agreement with the CH_3 peak perturbations of the same residues. The largest methyl-CSPs were measured for A74, L19 and V22, residues located in the peptide binding groove, which showed also a significant NH-CSP. Significant CSP values were displayed also by other residues in the βB/αB groove (I21, A23 and T78), in αA helix (A44 and A45) and in the βC strand (I34 and V36). It is worth mentioning that the chemical shift change of the methyl peak of A45 is not proportional to that of its amide peak. As A45 is not in direct contact with the peptide, the difference between its NH-CSP and methyl-CSP

could be explained by the higher sensitivity of NH chemical shifts to small conformational changes which occurs it the αA helix. On the other hand, a significant methyl-CSP was showed by V40, a residue with a negligible NH-CSP. This residue is located in the βC/αA loop with the side chain oriented towards the GMGL motif. Some residues show different peak perturbations between their methyl groups (L19, V22, V36 and V40) suggesting a high sensitivity of the CH_3 chemical shift response to binding (**Figure 2.37**). Moreover, K_ds measured for two methyl groups of L19 are significantly different (**Table 3**).

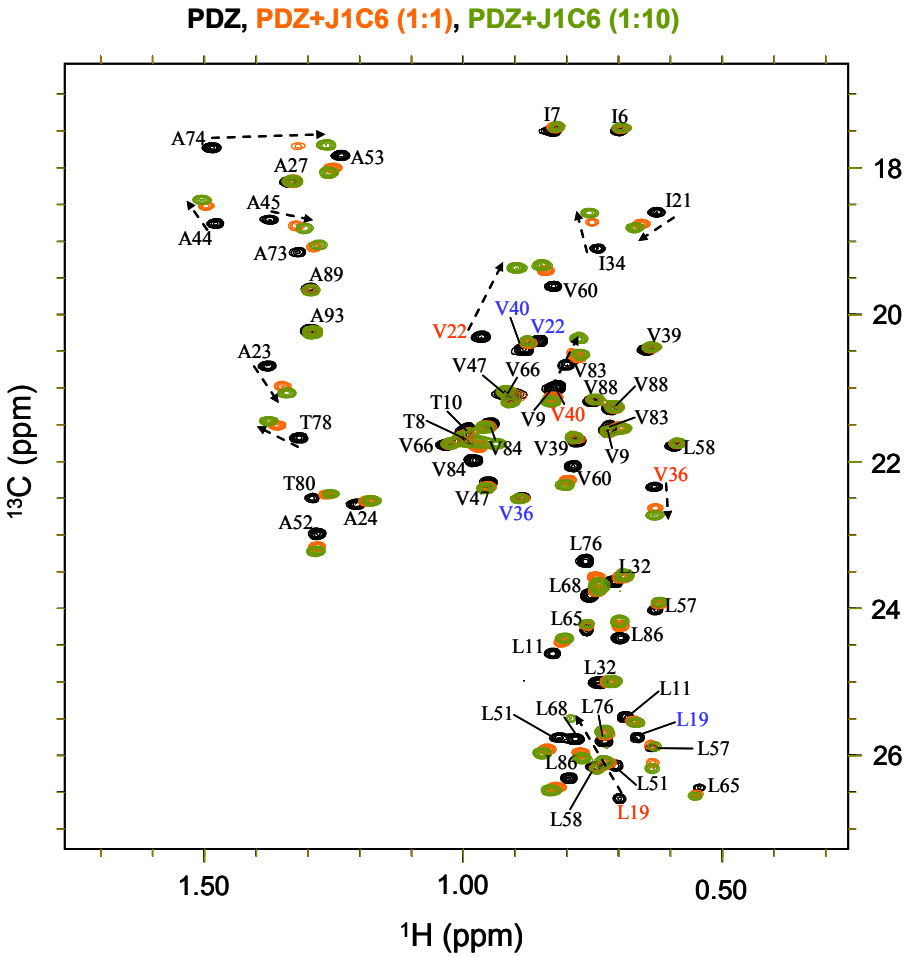


Figure 2.37. Chemical shift mapping. ^{1}H-^{13}C HSQC spectrum of ^{13}C,^{15}N labeled AF6_PDZ (black) superimposed on the HSQC spectra obtained in the presence of J1C6 at protein:peptide molar ratios of 1:1 (orange) and 1:10 (green). Methyl groups (L19, V22, V36 and V40) displaying larger CSP values are labeled in red, those displaying smaller CSP values are labeled in blue.

2.2.4 Jagged-1's binding is strictly local

In general, binding to PDZ domains is largely determined by the last four C-terminal residues of the target protein [144]. In the NMR-derived structure of the complex between the afadin PDZ domain and a peptide corresponding to the C-terminus of Bcr, however, several NOEs between the αB helix and the side chain of the residue at P-4 were identified [75]. Furthermore, deletion studies showed that at least six terminal residues of EphB2 and EphB3 are required to bind the PDZ domain of afadin [56]. Finally, binding of afadin PDZ domain to a combinatorial peptide library suggests that even residues in upstream positions might play an additional role [144]. To determine if the interaction with AF6_PDZ involves also residues in the upper C-terminal region of Jagged-1, the binding site on AF6_PDZ was also mapped titrating the ^{15}N-labeled AF6_PDZ with the J1C24 peptide and the J1_tmic recombinant protein. J1C24 corresponds to the last 24 C-terminal residues and J1tmic corresponds to the whole cytoplasmic tail of Jagged-1. NMR measurements showed no major changes in the backbone NH CSP patterns whether AF6_PDZ was titrated with J1C6, J1C24, or J1_tmic (**Figure 2.38**), (**Figure 2.39**). This shows that binding of Jagged-1 cytoplasmic tail to AF6_PDZ is strictly local, involving only the C-terminal residues of Jagged-1 and the binding groove formed by βB and αB in the PDZ.

The CSP differences profile of J1C6 and J1_tmic are not significantly different, except small changes for the G54 residue. Unlike titrations with J1C6 and J1_tmic, in the titration with J1C24 the G18 cross peak does not reappear. In principle, this may be due to exchange rate differences between J1C6 and J1C24. A similar behavior would be then expected for the titration with J1_tmic. This is not the case, which suggests that for J1C24, probably due to the slightly minor solubility, the titration did not reach the full saturation point. An estimation of the K_d from backbone NH CSP of selected residues lead to mean values of 17.2±2.3, 8.9±3.7, and 25.9±8.9 μM for J1C6, J1C24, and J1_tmic, respectively (**Table 2**).

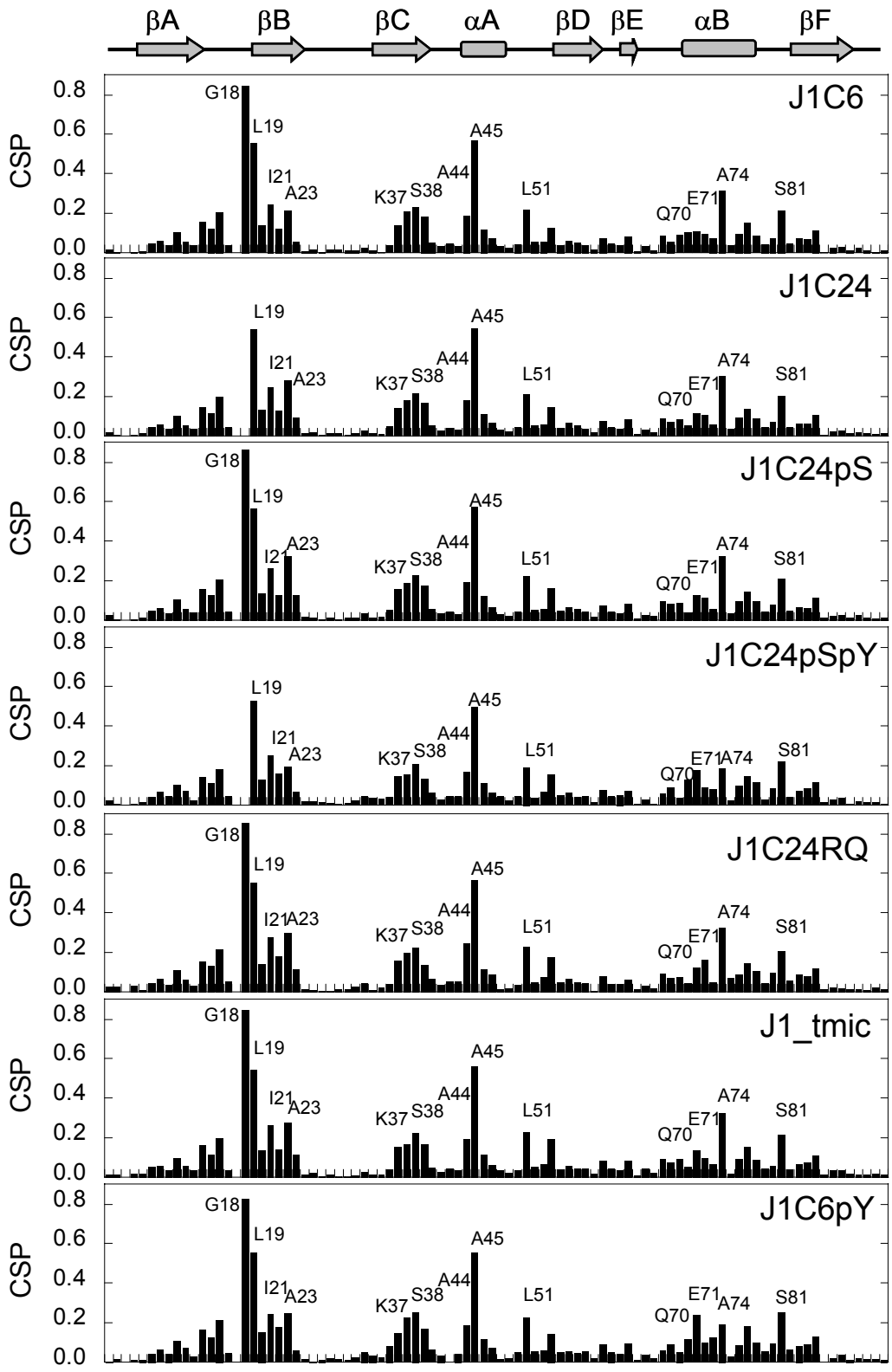


Figure 2.38. Chemical shift mapping. Combined chemical shift perturbation (CSP) values of AF6_PDZ backbone amides obtained from titration with different peptides (J1C6, J1C24, J1C24pS, J1C24pSpY, J1C24RQ, J1C6pY) and with the recombinant J1_tmic; secondary structure elements are also shown.

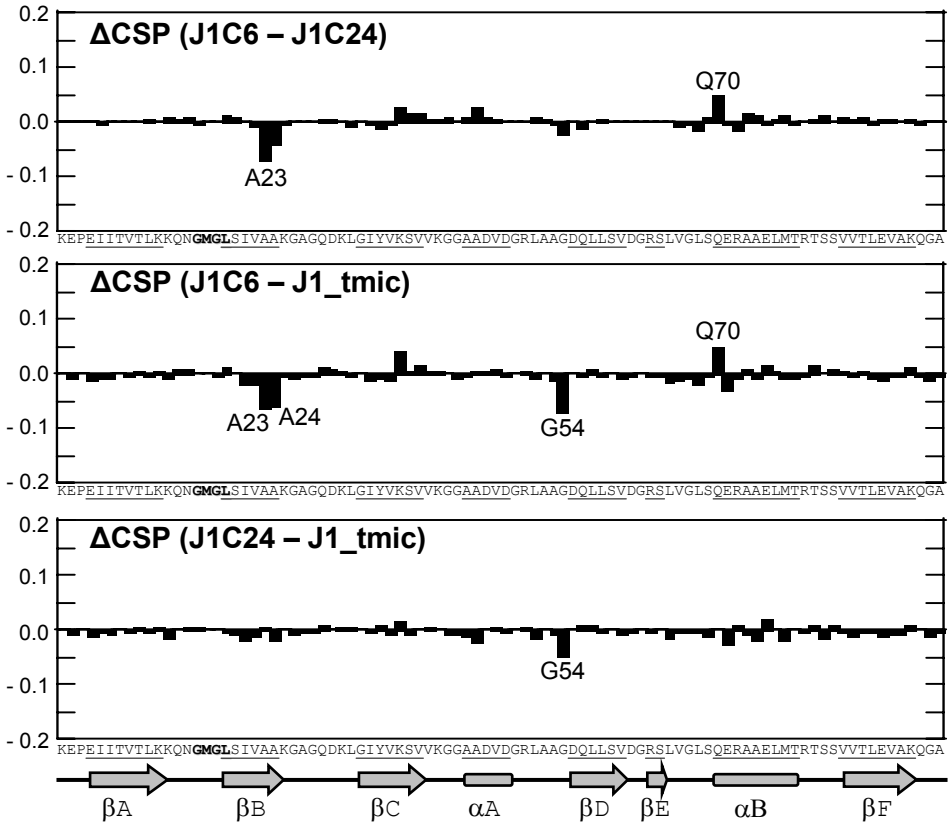


Figure 2.39. Chemical shift mapping. Histogram plot of CSP differences (ΔCSP) for AF6_PDZ backbone amides, obtained comparing different ligand pairs. ΔCSP values are plotted against the amino acid sequence and residues displaying higher values are labeled.

		J1C6	J1C24	J1C24pS	J1C24pSpY	J1C24RQ	J1_ic
G18		17.1	–	9.9	–	–	–
L19	βB	20.1	(15.7)	3.8	–	(1.5)	–
I21	βB	–	7.2	12.6	22.5	–	(40.6)
A23	βB	16.5	9.5	5.9	17.4	(1.0)	–
K37	βC	16.6	6.0	8.3	15.8	(1.6)	21.1
S38	βC	16.1	7.1	8.9	24.0	–	(39.1)
A45	αA	19.1	5.2	10.3	(20.5)	–	(7.7)
L51		15.8	3.2	7.0	18.5	–	13.4
Q70	αB	11.8	(12.8)	4.6	(49.8)	–	(23.8)
A74	αB	19.7	12.2	10.3	(19.2)	(1.1)	21.2
M77	αB	18.1	7.6	12.2	25.0	–	32.3
S81		18.7	14.6	12.1	40.0	–	25.9
mean		17.2±2.3	8.9±3.7	8.7±3.1	25.8±11.5	–	25.9±8.9

Table 2. Backbone chemical shift mapping. Residue specific Kd values (μM) calculated from combined CSP values of backbone amides; a mean value and a standard deviation for selected residues are also reported; values with high fitting errors are reported in parentheses and were not included in the calculation of the mean.

2.2.5 J1_tmic does not fold upon binding

The large majority of structural studies carried out so far on PDZ domains focused on the details of the interaction between the PDZ domain and short peptides. The fact that peptides bind in an extended β-strand conformation that complements the PDZ βB-strand [145] raises the question of whether the binding with AF6_PDZ could trigger a conformational change in the entire cytoplasmic tail of Jagged-1 [146]. To address this question, NMR spectra of the ^{15}N-labeled J1_tmic alone and in the presence of the unlabeled AF6_PDZ were recorded. The ^{1}H-^{15}N HSQC spectrum of J1_tmic alone does not show any significant ^{1}H chemical shift dispersion in the NH region, as expected for an intrinsically disordered protein (**Figure 2.40**). No significant chemical shift dispersion was induced by addition of AF6_PDZ, and the chemical shift of the majority of the NH peaks remained unaffected, with the notable exception of a few peaks that can tentatively be assigned to the C-terminal, PDZ binding region of J1_tmic. These results show that, with the conditions used, binding of J1_tmic to the PDZ domain of afadin does not trigger folding and confirm that this interaction is restricted to the C-terminal residues of Jagged-1 cytoplasmic tail.

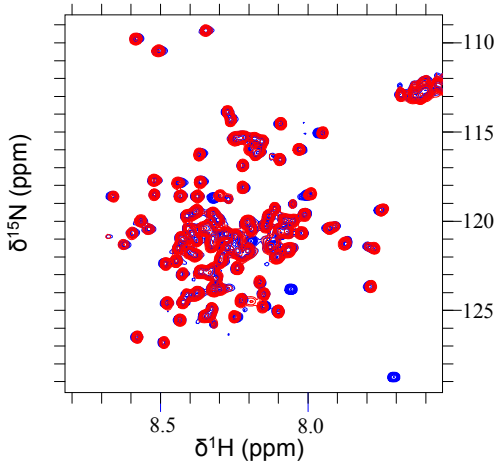


Figure 2.40. Binding of Jagged-1 cytoplasmic tail to AF6_PDZ. ^{1}H-^{15}N HSQC spectrum of the ^{15}N-labeled J1_ic (blue) free and (red) in the presence of unlabeled AF6_PDZ.

2.2.6 Tyrosine phosphorylation at P-2 reduces affinity

Whereas the interaction between PDZ domains and short peptides has been studied extensively, relatively little is known about the mechanisms that control this interaction. Post-translational modifications like S/T phosphorylation [147] and, more

recently, lysine acetylation [148], [149] were shown to play a role in the modulation of binding. In the C-terminal region of Jagged-1, several residues (T1197, S1207, S1210, Y1216) are potential targets for protein kinases. Although no experimentally confirmed phosphorylation site has been reported to date for Jagged-1, different prediction methods point to T1197, S1210, and Y1216 as potential phosphorylation sites. S1210 matches the [SX-X-S/T] CK1 phosphorylation pattern and is predicted by GPS [150] as a potential phosphorylation site also for CAMK. Y1216 matches the [E/D]pY[I/L/V] EGFR pattern and is predicted by GPS as a potential phosphorylation site for several receptor tyrosine kinases. As T1197 is far upstream of the PDZ-binding motif, the experiments focused on S1210 and Y1216, which are at P-8 and P-2, respectively, in the binding motif.

To investigate the possible effects of phosphorylation on the binding mode and affinities of J1C24, two phosphorylated variants of the J1C24 peptide, J1C24pS (Ac-NWTNKQDNRDLESAQpSLNRMEYIV-COOH) and J1C24pSpY (Ac-NWTNKQDNRDLESAQpSLNRMEpYIVCOOH), were prepared to study their interaction with AF6_PDZ by NMR and SPR. Titration of ^{15}N-labeled AF6_PDZ with J1C24pS resulted in a CSP profile very similar to that obtained for J1C24, the only major difference being the reappearance of the resonance corresponding to the G18 amide (**Figure 2.38**),(**Figure 2.41**). It is possible that the negatively charged phosphate group at P-8 slightly increases the peptide solubility, which could affect the saturation point of the titration (paragraph 2.2.4.). Eventually, serine phosphorylation at P-8 does not change significantly the mode of binding to AF6_PDZ.

Titration of AF6_PDZ with the doubly phosphorylated J1C24pSpY peptide, on the other hand, showed appreciable differences compared with the J1C24 titration. Residues A23 (βB strand) and A74 (αB helix) are significantly less perturbed, while Q70 and E71 (both in αB helix) show larger CSP values (**Figure 2.42**). These differences are more evident when comparing the CSP histograms of J1C24pSpY and J1C24pS. The effect of the phosphorylatyed P-2 tyrosine can be also seen monitoring CSP of the Q70 $^{\varepsilon}NH_2$ peaks (**Figure 2.36**). These results suggest that J1C24pSpY enters in the βB/αB groove in a slightly different mode in which the phosphorylated tyrosine interacts with Q70 whether with A74. Additionally, results from our preliminary studies also show binding differences between short peptides, J1C6 and J1C6pY (phosphorylation at P-2), confirming that tyrosine phosphorylation at P-2 must slightly change the mode of binding. As above, the A74 NH peak is less

perturbed, but in the case of J1C6pY the CSP increase of E71 is more evident while, on the other hand, A23 and Q70 show no chemical shift changes (comparing with the J1C6 binding) (**Figure 2.41**).

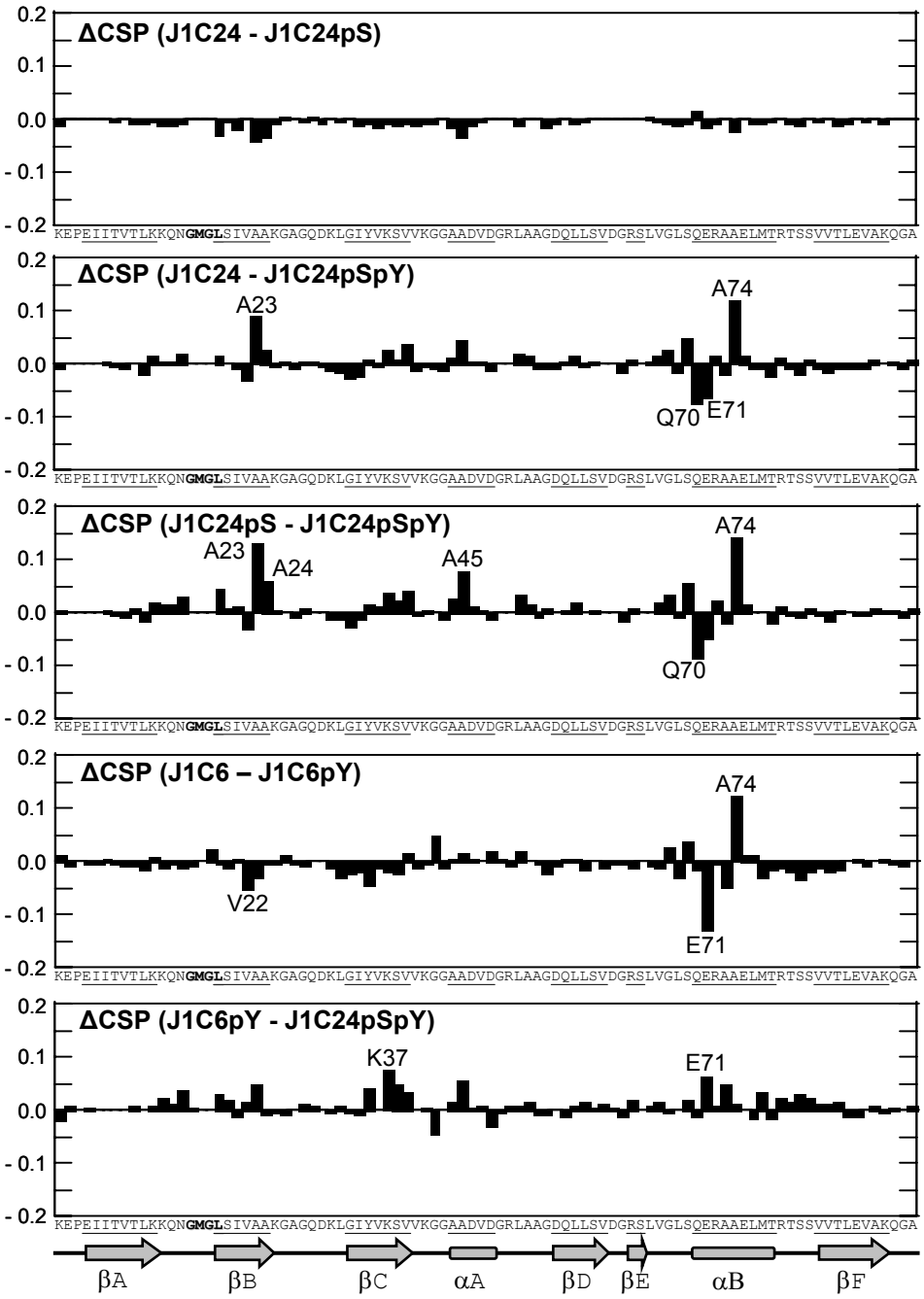


Figure 2.41. Chemical shift mapping. Histogram plots of CSP differences (ΔCSP) for AF6_PDZ backbone amides, obtained comparing different phosphorylated and not phosphorylated ligand pairs. ΔCSP values are plotted against amino acid sequence and residues displaying higher values are labeled.

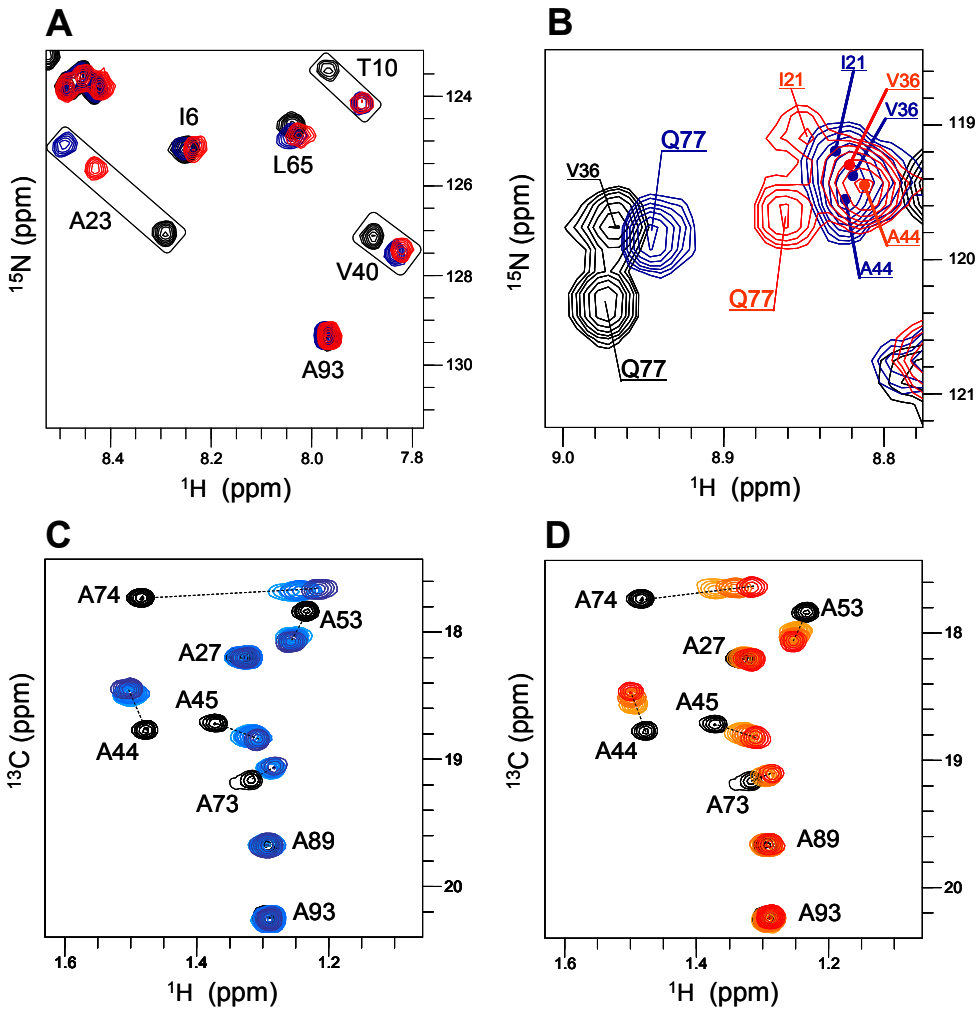


Figure 2.42. Titration of AF6_PDZ with J1C24 and J1C24pSpY. (**A**) and (**B**): backbone NH cross peaks in different regions of the ^{1}H-^{15}N HSQC spectrum of the ^{13}C,^{15}N-labeled AF6_PDZ, free (black), and in the presence of J1C24 (blue) or J1C24pSpY (red). Resonances of the peptide bound AF6_PDZ (red and blue) are at the saturation point. (**C**): methyl cross peaks of alanines in the ^{1}H-^{13}C HSQC spectrum of the ^{13}C,^{15}N-labeled free AF6_PDZ (black), and in the presence of increasing amounts of J1C24 (very light blue, light blue and blue). (**D**): methyl cross peaks of alanines in the ^{1}H-^{13}C HSQC spectrum of the ^{13}C,^{15}N-lebeled free AF6_PDZ (black), and in the presence of increasing amounts J1C24pSpY (light orange, orange and red). Titration points in figures (**D**) and (**E**) are equivalent.

Residue-specific K_d values estimated from CSP values of backbone amides (**Table 2**) show that, whereas no significant difference can be detected between J1C24 (mean K_d = 8.9±3.7 μM) and J1C24pS (mean K_d = 8.7±3.1 μM), tyrosine phosphorylation at P-2 decreases affinity for AF6_PDZ (mean K_d = 25.8±11.5 μM). Also in this case, the K_d values estimated from backbone NH CSP and from side chain methyl groups CSP (10.0±2.1, 10.2±3.6, and 30.9±13.1 μM for J1C24,

J1C24pS, and J1C24pSpY, respectively) (**Table 3**) are in very good agreement and display the same trend. SPR measurements support the NMR results (**Figure 2.43**). The K_d values measured for the J1C24 and J1C24pS (37±3.8 μM and 28±5.7 μM, respectively) are only slightly different, given the experimental error, whereas tyrosine phosphorylation at P-2 markedly decreases the affinity (K_d = 150±69 μM). Although there is a discrepancy between the K_d values measured by NMR and by SPR, the relative affinities are conserved (J1C24 ~ J1C24pS > J1C24pSpY).

		J1C6	**J1C24**	**J1C24pS**	**J1C24pSpY**
L19	βB	6.7/19.8	–	16.0	7.2/20.3
I21	βB	16.4/14.5	–	7.4	18.2 /(45.12)
V22	βB	18.4	6.6	7.0	28.4
A23	βB	15.5	10.2	8.7	33.3
A24	βB	11.1	14.2	9.5	(60.3)
V36	βC	21.8	10.8	16.2	22.4
A44	αA	18.5	10.2	5.4	26.3
A45	αA	15.8	8.0	10.8	36.1
L51		19.6	(62.2)	(47.5)	–
A53		–	8.1	–	25.3
V60	βD	16.0/22.1	9.7 /9.7	10.0	32.8/ 26.3
A73	αB	15.4	12.5	13.8	48.5
A74	αB	14.7	10.0	7.6	32.9
mean		16.4±4.0	10.0±2.1	10.2±3.6	30.9±13.1

Table 3. Side chain chemical shift mapping. Residue-specific Kd values (μM) calculated from combined CSP values of side chain methyl groups; a mean value and a standard deviation for selected residues are also reported; values with high fitting errors are reported in parentheses and were not included in the calculation of the mean.

The differences between NMR and SPR measurements may arise from the experimental set up specific to each technique, the different conditions used, and from fitting errors in the estimation of K_d values from CSP. Preliminary experiments were carried out by immobilizing AF6_PDZ on a CM5 sensor chip and injecting the different peptides. Reproducibility of measurements was however poor, and K_d values obtained with this experimental approach (400-800 μM) were more than one order of magnitude larger than K_ds measured from NMR spectra and significantly larger than

the average K_d values reported for the PDZ-peptide binding (low micromolar [151], [69]).

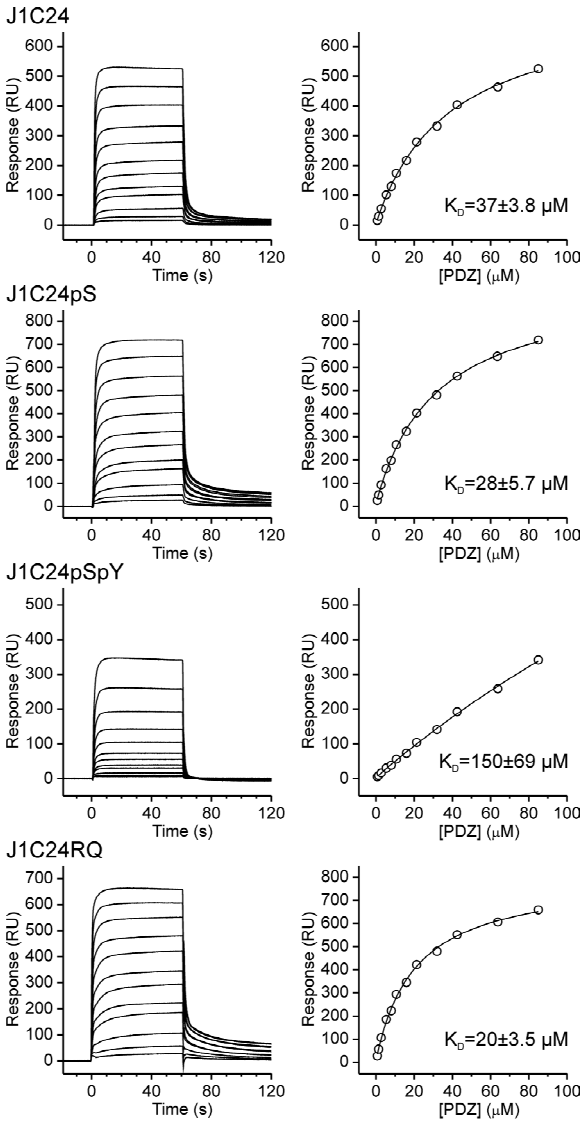


Figure 2.43. Surface Plasmon Resonance. Sensorgrams of AF6_PDZ (0.66-85 μM) binding to biotinylated peptides immobilized on the surface of avidin-CM5 sensor chip (left); plots of the response at steady state levels versus the concentration of AF6_PDZ and curve fitting (right).

A reciprocal experimental strategy was used to overcome this problem. Briefly, biotinylated peptides were immobilized on the streptavidin-coated sensor chip (SA sensor chip) through non-covalent interactions. This approach gave better results, but suffered from a significant degree of non-specific binding to the control flow cell without the peptide, possibly because of charge-charge interactions between

streptavidin (calculated pI = 6.1) and AF6_PDZ (calculated pI = 9.4). Eventually, better results were obtained employing avidin instead of streptavidin. The calculated pI = 9.7 of avidin is similar to that of AF6_PDZ. Indeed, the non specific binding of AF6_PDZ to the control flow cell was almost completely eliminated.

2.2.7 The R1213Q mutation at P-5 increases affinity

A G→A missense mutation leading to the R1213Q substitution in the intracellular region of Jagged-1 was identified as a sporadic mutation in a severe case of extrahepatic biliary atresia, a congenital obstruction of the bile ducts [28]. This mutation involves the residue at P-5 of the PDZ-binding motif. To verify whether this mutation can affect the mode of binding of Jagged-1 to afadin, the ^{15}N-labeled AF6_PDZ was titrated with J1C24RQ (Ac-NWTNKQDNRDLESAQSLNQMEYIV COOH), ^{1}H-^{15}N HSQC spectra recorded, and the binding site mapped. The CSP plot shows that J1C24RQ binds into the groove between βB and αB, in a similar way as the wild-type J1C24 peptide, but with slightly larger CSP values measured for residues V22, A44 and R72, which are located in the βB-strand, the αA-helix and the αB-helix, respectively (**Figure 2.38**), (**Figure 2.44**).

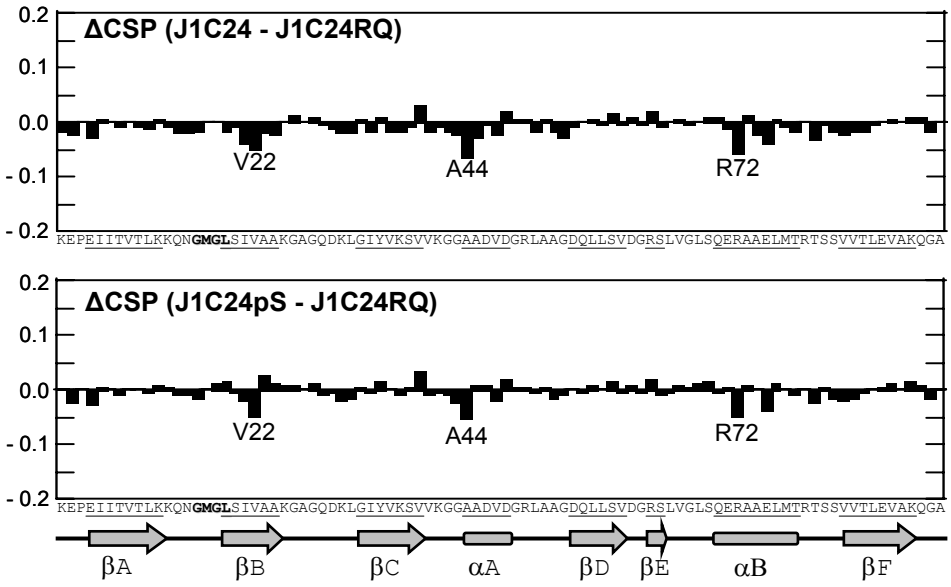


Figure 2.44. Chemical shift mapping: the effect of the R1213Q mutation. Histogram plot of CSP differences (ΔCSP) for AF6_PDZ backbone amides, obtained comparing different ligand pairs. ΔCSP values are plotted against the amino acid sequence and residues displaying higher values are labeled.

The R1213Q mutation removes the positive charge at the P-5 position, which could affect the interaction with the positively charged R72 of the αA-helix. Calculation of residue specific K_d values from CSP of backbone amides was difficult because of the limited solubility of J1C24RQ, but appear to be in the low micromolar range. Consistently, the K_d measured by SPR for J1C24RQ (20±3.5 μM) is smaller than that of the wild type J1C24 (K_d = 37±3.8 μM). This suggests that residues upstream of the canonical C-terminal tetrapeptide can subtly change the binding affinity. In this specific case, a significant but not dramatic change in the binding affinity induced by R1213Q mutation is associated with a severe morphological disorder.

2.2.8 Phosphorylation affects binding to lipid micelles

The interaction of J1_tmic with AF6_PDZ is local, folding-independent, and tuned by phosphorylation (paragraph 2.2.). It was also shown that the last six residues of J1_tmic are sufficient for binding to AF6_PDZ. Additionally, tyrosine phosphorylation at P-2, but not serine phosphorylation at P-8, reduces the binding affinity. Moreover, it was demonstrated that J1_tmic is disordered in solution, but partially folds upon binding to negatively charged liposomes and micelles (paragraph 2.1). If partial folding encompasses the C-terminal region of J1_tmic, it could make it not available for the PDZ binding. Additionally, phosphorylation in the C-terminal region of J1tmic could affect binding to the membrane. To address these hypotheses, far-UV CD and fluorescence experiments were carried out on J1C24 peptides (J1C24, J1C24RQ, J1C24pS, J1C24pSpY) in the presence of LMPG and LMPC micelles.

In buffer, the CD spectrum of J1C24 is very similar to that of the full-length J1_tmic. The strong minimum at 198 nm is typical of disordered proteins having little or no secondary structure (**Figure 2.45**).

CD spectra of J1_tmic in the presence of trifluoroethanol (TFE) showed that J1_tmic possesses intrinsic helical propensity and the measured α-helical content is consistent with the predicted one (**Figure 2.11**). Of the four helical stretches identified in the consensus secondary structure predictions, one is localized in the last 24 C-terminal residues (the J1C24 sequence) (**Figure 2.10**). Indeed, CD spectra showed a coil-helix transition in J1C24 upon addition of increasing amounts of TFE (**Figure 2.45 A**), although with a slightly lower intensity with respect to J1_tmic. In the

presence of LMPG, but not LMPC micelles, J1C24 shows a coil-helix transition (**Figure 2.46 A**). This secondary structure formation is strongly pH-dependent, with a sharp increase in the helical content when the pH is lowered from pH 7 to pH 6.

The above experiments were carried out also on the phosphorylated peptides. CD spectra show that J1C24pS and J1C24pSpY are mainly disordered in buffer solution (**Figure 2.45 B and C**). In the presence of trifluoroethanol (TFE) both peptides showed far less intrinsic helical propensity compared to J1C24. Furthermore, J1C24pS and J1C24pSpY remain disordered also in the presence of lipid micelles, LMPC and LMPG as well (**Figure 2.46 B and C**).

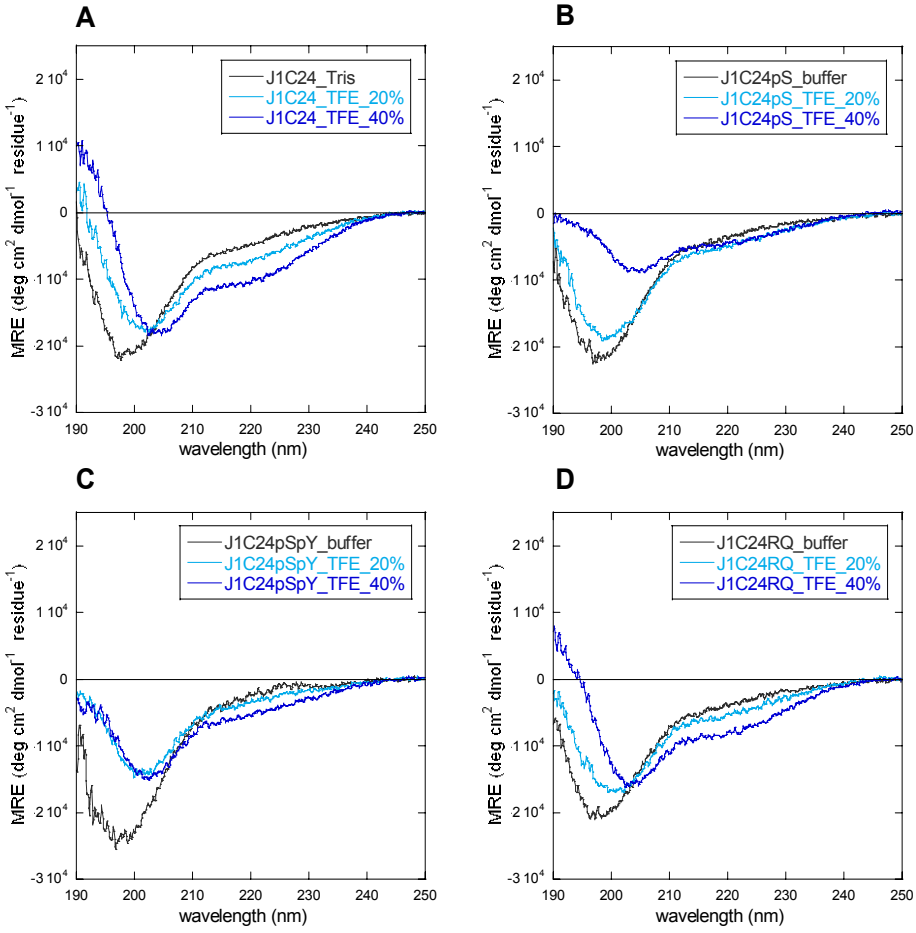


Figure 2.45. **Circular dichroism.** Far-UV CD spectra of different J1C24 peptides (45 μM) in 5 mM Tris⁄HCl buffer, pH 7.4, in the presence of different concentrations of 2,2,2-trifluoroethanol (TFE) (%, v/v): (**A**) J1C24, (**B**) J1C24pS, (**C**) J1C24pSpY and (**D**) J1C24RQ.

These spectra show no coil-helix transition in J1C24pS and J1C24pSpY, and suggest that phosphorylation in P-2 and P-8 may block binding of the J1_tmic C-terminal region to the membrane.

CD measurements were carried out also on J1C24RQ, which bears a mutation which involves the R→Q substitution at P-5. In buffer, the CD spectrum of J1C24RQ is almost identical to that of J1C24, typical of disordered peptides (**Figure 2.45 D**). In the presence of TFE, J1C24RQ is slightly less helical than J1C24, but still significantly more than the two phosphorylated peptides. In the presence of LMPG micelles, on the other hand, J1C24RQ behaves very similarly to phosphorylated peptides, with no coil-helix transition (**Figure 2.46 D**). J1C24RQ has a glutamine instead of an arginine, which abolishes the +1 charge at P-5. This charge change does not disturb the intrinsic helical propensity of the peptide, but evidently affects binding to negatively charged LMPG micelle.

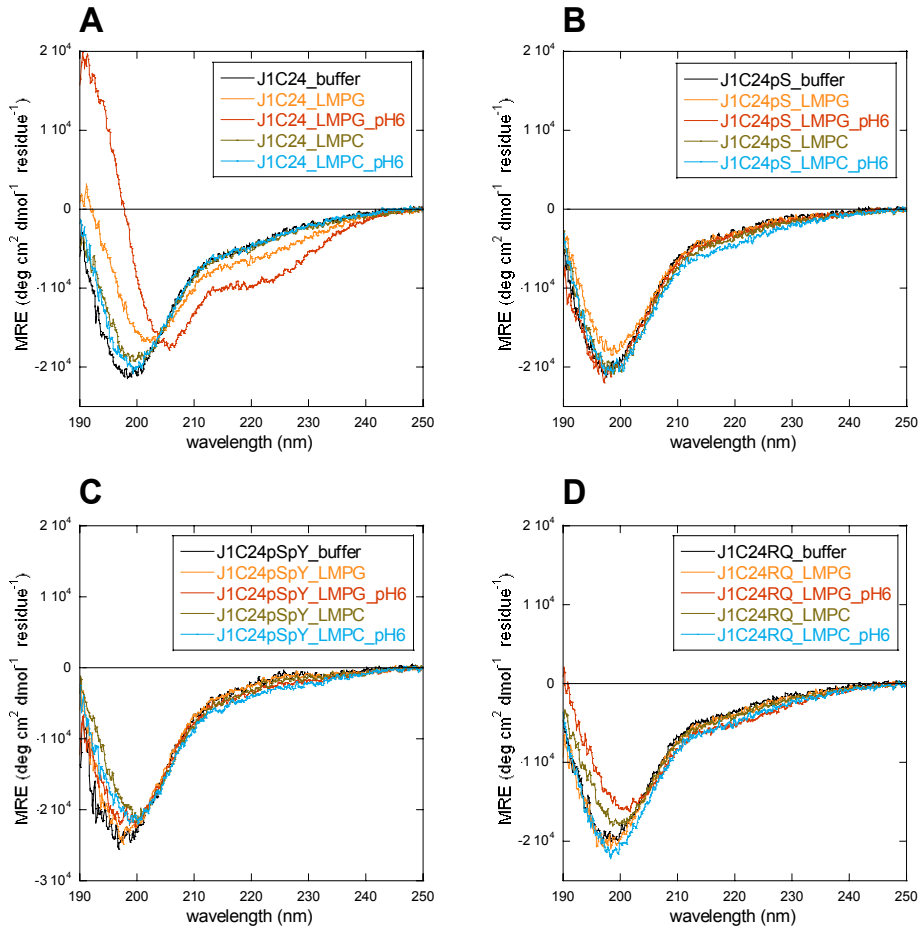


Figure 2.46. **Circular dichroism.** Far-UV CD spectra of different J1C24 peptides (45 μM) in 5 mM Tris⁄HCl buffer, pH 7.4, in the presence of LMPG and LMPC micelles, at pH 7.4 and pH 6: (**A**) J1C24, (**B**) J1C24pS, (**C**) J1C24pSpY and (**D**) J1C24RQ.

To confirm the CD results, tryptophan fluorescence experiments were also performed on all four peptides (J1C24, J1C24pS, J1C24pSpY and J1C24RQ) using W1196 as a probe (**Figure 2.4**). The emission spectrum of J1C24 in buffer shows a maximum (λ^{MAX}) around 360 nm (**Figure 2.47 A**), which is consistent with a solvent-exposed tryptophan and confirms the structural disorder of the peptide in solution.

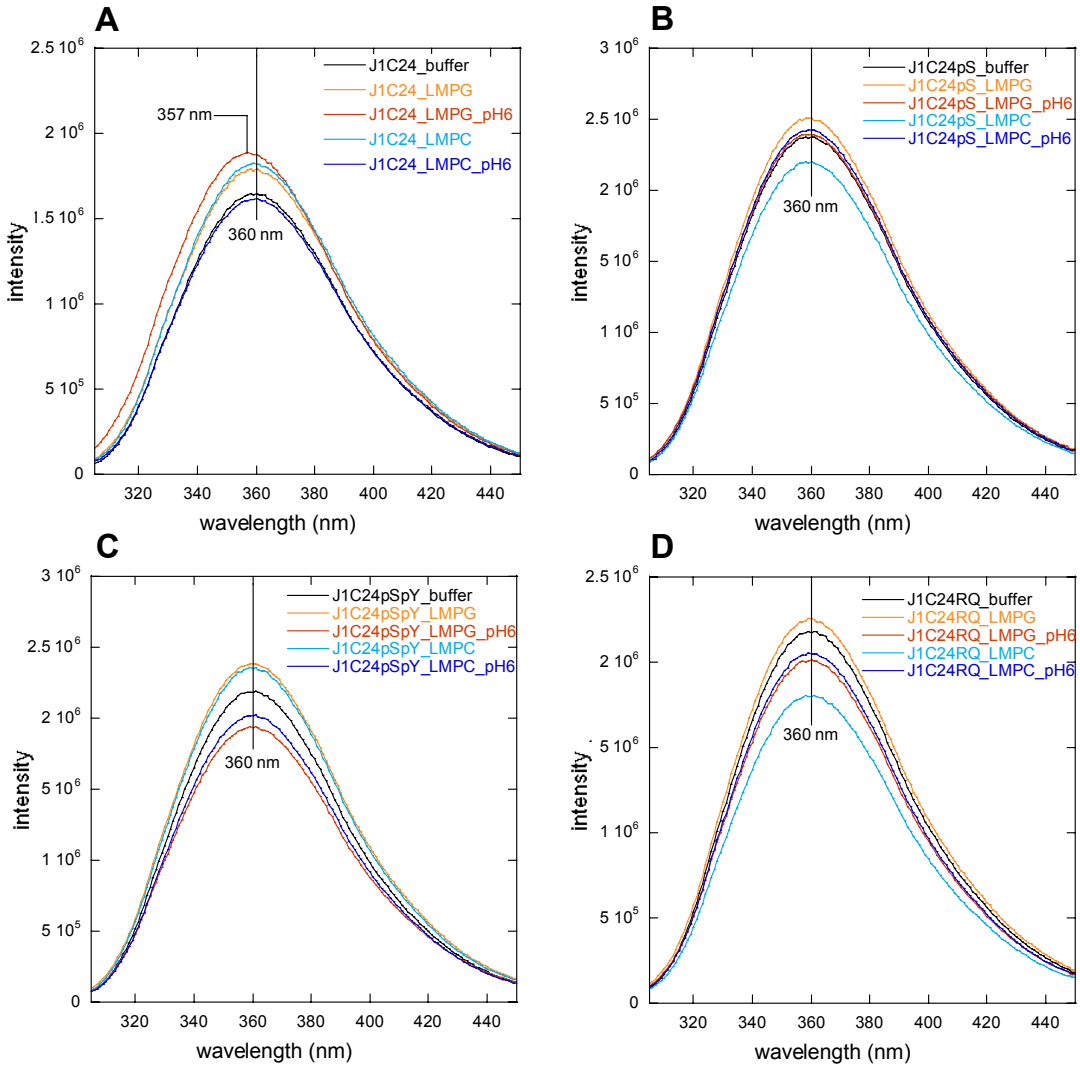


Figure 2.47. Fluorescence spectroscopy. Tryptophan fluorescence emission spectra of different J1C24 peptides (35 μM) in 5 mM Tris / HCl buffer, pH 7.4, in the presence in the presence of LMPG and LMPC micelles, at pH 7.4 and pH 6: (**A**) J1C24, (**B**) J1C24pS, (**C**) J1C24pSpY and (**D**) J1C24RQ.

In the presence of both type of micelles, the fluorescence emission spectra of J1C24 shows only small intensity changes with no λ^{MAX} shift, suggesting that W1196 remains solvent exposed even when the peptide partially folds upon binding to LMPG

micelles. In fact, secondary structure predictions suggest that W1196 is not included in the region of the predicted helix (N**W**^{1196}TNKQDNRDL **ESAQSL**NRMEYIV). However, experiments carried out on the full-length protein (J1_ic) showed the interaction of W1196 with negatively charged micelles (**Figure 2.13**). This discrepancy could be explained by the fact that other regions, present in J1_ic but not in the J1C24 peptide, might be required for the interaction of W1196 with micelles.

Tryptophan fluorescence anisotropy of J1C24 W1196 was also measured, but the values obtained were very low ($R<0.04$) and with large measuring errors.

The spectra of J1C24pS, J1C24pSpY and J1C24RQ showed all λ^{MAX} of ~360 nm (**Figure 2.47 B, C and D**), whether acquired in buffer or in the presence of micelles (LMPG and LMPC). Each of the three peptides, upon addition of LMPG or LMPC micelles, showed only small changes in the emission intensity. These results, together with the above λ^{MAX} values, support the presence of solvent exposed W1196.

3. DISCUSSION

3.1 Life on the membrane: Jagged-1 intracellular region as an interfacial protein

As a membrane-tethered protein, Jagged-1 intracellular region experiences neither the acqueous environment of the cytosol, nor the hydrophobic environment of the hydrocarbon core of the membrane bilayer, but is rather located at the interface between the two. The phospholipid interfacial layer is ~15 Å thick, is made by the polar groups of the fatty acid esters, the negatively charged groups of the phosphate moiety, and by the phospholipid head groups, which in turn determine the overall net charge of the phospholipid.

The interaction of a proximal protein with the phospholipid interfacial layer is mediated by charge-charge interactions, insertion of hydrophobic side chains into the hydrocarbon core, desolvation, and possibly reorganization of the lipid matrix itself. The complexity and plasticity of this peculiar environment make biophysical studies very challenging. This environment can be mimicked as an interface between an acqueous buffer and an artificial membrane. A wide range of different artificial membrane systems were developed in time, especially to address the study of integral membrane proteins ("***Biophysical methods***"). These systems include SDS and DPC micelles, lysophospholipid micelles, bicelles, nanodiscs, and liposomes (**Figure 2.21**). Their use has been recently extended to the study of proteins that are associated to the membrane through covalently linked lipid anchors, proteins that bind specific phospholipids, and proximal proteins, like the intracellular regions of transmembrane proteins and receptors.

In the presence of SDS micelles, negatively charged lysophospholipid micelles (LMPG) or negatively charged phospholipid vesicles (DMPG, DMPS), which are prevalent components of the inner layer of the plasma membrane in eukaryotes, J1_tmic gains secondary structure (**Figure 2.14**), (**Figure 2.19)**. The helical content measured by CD is consistent with secondary structure predictions (**Table 1**), (**Figure 2.10**). No changes in CD spectra were observed in the presence of vesicles formed by a zwitterionic phospholipid like DMPC, suggesting that the negative charge density at the surface of SDS micelles or phospholipid vesicles is required to promote binding

and secondary structure formation. Binding of J1_tmic to SDS micelles and DMPG liposomes was confirmed by tryptophan fluorescence experiments.

In the presence of SDS micelles and DMPG liposomes, the formation of secondary structure is strongly pH-dependent, with a sharp increase in the helical content from pH 7 to 6 (**Figure 2.15),** (**Figure 2.20)**. As J1_tmic contains six endogenous histidines, it is possible that protonation of one or more of the histidines is promoting helix formation or extension (**Figure 2.10**). On the other hand, a similar pH effect was observed also for J1C24, which contains no histidine residues. It is also possible that the particular environment of the interface affects the effective pK_a of other protonable groups, such as the carboxylate side chain of aspartic and glutammic acid. The possible biological relevance of this observation is not clear. The biophysical properties of the interface between the cytoplasm and the plasma membrane are not very well known [152], and it is plausible that the negatively charged head groups of phospholipids present in the membrane of eukaryotic cells can generate a pH gradient [153]. From pH mapping by fluorescence, it was actually reported in an early study that the effective pH in proximity of the membrane in yeast cells is ~6.0 [154], which supports the physiological relevance of the pH-dependent secondary structure formation in J1_tmic.

It may be argued that the conformational changes observed are induced by the hydrophobic interaction of SDS or phospholipid vesicles with J1_tmic, rather than by the charged surface of micelles or vesicles. It can be remarked, however, that the fatty acid chains in SDS and in the phospholipids used (DMPG, DMPS and DMPC) are similar, if not identical (**Figure 2.21)**. If the conformational change is induced by hydrophobic interactions, similar effects should be observed. On the contrary, CD spectra display distinct features, depending on the conditions. Most significantly, different types of phospholipids, depending on the charge of the polar head, have different effects. Moreover, hydrophobic interactions are expected to be rather insensitive to pH changes. On the contrary, in the presence of SDS micelles and DMPG phospholipid vesicles, the helical content of J1_tmic is markedly dependent on pH, suggesting that the conformational change is driven by polar, rather than hydrophobic interactions.

The titration with SDS revealed that SDS triggers binding below its critical micellar concentration (2–8 mM, depending on ionic strength), with saturation around

1 mM for 7.5 μM J1_tmic, suggesting that J1_tmic can drive the formation of SDS micelles while binding on their surface. This effect is not unusual, as it has already been observed with α-synuclein, another membrane-interacting protein [155]. The most evident helical transition occurs from 0.5 mM to 1 mM SDS, which is the same SDS concentration range of the anisotropy increase. Additionally, in the range below 0.5 mM, the CD spectra show no ellipticity increase at 206 nm and 222 nm (there is no helical increase in J1_tmic), but show significant signal decrease (**Figure S3)**. These results are consistent with abnormally high anisotropy at the SDS concentration <0.5 mM and are probably due the interaction of J1_tmic with the aliphatic chains of SDS molecules which most probably do not assemble into micelles at that concentration. Altogether, J1_tmic drives the micelle formation in the range of 0.5 to 1 mM of SDS.

In the attempt to get a better insight into the J1_tmic-membrane interaction, we carried out solution NMR studies using lysophospholipid micelles (**Figure 2.23),** (**Figure 2.24)**. These micelles are especially attractive because they are prepared in a straightforward manner, they are relatively stable in time at the temperature and pH conditions used for acquisition of NMR data, they have a well defined composition, and they are biologically relevant, in that not only the corresponding lysophospholipids are natural components of eukaryotic plasma membranes, but their head groups are identical to those found in natural phospholipids. The phospholipid vesicles have shown to be inadequate for high-resolution NMR studies of membrane-associated peptides and proteins as their reorientation time is too long on the NMR timescale, giving rise to large bandwidths [156]. Therefore, the rapidly reorienting phospholipid micelles are more suitable for high resolution NMR studies [157]. The head groups of the lysophospholipid used (LMPG and LMPC) are identical to those of the corresponding liposomes (DMPG and DMPC).

The ^{15}N-HSQC spectrum of J1_tmic in the presence of LMPG micelles, and not LMPC micelles, displays significant differences compared to the spectrum recorded in buffer. These chemical shift changes confirm the interaction with LMPG micelles and a possible conformational change in J1_tmic, which is consistent with CD results.

While LMPC is zwitterionic, with the external layer of the micelle dominated by the positive charge of the choline moiety, LMPG is negatively charged because of

the phosphate moiety. It is thus possible that the interaction between J1_tmic, which has a calculated pI of 9.3, a total of 29 positively charged residues (R+K+H) and a net charge of +10, with the negatively charged LMPG micelle is favored by electrostatic interactions. Charge-charge interactions, however, are not likely to be the only driving force for binding. The small but measurable shifts in the positions of the W side chain $^{\varepsilon 1}$NHs might be indicative of the insertion of J1_tmic hydrophobic N-terminal residues in the lipid core of the micelle. The fairly good chemical shift dispersion in the N/Q side chain NHs suggests a difference in the environment experienced and possibly specific hydrogen bonding. Even more interestingly, the appearance of relatively sharp, distinct resonances that can be tentatively assigned to $^{\varepsilon}$NHs of the R side chain suggests the presence of hydrogen bonds between these atoms and possibly the negatively charged phosphate groups of the LMPG micelle. Indeed, binding to negatively charged but not to zwitterionic lysophospholipid micelles was observed also with the Jagged-1 C-terminal peptide, J1C24, which has a total net charge of -1 (-2 including the C-terminal carboxylate group) (**Figure 2.46)**. This result confirms that the interaction between the cytoplasmic tail of Jagged-1 and the inner leaflet of the membrane is quite complex, and is governed not only by charge interactions, but also by the formation of specific hydrogen bonds and possibly by the insertion of hydrophobic side chains into the hydrocarbon core of the membrane. Recently, it has been reported that the interaction between the cytoplasmic domain of the CD3-ε chain of the T-cell receptor and the membrane involves the insertion of a tyrosine side chain into the lipid bilayer [158]. This has important functional consequences, because this specific tyrosine is then subtracted to the cytosolic environment and cannot be phosphorylated. Although fluorescence spectroscopy is in principle capable of detecting differences in the environment of aromatic residues, its use is mostly restricted to tryptophan residues, which display a blue shift in the emission maximum when buried in a hydrophobic environment. For tyrosine residues, only a difference in the emission intensity, but not in the wavelength, can be detected upon changes in the solvation. When several tyrosines and tryptophans are present in the sequence, it is very difficult to extract any useful information from fluorescence spectra. Specific information can be instead obtained from NMR. Upon insertion of a side chain in to the hydrocarbon layer, specific NOEs between amino acid side chains and the lipid

chain can be detected, provided that the system is suitable for high resolution NMR studies in solution.

A short segment of basic amino acids is present in the intracellular region of most type I membrane proteins, close to the transmembrane region. Lipid binding of the intracellular region of several multichain immune receptors was shown to be related to the presence of clusters of basic residues [159]. Targeting of proteins to the membrane through basic segments occurring in mainly disordered regions have been shown for src [160], K-ras4B [161], rac1 [162] and PLC-ζ [163]. Recombinant fluorescent probes exploiting electrostatic interactions with the plasma membrane were used to study the changes in charge distribution occurring during phagocytosis [164]. On the other hand, basic amino acids are also found in nuclear localization signals [165]. For Jagged-1, which undergoes regulated intramembrane proteolysis, basic segments may thus play a dual role, stabilizing the interaction with the inner leaflet of the plasma membrane, and targeting the intracellular region to the nucleus, once the cytoplasmic tail has been proteolytically cleaved.

If electrostatic interactions between the basic intracellular region of Jagged-1 and the negatively charged phospholipids of the membrane are important, it can be expected that phosphorylation, which is reducing the positive net charge, might significantly reduce the affinity of J1_tmic for the membrane. Desorption from the membrane induced by phosphorylation has been actually proposed for the myristoylated alanine-rich C kinase substrate (MARCKS) [166] and for K-ras [167]. Because potential phosphorylation sites are predicted in the helical region preceding the PDZ-binding motif, it can be speculated that multiple phosphorylation of this region may decrease the affinity for the membrane, thus increasing the capture radius of the PDZ-binding motif.

Despite the overall positive net charge, J1_tmic contains a stretch of acidic residues (EVEEDDMD) (**Figure 2.10)**. Although the negative charge of this segment is expected to locally reduce the affinity for the negatively charged phospholipids, it should be born in mind that the peculiar environment of the membrane/cytoplasm interface could affect the effective charge in at least two manners. First, coordination of calcium ions could reduce or neutralize the negative charge of E/D residues [168]. Second, desolvation at the interface between the protein and the membrane could destabilize the carboxylate group in favor of its protonated form, thus reducing the

effective pK_a value [169]. The same effect could be achieved by the presence of the lipid phosphate moieties, which would also destabilize the charged carboxylate.

A yet to be addressed issue is if the intracellular region of Jagged-1 displays any specificity in the recognition and binding of phospholipids such as phosphoinositides which, despite their relatively low abundance (~1%) have a key role in several signaling pathways [170], [171].

Preliminary ^{15}N relaxation studies showed very variable R_2 values among NH peaks (**Figure 2.25**). A possible scenario consistent with the wide range of R_2 observed is given by the simultaneous presence of regions that are partially or totally embedded in the micelle, other regions that bind the surface of the micelle, probably in a dynamical way, and other regions that do not significantly interact with the micelle and freely tumble in solution.

Sequential assignment of backbone resonances followed by the analysis of Hα, Cα, Cβ, C' secondary shifts and the estimation of $^3J_{HNH\alpha}$ coupling constants from 3D HNHA spectra will supply essential information on the secondary structure of J1_tmic bound to LMPG micelles [172]. Measurement of temperature and protection factors of backbone and side chain HNs will be helpful in mapping the protein/micelle interface [173]. Determination of ^1H-^{15}N NOEs, together with a more detailed analysis of relaxation data, will provide a view of the dynamics of J1_tmic bound to micelles [174].

The partial folding of the cytoplasmic domain of Jagged-1 accompanied by its association with the inner side of the cell membrane may have relevant effects on the function of Jagged-1 in Notch signaling [57], [175], [51]. For instance, it may selectively mask certain residues that are potential targets for post-translational modifications such as phosphorylation, ubiquitination, or O-glycosylation by β-N-acetylglucosamine [176], [177] while leaving others exposed for the same modifications. In a similar way, it may mask or expose selected binding motifs with respect to binding partners (**Figure 3.1**). The partial folding and association of the intracellular region of Jagged-1 with the membrane is also expected to reduce its "capture radius" [178] towards protein targets like PDZ-containing proteins. Despite the high number of single pass membrane proteins involved in signaling, little is known about the structure and function of their cytoplasmic tails and, to our knowledge, only few examples have been reported [179], [180]. The cytoplasmic tail

of the T-cell receptor ζ-chain [181], [182] binds to lipid membranes through a lipid-induced coil–helix transition dependent on phosphorylation [180]. Other cytoplasmic domains related to multichain immune recognition receptors were found to be intrinsically disordered even when bound to lipids [159]. A role of the cytoplasmic tail of membrane-spanning proteins in protein–protein interactions has also been proved, e.g. the case of the association between the N-terminal region of the membrane-bound tyrosine kinase Lck with the cytoplasmic tail of the T-cell coreceptors CD4 or CD8 [183]. In solution, on the contrary, J1_tmic is mainly disordered (**Figure 2.3**).

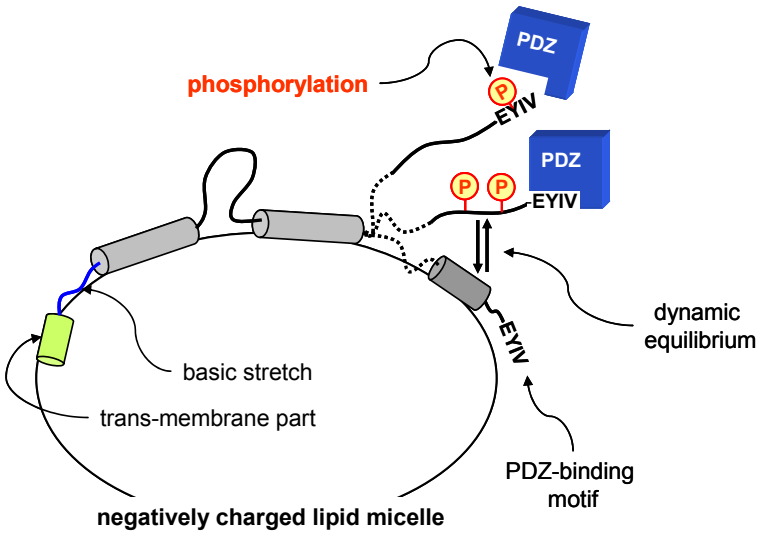


Figure 3.1. Model of J1_tmic bound to micelles. Proposed model of J1_tmic bound to a negatively charged phospholipid micelle. J1_tmic is shown as a cartoon model; the number, location and length of helical regions (cylinders) are based on secondary structure predictions; coil regions are drawn as black lines. The C-terminal region (last 24 residues) is shown in dynamic equilibrium between a membrane-bound unphosphorylated form and a PDZ-bound form phosphorylated at T1197 (P-22) and S1210 (P-8). Phosphorylation at Y1216 (P-2) reduces the binding affinity for AF6_PDZ.

3.2 Leaving the membrane: Jagged-1 intracellular region as a nucleocytoplasmic protein

The intracellular region of Jagged-1 drives endocytosis, mediates association with proteins implicated in the organization of cell-cell junctions and, through

regulated intra-membrane proteolysis, is released from the membrane as a signaling fragment. Although the precise cleavage site in Jagged-1 is not known, experimental evidence from the cleavage of Notch receptors suggests that it is placed at the first valine close to the inner side of the cytoplasm [51]. Thus, the putative Jagged-1 fragment released into the cytoplasm (J1_tmic) comprises the whole intracellular region (J1_ic) plus the 8 residue long transmembrane part at its N-terminus. J1_tmic may then exist in at least two forms: as a membrane-tethered protein located at the interface between the membrane and the cytoplasm, and as a soluble nucleocytoplasmic protein. Overall, far-UV circular dichroism (CD), tryptophan fluorescence (emission spectra and anisotropy) and ^{1}H-NMR data are consistent with the lack of tertiary structure and the presence of very little or no secondary structure, thus suggesting that in the experimental conditions tested J1_ic behaves as a mainly disordered protein in solution (**Figures 2.3 – 2.6**). The disordered conformation of J1_ic was further confirmed by the limited proteolysis experiments in which J1_ic was degraded by chymotrypsin very rapidly with respect to small-globular proteins (**Figure 2.7**). This result is indirectly confirmed by the extent of protein degradation occurring during the cell lysis. Using the same spectroscopic techniques, it was shown that J1_tmic is mainly disordered in solution as well, indicating that the additional intramembrane N-terminal residues do not promote folding of J1_ic. Additionally, results obtained from heteronuclear NMR experiments and size exclusion chromatography carried out on J1_tmic are consistent with a mainly disordered, but rather compact, state of J1_tmic in solution (**Figures 2.8** and **Figure 2.9**).

J1_tmic can be thus assigned to the family of "intrinsically disordered proteins" (IDP), which are also described as "intrinsically unstructured proteins" (IUP) or "natively unfolded proteins" [116], [184], [129], [185], [186]. From the biophysical point of view, these proteins (or regions) are characterized by a low content of secondary structure, no tertiary structure, and an expanded hydrodynamic radius compared to globular proteins. This is the consequence of a characteristic amino acid composition, which is enriched in charged and flexible amino acids and depleted in hydrophobic amino acids (Ile, Leu, Val, Trp, Tyr and Phe) that usually form the hydrophobic core of globular proteins. Based on this peculiar amino acid composition, intrinsic disorder propensity can be predicted using a plot (**Figure 3.2**) of the protein mean net charge versus mean hydrophobicity, where the mean net charge is defined as the calculated total net charge of the protein at pH 7.0, divided by

the total number of residues and the mean hydrophobicity is defined as the sum of the normalized hydrophobicity (scale if 0 to 1) of all residues divided by the number of residues [186], [115].

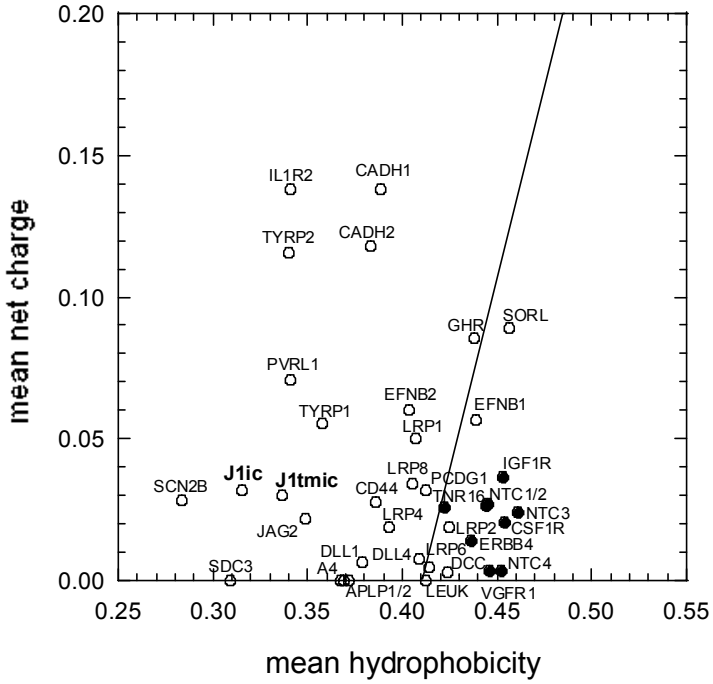


Figure 3.2. Intrinsic disorder. Mean net charge versus mean hydrophobicity calculated for the intracellular region of 37 human presenilin/γ-secretase substrates that undergo regulated intramembrane proteolysis. Proteins that contain globular domains in the cytoplasmic tail are shown as filled circles. The line ideally separates intrinsically disordered proteins (left-hand side of the plot) from natively folded ones (right-hand side of the plot). A4, amyloid β A4 precursor; APLP1/2, amyloid-like proteins 1/2; CADH1, E-cadherin; CADH2, N-cadherin; CD44, CD44 antigen; CSF1R, colony stimulating factor 1 receptor; DCC, netrin receptor; DLL1/4, delta-like 1/4; EFNB1/2, ephrin-B1/2; ERBB4, receptor protein tyrosine kinase erbB4; GHR, growth hormone receptor; IGF1R, insulin-like growth factor 1 receptor; IL1R2, interleukin 1 receptor II; J1ic and J1tmic, jagged-1; JAG2, jagged-2; LEUK, leukosialin (CD43); LRP, low-density lipoprotein receptor related proteins; NTC1-4, Notch receptors 1–4; PCDG1, proto cadherin c A1; PVRL1, nectin-1; SCN2B, sodium channel b2 subunit; SDC3, syndecan-3; SORL, sortilin-related receptor; TNR16, tumor necrosis factor superfamily member 16; TYRP, tyrosinase-related proteins; VGFR1, vascular endothelial growth factor receptor 1.

Although very simple, this plot is quite effective in clustering globular proteins and IDPs into two distinct groups. The biophysical characterization of the

Jagged-1 intracellular region is in accordance with the high degree of intrinsic disorder predicted from the protein amino acid sequence (**Figure 2.10**).

It was proposed that IDPs can be further classified in different groups, pure random coils, pre-molten globules and molten globules [187]. This terminology derives from classical biochemical studies on protein folding, and was devised to rationalize different degrees of protein compactness. While all these states lack a tightly packed hydrophobic core and a stable tertiary structure, random coils display little or no secondary structure and a high hydrodynamic radius, and molten globules are more compact and may contain some secondary structure elements. The difference between pre-molten globules and molten globules is essentially dynamic, with pre-molten globules and molten globules showing fast and slow exchange, respectively, between the different conformational states.

Size exclusion chromatography, NMR, and CD results tend to exclude a pure random coil structure for J1_tmic, and rather suggest a molten globule state. The hydrodynamic radius, R_S, calculated for J1_tmic by size exclusion chromatography (23.57 ± 0.35 Å) is intermediate between the value (19.6 ± 0.3 Å) expected for a globular protein of the same MW (15.5 kDa) and that of a random coil (36.4 ± 0.7 Å). According to the calculated R_S, J1_tmic can be placed closer to molten globules (R_S = 22.2 ± 0.3 Å) than to pre-molten globules (R_S = 27.1 ± 0.5 Å).

Finally, pre-molten globules and molten globules, unlike random coils, usually possess a propensity to form secondary structure. Indeed, four helical stretches can be identified in J1_tmic from the consensus secondary structure predictions (in total 23–35%, depending on the threshold set for the probability score). CD spectra of J1_tmic in the presence of TFE confirmed its intrinsic helical propensity (50% of α-helical content at 50%, v/v of TFE) (**Figure 2.11**), (**Table 1**). TFE is a widely used co-solvent in the biophysical characterization of the solution structures of peptides and proteins. It promotes secondary structure formation by reducing the protein backbone exposure to the aqueous solvent and favoring the formation of intramolecular hydrogen bonds. Although considered as a helix-inducing co-solvent, TFE in fact stabilizes specific secondary structure elements depending on the intrinsic conformational propensity of the polypeptide chain. It was found to stabilize helical structure in regions with a high helical propensity [188], [189], [190] to stabilize β-sheets in proteins and in polypeptides with an intrinsic preference for the extended conformation [191], [192] while leaving unaffected polypeptides that have no intrinsic

conformational preference [193]. These results also suggest that Jagged-1 intracellular region could undergo some disorder-to-order transitions in certain cellular environments or specific interactions.

A more thorough characterization of J1_tmic would require a more precise determination of the hydrodynamic radius, for example by analytical ultracentrifugation, dynamic light scattering, or small angle X-ray scattering (SAXS) [194]. Although technically more complex, the latter is the technique of choice because is more sensitive than dynamic light scattering and ultracentrifugation, and measurements can be performed using relatively diluted protein samples. Furthermore, it may provide, in some cases, also information on the shape of the molecule, and it has extensively used to determine the hydrodynamic radius of globular, partially folded, and unfolded proteins.

Additional information on transient secondary structures and local interactions could in principle be provided by solution NMR which, however, in this specific case would be of limited utility given the line broadening due to slow exchange processes.

Fourier-transform infrared spectroscopy (FTIR) has also been used to study protein secondary structure and conformational changes [195]. By tracking changes in hydrogen-deuterium exchange in the protein amide groups, this method can be also used to monitor protein structure dynamics. Additionally, when estimating the percentage contents of protein secondary structures, a combination of FTIR and CD can increase prediction accuracy.

It can be argued that the structural disorder of J1_tmic is not an intrinsic property, but a result of the harsh conditions used in the protein purification process or of the non-native environment in which the biophysical characterization is carried out. Unwanted effects of the purification protocol could be ruled out purifying the protein using mild, non-denaturing conditions. The most effective way to solve this issue is, however, to determine the structure of the protein within the cellular environment. This is now feasible using in-cell NMR ("***Biophysical methods***") [196], [197], [198]. NMR experiments carried out on the ^{15}N-labeled J1_tmic overexpressed in *E. coli* cells confirmed that J1_tmic is mainly disordered. The main difference between the in-cell HSQC spectrum and the HSQC spectrum spectra recorded on the purified J1_tmic in PBS is a significant line broadening of NH peaks in the in-cell NMR HSQC (**Figure 2.8**). This is probably due to the reduced protein tumbling in viscous media such as the cytosol, although association with other cellular factors to yield

higher molecular weight complexes cannot be ruled out. In-cell NMR experiments on the intrinsically disordered FlgM demonstrated its partial folding in *E. coli* [130]. FlgM is a bacterial but not an endogenous protein of *E. coli*. The endogenous cellular factors were thus excluded as a trigger of its folding and it was proposed that folding is mainly due to molecular crowding effects. In-cell results were reproduced by *in vitro* NMR experiment, in which FlgM was dissolved in high concentrations of either other proteins (bovine serum albumin) or small molecules (sugar). However, not all intrinsically disordered proteins become more compact in the crowded cellular environment. The same research group also demonstrated that α-Synuclein [199], a natively disordered protein that adopts secondary structure elements in very dilute solutions, does not show conformational changes in the cytoplasm of *E. coli* cells.

It can be argued that, as J1_tmic is a human protein, the environment provided by *E. coli* cells, although closer to the native environment than a buffer solution, is not the native environment provided by a human cell. Actually, in-cell NMR experiments carried out using a eukaryotic expression system like yeast, *Pichia pastoris*, or SF9 insect cells, provided that the expression levels are sufficient for detection, would represent a further step in mimicking the native cellular environment ("***Biopysical methods"***). As an alternative approach, in-cell NMR experiments could be carried out on *Xenopus* oocytes [200]. Oocytes are very large eukaryotic cells that can be inoculated with the purified, recombinant labeled protein, thus providing an alternative to endogenous expression of the target protein. As a further limitation of in-cell experiments carried out with *E. coli* cells, it cannot be excluded that a protein, although intrinsically disordered, might fold, or undergo a conformational change upon post-translational modifications. In the case of a human protein, the molecular machinery required would be probably absent in *E. coli*.

The association between a well defined globular structure and a specific biological function is a paradigm that has governed structural biology for decades, and it derives partially from early biochemical studies on enzymes, and partially from the overwhelming role of crystallography [201]. From more recent studies, it is now clear that well defined globular structures represent only part of what can be defined as the structural continuum [187],[119]. While there can be no doubt that protein structure and function are intimately linked, numerous intrinsically disordered proteins have now been recognized. The evolutionary persistence of such proteins represents strong

evidence in favor of their importance and raises intriguing questions about their role in biological processes.

Regions that are predicted to be intrinsically disordered are quite common, especially in eukaryota [116]. These regions, or entire proteins, often play an important role in protein–protein interactions, especially in complex signaling networks [117] like those in multicellular organisms. With their extended conformations, IDPs offer a very large interaction surface compared to their volume, which allows them to play as docking stations hosting several protein interactions motifs. Moreover, the lack of a rigid globular structure is a functional advantage, as their plasticity allows them to interact efficiently with several different targets.

Recognition of a target protein or DNA by an intrinsically disordered protein or region is often associated with a gain of structure in the IDP or region. Folding upon binding is accompanied by a significant decrease in conformational entropy, but also involves the formation of a large number of interactions between the IDP and the ligand, leading to a negative binding enthalpy which compensates the entropy loss [119]. This balance between the enthalpic and entropic effects enables the inducible binding characterized by high specificity and low affinity. These interactions are easily reversible and thus very suitable for the signaling pathways where the associated proteins must quickly dissociate when the signaling is over. Folding upon binding to negatively charged artificial membranes was demonstrated for Jagged-1 cytoplasmic tail (discussed in the previous paragraph).

The only experimentally determined binding partner of the cytoplasmic tail of Jagged-1 is afadin. This interaction is strictly local, in that it involves the PDZ domain of afadin and the C-terminal PDZ binding motif of Jagged-1, and does not imply folding of Jagged-1 upon binding to the PDZ (**Figure 2.34**, **Figure 2.35**, **Figure 2.39**, **Figure 2.40**). Afadin is however a large multidomain protein, and Jagged-1 intracellular region contains several docking sites. It cannot be then ruled out that the interaction between afadin and Jagged-1 cytoplasmic tail might not be restricted to the characterized interaction, and might involve additional regions, either in Jagged-1 or in afadin, or both. It has been proposed that phosphorylation of the PDZ binding motif may act as a modulator of the interaction between Jagged-1 and afadin, but this hypothesis will have to be verified by *in vitro* and *in vivo* experiments, as well as the presence of other post-translational modifications. Also, it is not known how ubiquitination, or RIP may affect the interaction with afadin.

Intrinsic disorder in the cytoplasmic region of type I membrane proteins that undergo regulated intramembrane proteolysis mediated by the presenilin/γ-secretase complex is probably not unique to Jagged-1 [48]. The charge/hydrophobicity plot [186], [115] calculated for the intracellular region of a series of human membrane proteins that are cleaved by presenilin shows that most of the RIP substrates, including Jagged-1, actually fall in the left-hand side of the plot (natively unfolded proteins) (**Figure 3.2**). All the proteins that clearly fall in the right-hand side of the plot contain, along with disordered stretches, structured domains (**Figure S15**). Nevertheless, TFE can induce helix formation in J1_tmic in even a more effective way than SDS micelles or phospholipid vesicles. This is of particular significance in view of the fact that most of the presenilin/γ-secretase substrates considered in (**Figure 3.2**) release fragments that are translocated to the nucleus and are involved in transcriptional regulation. This is the case also for Jagged-1, which was shown to activate gene expression through the AP1 element [51]. Control of transcription by the released signaling fragments probably does not occur in a straightforward manner, but through the interaction with transcription factors and transcriptional complexes that have not been identified yet. In this scenario, the intrinsic propensity to adopt a particular type of secondary structure may facilitate folding when binding to target proteins occurs [202]. The identification of post-translational modifications that can play a role in the function and structure of Jagged-1 cytoplasmic tail, as well as the identification of binding partners at the membrane/cytoplasm interface, in the cytosol, and in the nucleus, represent issues that need further investigation.

3.3 Social life: the interaction with afadin

Jagged-1, through its C-terminal EYIV sequence, was shown to interact with the PDZ domain of the protein AF6 in a PDZ-dependent manner [56], [57]. While this interaction couples Notch signaling with proteins involved in the organization of cell-cell contacts, its precise biological role is still unclear.

3.3.1 A qualitative analysis of the binding

NMR chemical shift mapping showed that Jagged-1 C-terminus binds into the βB/αB groove of AF6_PDZ in a canonical way (**Figure 2.35**) [203], [67]. The peptide binds in an extended conformation in the groove between the βB-strand and the αB-

helix of the PDZ domain, complementing the PDZ βB-strand and in an antiparallel orientation with respect to it, with the peptide C-terminal carboxylate group pointing towards the conserved loop ((R/K)xxxGψGψ) that precedes the βB-strand. Chemical shift mapping is a very powerful technique ("***Biophysical methods***") [204], [205], [140], but it should be born in mind that, alone, does not provide any direct information on the conformation of the bound ligand, but only on the binding site. In this specific case, the mode of peptide binding could be inferred comparing the CSP profiles with those obtained for a related peptide bound to the same PDZ domain, for which the NMR structure of the complex was calculated, and from the structural data available in the literature. The structure determination of the complex would require additional data, like the acquisition of filtered and half-filtered NOESY spectra to identify intermolecular NOEs (between the unlabeled peptide and the labeled PDZ) and intramolecular NOEs (within the unlabeled peptide) [206], or NOE transfer experiments [207]. In PDZ/ligand interactions, binding of the peptide to the PDZ is largely due to the formation of hydrogen bonds between the backbone of the peptides and the βB-strand of the PDZ [208]. NMR experiments finalized at the detection of hydrogen bonds formed by the βB-strand of the PDZ can be a straightforward way to assess the binding mode. For example, HSQC spectra acquired with and without solvent presaturation on the free and on the bound PDZ would be helpful in providing such an information [209], [210]

As an alternative to the direct structure calculation of the complex, docking protocols that are able to use a wide variety of restraints, including chemical shift perturbation data, are becoming popular[211], [212].

The large majority of chemical shift mapping studies are based on chemical shift perturbation of protein backbone amides. This approach has certainly several advantages: the protein needs to be labeled with ^{15}N only, ^{1}H,^{15}N HSQC spectra usually display well dispersed signals, are highly sensitive experiments, and provide residue specific information (except on prolines). On the other hand, the chemical shift of backbone NHs is strongly affected by subtle changes in secondary structure, hydrogen bonding, exchange with the solvent, pH and temperature variations. Moreover, the chemical shift of backbone NHs is expected to be significantly affected when ligand binding involves the formation of hydrogen bonds or, more generally, interactions of the protein backbone with the ligand, like in the case of PDZ/peptide

binding. They are expected to be much less affected when side chain/side chain or hydrophobic interactions are involved.

To complement the information provided by backbone NH chemical shift mapping, ^{13}C chemical shift mapping experiments were also carried out. More specifically, CSP analysis was used to monitor side chain perturbation of alanine, valine, leucine, isoleucine, threonine and methionine methyl groups.

This analysis proved to be very useful as there is a significant number of methyl-bearing residues in the regions (βB, βC, αA and αB) involved in the interaction with the ligand, and was in very good agreement with backbone NH CSP results (**Figure 2.37**).

Interestingly, significant CSP value were also displayed by residues belonging to the αA-helix and the βC-strand, regions that are distal from the peptide binding groove and not directly involved in the binding. The perturbations in the βC-strand might be partially explained by the hydrogen bonds connections with the βB-strand. However, it has been suggested that the perturbation in regions distal from the peptide binding might arise from a cascade effect [142], [143]. This effect is due the long-range energetic couplings between distal residues. It could be verified experimentally through double-mutant cycles method [213] and is supported by multiple sequence alignment of the members of a conserved protein family [142]. This allosteric communication may occur also in the absence of a protein conformational change [214] as was demonstrated by NMR relaxation experiments on the second PDZ domain from the human tyrosine phosphatase [143].

NMR chemical shift mapping also showed that binding is strictly local, in that it involves mainly the last six residues of Jagged-1 intracellular region and the binding groove of the PDZ. Binding of short C-terminal sequences offers an easy way for PDZ proteins to interact with target proteins without disrupting the overall structure and function of their ligands. Although specificity is mainly determined by the four C-terminal residues (P0, P-1, P-2 and P-3) [144], [215] further ligand positions may also contribute to specific interactions [69], [216], [217]. The contribution of residues upstream P-3 was also demonstrated in some ligands of AF6_PDZ. This is most evident in the case of EphB2 and EphB3 ligands, which need at least six terminal residues to bind the AF6_PDZ [56]. In Jagged-1, the sequence conservation of the C-terminal region is not limited to the four residue long PDZ-interacting motif, but

extends well beyond and encompasses at least 20 residues. For this reason, mapping of the AF6_PDZ binding site was carried out using peptides that corresponds to the last 24 C-terminal residues of Jagged-1. The obtained NH-CSP and methyl-CSP profiles show great similarity with CSP profiles of the AF6_PDZ-J1C6 interaction. Similar perturbation of residues in AF6_PDZ confirms the same binding mode of the two peptides. Small differences in NH-CSPs were displayed only by A23, A24 and Q70 (**Figure 2.39**). These residues are all located at the edge of the peptide binding groove. As A23 and Q70 are significantly perturbed already upon binding to J1C6, it is possible that these differences arise from the fact that the positively charged amino terminus in J1C6 is replaced by a neutral amide in the longer J1C24 peptide. Moreover, residues in the βB/βC loop are not perturbed, neither with J1C6 nor with J1C24, and this is the PDZ region which was shown to interact with residues upstream P-3 [217], [208]. This loop is highly divergent in sequence, length, and main chain conformation among different PDZ domains. In AF6_PDZ, the βB/βC loop ($K_{25}G_{26}A_{27}G_{28}Q_{29}D_{30}K_{31}L_{32}$) is four residue longer than in PDZ3 of PSD-95, the PDZ domain with the most similar structure (RMSD 2.1Å) [75]. Usually, PDZ domains with a longer βB/βC loop use it for the specific interaction with ligands and can serve as an anchoring site for residues as far as the -7 position in the peptide ligands [216].

However, structural analysis demonstrated that the βB/βC loop of the AF6 PDZ domain is not involved in the interaction with Neurexin and Bcr [75]. Moreover, from NH-CSP profiles it is evident that the βB/βC loop region is equally perturbed in the C-terminal peptides from Jagged-1 and in Neurexin. It is possible that the lack of interaction between this loop and the ligand can be responsible for the AF6_PDZ binding promiscuity. Binding sites in AF6_PDZ were also mapped upon binding with the full-length Jaggged-1 intracellular region (J1_tmic). Chemical shift perturbations, when compared to the J1C24 binding, showed no significant differences.

3.3.2 A quantitative analysis of the binding

Binding of the C-terminal region to AF6_PDZ was studied not only in a qualitative way, through chemical shift mapping, but also quantitatively, measuring dissociation constants (K_d) of the complexes both by NMR and by SPR (**Table 2**), (**Table 3**), (**Figure 2.43**).

Measurement of K_d values from NMR data is time consuming and require high quality data, but provides invaluable residue specific information on binding in

solution, even when the affinity is low or very low (K_d in the μM-mM range) [204]. For all Jagged-1 C-terminal peptides, K_ds were estimated both from backbone NH and side chain methyl groups CSP values. In the fast exchange rate, which is the case for the Jagged-1-AF6_PDZ binding, dissociation constants can be calculated plotting CSP values against the molar ratios (peptide:PDZ), and fitting by nonlinear regression analysis. The K_ds were calculated for selected residues in βB, βC, αA, and αB regions mostly because of the large, and thus easy measurable, CSP values displayed by these residue. Mean K_d values measured from NH-CSP and methyl-CSP are in excellent agreement and lie in the micromolar range [151]. Results showed values of 17.2±2.3, 8.9±3.7, and 25.9±8.9 μM for J1C6, J1C24, and J1_tmic, respectively. The higher K_d obtained for J1C6 may arise from unfavorable interactions between the charged amino terminus of the peptide and the PDZ. The difference between J1C24 and J1_tmic, on the other hand, might arise from unfavorable electrostatic interactions between AF6_PDZ (calculated pI = 9.4) and J1_tmic (calculated pI = 9.3) or from a partially compact form of J1_tmic that would limit the availability of the C-terminal region for binding. Overall, results suggest that the binding is strictly local and generally does not involve Jagged-1 residues upstream the position P-5.

Binding affinities were also measured by SPR ("***Biophysical methods***"). In our experiments, K_ds estimated by SPR were higher compared to those measured by NMR. The differences between NMR and SPR measurements may arise from the experimental set up specific to each technique. While by NMR the protein-peptide binding affinity is measured in solution and at the equilibrium, the experimental conditions used in SPR experiments are quite different, in that one of the binding partners is immobilized on a solid phase, and binding is measured in a flowing system, i.e. not strictly in equilibrium conditions. However, because many PDZ ligands are membrane-associated and clustered [218], it is difficult to know whether solid phase or solution methods better simulate the *in vivo* situation.

It should also be considered that while K_ds estimated by NMR are residue specific (in fact, atom specific), the corresponding values obtained by SPR are "global" or, in other words, refer to a molecular interaction. Whether a molecular interaction can be quantitatively treated as the average of a series of residue specific interactions is a problem that, to our knowledge, has never been addressed.

3.3.3 J1_tmic does not fold upon binding

It was shown that the Jagged-1 binds the PDZ domain of afadin in a canonical mode, which implies an extended β-strand conformation of the Jagged-1 C-terminal region complementing the AF6_PDZ βB-strand [145]. The question that arises is whether binding could trigger folding or any significant conformational change in the Jagged-1 cytoplasmic tail. To address this question, NMR spectra of the ^{15}N-labeled J1_tmic in the presence of the unlabeled AF6_PDZ were recorded. ^{15}N-HSQC spectra show that, within the conditions used, binding of J1_tmic to the PDZ domain of afadin does not trigger folding (**Figure 2.40**) and confirm that this interaction is restricted to the C-terminal residues of Jagged-1 cytoplasmic tail. Whether the PDZ-binding motif is already preorganized in an extended conformation prior to binding (conformational selection) or if it becomes ordered upon binding (induced folding), remains to be determined [118]. Specific environmental factors might also influence this interaction. It was shown that the intracellular region of Jagged-1 has a remarked propensity to acquire helical structure upon binding to artificial membranes that mimic the inner leaflet of the cytoplasmic membrane. The interaction between Jagged-1 cytoplasmic tail and afadin occurs at the interface between the plasma membrane and the cytoplasm, and it is not known if this specific environment can affect binding. Furthermore, afadin is a very large multidomain protein (**Figure 1.8**), and it cannot be ruled out that additional regions could act as docking sites for the Jagged-1 cytoplasmic tail.

3.3.4 Tyrosine phosphorylation at P-2 reduces affinity

More than 400 different PDZ domains were identified in the human genome and many of them share remarkably similar peptide ligand-binding properties [219]. While the interaction between PDZ domains and short peptides has been studied thoroughly, most of the biological interactions, and especially those involving signaling, are dynamic and reversible, i.e. they must be capable of up- and down-regulation. How regulation of PDZ-mediated interactions is achieved, is in most cases a matter of speculation. One possible mode of control is through post-translational modifications, and in particular phosphorylation of residues within the C-terminal sequences that bind the PDZ domain.

Bearing serine or threonine residue at P-2, PDZ ligands of Class I are commonly phosphorylated by kinases. Serine phosphorylation at P-2 of β2 adrenergic

receptor abolished binding to EBP50 [220]; serine phosphorylation at P-2 of the Inward Rectifier K+ channel Kir 2.3 abolished binding to PSD-95 [81]; serine phosphorylation at equivalent position of the NR2B subunit of the NMDA receptors disrupts its interaction with PSD-95 and SAP102 [82]; threonine phosphorylation at P-2 of stargazin inhibits binding to PSD-95 [221] and so on. Beside the P-2 residue, which is critical for the binding with PDZ, there are evidences of phosphorylation in other positions of the ligand C-terminal region. Phosphorylation of serine in P-3 has been reported to prevent binding of GluR2 to glutamate receptor interacting protein (GRIP) [83] but not to PICK [85] and to enhance binding of MRP2-derived peptides to a series of PDZ proteins [84]; phosphorylation of serine at P-5 in LDL receptor-related protein 4 suppressed its interaction with PSD-95 and SAP97 [222]; phosphorylation of a tyrosine at P-7 in ErbB2 induces a reorientation of the phenolic ring but does not affect binding to erbin [216]. The interaction between syntenin-1 and syndecan-1 was recently shown to depend on the dephosphorylation of the tyrosine at P-1 in the EFYA PDZ-binding motif of syndecan-1 [223]. Binding of afadin PDZ domain to a library of immobilized peptides phosphorylated at different positions showed that strong binding inhibition (<50% binding) is associated with phosphorylation at P-2, a slightly smaller effect at P-1 (~50% binding) and a small but measurable effect at P-8 (~80% binding) [224].

In the Jagged-1 C-terminal region (last 24 residues), although no experimental evidence has been reported so far, there are four potential phosphorylation sites, and three of them are also predicted to be phosphorylated (**Figure 2.27 B**). One of them is in the PDZ binding motif at P-2 (Y1216), one at P-8 (S1210) and one far upstream the binding motif at P-21 (T1197). Phosphorylation of target proteins can be identified by *in vivo* or by *in vitro* experiments, or can be predicted by in silico methods. Although there are no doubts that determinations obtained by *in vivo* experiments are the closest to the real situation in the cell, these experiments are often difficult and time consuming. Moreover, traditional Mass Spectrometry methods were shown to underestimate the identification of phosphate-modified residues [225]. Commonly used *in vitro* experiments are based on the incubation of the target protein with the recombinant kinase, but the demonstration of effective phosphorylation *in vitro* does not necessarily imply that the substrate is phosphorylated also *in vivo*. In silico predictions are based on the prior knowledge of the kinase specificity, but do not include the role of substrate recruitment. To improve the likelihood that the predicted

modification makes sense in a biologically relevant situation, new in silico tools take in consideration also the subcellular co-localization of the kinase and its substrate, protein domain structure and evolutionary conservation [226]. Several on-line Web services exist to predict kinase-specific phosphorylation sites: GPS [227], PPSP [228], NetPhosK [226], ScanSite [229], KinasePhos [230], PredPhospho [231], Predikin [232], PhoScan [233], pkaPS [234] and others.

To examine the effects of phosphorylation on binding, chemical shift mapping was carried out on the $^{13}C,^{15}N$ labeled AF6_PDZ titrated with two phosphorylated J1C24 peptides, J1C24pS (phosphorylated at P-8) and J1C24pSpY (phosphorylated at P-8 and at P-2). No significant differences between the CSP plots of AF6_PDZ titrated with J1C24pS or J1C24 were observed (**Figure 2.41**).

However, from the comparison either with the CSP profile of J1C24 or with the CSP profile of J1C24pS, it is evident that J1C24pSpY interacts in a different mode with the residues of the βB/αB groove. This is most likely the consequence of the phosphorylated P-2 tyrosine. A similar effect was also shown in the CSP profile of the AF6_PDZ titrated with J1C6pY, the short peptide containing a phosphorylated tyrosine at P-2. Binding differences between J1C6pY and J1C6 are apparent from the different perturbation of E71 and A74. As in the case of J1C24pSpY, the effect of tyrosine phosphorylation at P-2 reduces the perturbation of A74 while increasing the perturbation of E71. The CSP profiles of the short peptides do not show significant differences for A23 and Q70. The discrepancy between J1C6pY and J1C24pSpY is not evident directly comparing CSP profiles of these two peptides, but can be explained by taking into consideration the CSP values of A23 and Q70 in the titrations with all four peptides (J1C6, J1C6pY, J1C24 and J1C24pSpY). As discussed, the different perturbation of A23 and Q70 could be a consequence of the positively charged amino terminus at P-5 of the short peptides. Therefore, the CSP profiles obtained for long peptides should be more reliable. Upon binding to J1C24, the most perturbed residue in the αB-helix is A74, which probably makes hydrophobic contacts with the aromatic ring of tyrosine at P-2, as proposed in the interaction with Neurexin [75] (**Figure 2.35**). With phosphorylation at P-2, the perturbation of A74 significantly decreases while increasing in Q70 and E71. It is likely that the negatively charged phosphate moiety at P-2 is electrostatically repelled from the negatively charged E71, which is on the same side of the αB-helix and close

to the phosphotyrosine side chain. This repulsion could increase the distance between phosphotyrosine at P-2 and A74, therefore reducing the CSP value (**Figure 3.3**). This is consistent with a significantly lower binding affinity of J1C24pSpY. Whereas no significant difference can be detected between J1C24 (mean K_d = 8.9±3.7 μM) and J1C24pS (mean K_d = 8.7±3.1 μM), tyrosine phosphorylation at P-2 decreases affinity for AF6_PDZ (mean K_d = 25.8±11.5 μM) (**Table 2**).

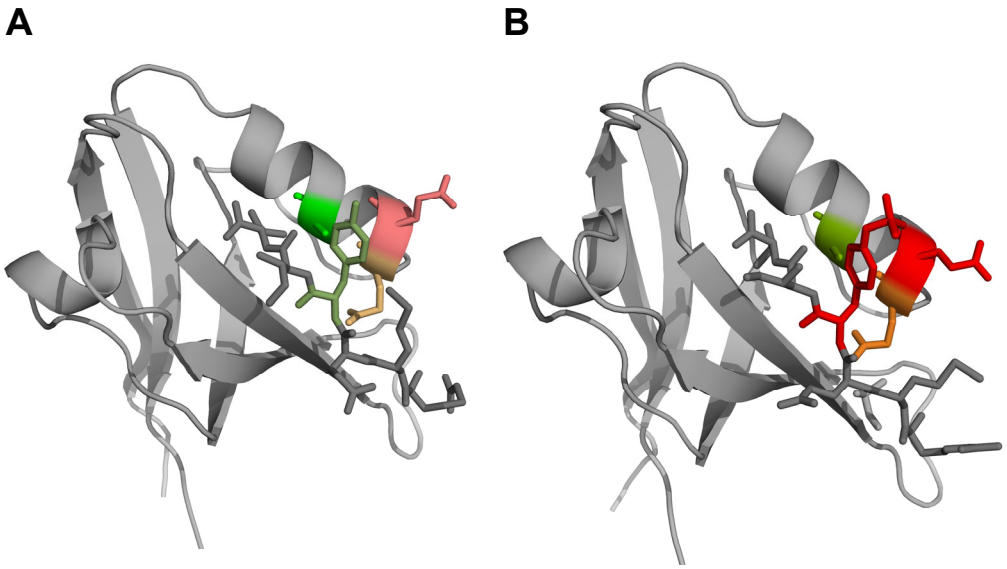


Figure 3.3. Model of AF6_PDZ bound to J1C6 and J1C6pY. The structural models of AF6_PDZ bound to J1C6 (**A**) or J1C6p (**B**) were built using the Modeller software. The NMR-derived structure of the complex between AF6_PDZ domain and the LFSTEV peptide was used as template (PDB: 2AIN). Only residues in the αB-helix of AF6_PDZ that are affected in a different way upon binding to either J1C6 or J1C6p are colored (Q70 in orange, E71 in red, and A74 in green). The intensity of the color is proportional to the NH-CSP value. In both peptides, Y at P-2 is colored in dark green (A, not phosphorylated) or red (B, phosphorylated). The figure was drawn using PyMol.

The relative binding affinities measured by SPR follow the same trend in which K_ds for J1C24 and J1C24pS are similar (37±3.8 μM and 28±5.7 μM) and significantly lower than in J1C24pSpY (K_d = 150±69 μM) (**Figure 2.43**).

It can be thus speculated that tyrosine phosphorylation of Jagged-1 at Y1216 may be used to modulate the interaction with afadin, either switching the binding on and off or shifting the binding specificity from one partner to another.

3.3.5 The R1213Q mutation at P-5 increases affinity

According to present knowledge, there are three genetic disorders associated with mutations in *JAG1*. In all of them, Alagille syndrome (AGS) [25], tetralogy of Fallot (TOF) [29], [26] and extrahepatic biliary atresia (EHBA) [28], truncations and mutations are related to the Jagged-1 extracellular region. In the EHBA patients, besides mutations in the extracellular region, one missense mutation was also detected in the cytoplasmic tail of Jagged-1 (R1213Q). This mutation involves the residue at P-5 of the PDZ-binding motif, and we therefore hypothesized that it could affect the binding with AF6_PDZ.

NMR chemical shift mapping experiments showed that the R1213Q mutation does not significantly affect the mode of binding of J1C24RQ to AF6_PDZ. SPR measurements showed, however, a nearly twofold higher binding affinity (K_d = 20±3.5 μM) of J1C24RQ compared to the wild-type peptide (K_d = 37±3.8 μM). A higher affinity is also supported by NMR results, although the limited solubility of J1C24RQ reduced the quality of the available data. (**Figure S14**).

This suggests that in developmental processes even a slight imbalance in the PDZ-mediated protein interaction network may have dramatic effects. Indeed, this case is not unique. In SANS (Scaffold protein containing ankyrin repeats and SAM domain), a D458V mutation at P-3 of the PDZ-binding motif (ALEDTEL-COOH) was found to be associated with an atypical form of Usher syndrome [235]. Because of the frequency of PDZ-mediated protein-protein interactions, it is likely that more disease-associated mutations in the PDZ-binding motifs are yet to be discovered.

3.4 All together now: phosphorylation may affect both binding to the membrane and to the PDZ

As a multidomain scaffold protein, AF6/afadin recruits different protein complexes near the plasma membrane cell-cell junctions. Through its single PDZ domain, it binds mainly transmembrane proteins. Thus, the interaction of the Jagged-1 intracellular region (J1_tmic) with AF6_PDZ most likely occurs at the membrane/cytoplasm interface. J1_tmic is disordered in solution, but partially folds upon binding to negatively charged liposomes and micelles. The partial folding and association of the intracellular region of Jagged-1 with the membrane is expected to reduce its “capture radius” [178] towards protein targets like PDZ-containing proteins.

As discussed previously, further NMR experiments are necessary for the mapping of helical regions and to determine the dynamics of J1_tmic bound to negatively charged micelles (LMPC). Nonetheless, from the consensus secondary structure prediction, one of the four helical stretches is located in the C-terminal region of J1_tmic (**Figure 2.10**). Indeed, far-UV CD showed that J1C24 is disordered in solution, but gains a significant helical content upon addition of TFE. In the presence of LMPG, but not LMPC micelles, J1C24 also gains a partial helical structure. Moreover, this coil-helix transition is pH dependent, as observed for J1_tmic. As J1_tmic has a positive total net charge (+10) whereas the C-terminal J1C24 peptide a negative total net charge (-1), binding of J1C24 to LMPG micelle cannot be driven only by electrostatic interactions. As already discussed, formation of specific hydrogen bonds and insertion of hydrophobic side chains into the hydrocarbon core of the phospholipids can also contribute to the binding. Furthermore, it can be speculated that desolvation at the interfacial layer may alter the effective pK_a of side chain carboxylates of the four acidic residues in J1C24. Protonation of these residues would change the total net charge of J1C24 from -1 to +3, leading to a more favorable electrostatic interaction with negatively charged phospholipids. If the interaction is mainly electrostatic, it can be expected that phosphorylation, which introduces a double negative charge, might significantly reduce the affinity of J1C24 for LMPG. Indeed, CD spectra show that phosphorylated peptides, J1C24pS and J1C24pSpY, show no coil-helix transition either in the presence of LMPG or LMPC micelles. Surprisingly, phosphorylation also abolishes the helical propensity of the same peptides, as shown by CD in water/TFE mixtures. Phosphorylation introduces a doubly negative charge on the side chain of an otherwise neutral residue (serine, threonine or tyrosine). This not only alters the total net charge of the peptide, but may also affect the interactions between the side chains of the peptide. The introduction of a double negative charge at P-2 and P-8 residues of J1C24 may influence the interaction with other charged residues (J1C24 contains 4 negatively and 3 positively charged residues) and eventually prevent the helical formation. The above results show that phosphorylation at P-8 is sufficient to reduce the helical propensity and block binding of the Jagged-1 C-terminal region to the membrane. Indeed, serine at P-8 is located in the predicted helical region (ESAQS$_{-8}$L).

It was also shown that the last six residues of J1_tmic are sufficient for the binding to AF6_PDZ and that serine phosphorylation at P-8 does not change the

binding affinity. As the above helical region precedes the PDZ-binding motif, it can be speculated that serine phosphorylation at P-8 may decrease the affinity for the membrane, thus increasing the capture radius of the PDZ-binding motif. On the other side, tyrosine phosphorylation at P-2 was shown to reduce the affinity for AF6_PDZ.

Circular dichroism spectra of J1C24RQ, the peptide with R1213Q substitution, show that the charge change at P-5 does not affect the intrinsic helical propensity of the peptide, but disrupts coil-helix transition in the presence of LMPG micelles. The same mutation increases the binding affinity to AF6_PDZ from K_d = 37±3.8 μM to K_d = 20±3.5 μM. This increase, even if not so dramatic, might have an additional indirect increment, when coupled with the above disruption of membrane binding.

3.5 Conclusions and future directions

This thesis focuses on some of the most puzzling and only partially explored issues in current structural biology and biophysics. The conformational plasticity of intrinsically disordered proteins and their ability to adapt to different environments and to bind to different partners, the peculiar biophysical properties of the membrane/cytosol interface and their impact on the structure and function of interfacial proteins, as well as the role of phosphorylation in protein-protein and protein-membrane interactions were investigated through the biochemical and structural characterization of the recombinant protein corresponding to the cytoplasmic tail of one of the five human Notch ligands, Jagged-1.

NMR studies aimed at the mapping of the helical regions and structure determination of Jagged-1 cytoplasmic tail bound to lysophospholipid micelles used as models of the cell membrane are currently underway. SPR experiments finalized at the thermodynamic characterization of the interaction between artificial membranes and Jagged-1 intracellular region are in progress as well.

A recombinant protein encoding the transmembrane segment and longer constructs encoding both the transmembrane and the intracellular regions, thus allowing for full insertion in lipid bilayers, as found in bicelles or nanodiscs, will provide a more accurate model system for the biochemical and biophysical studies of the membrane/cytosol interface. On the other hand, the intracellular region of Jagged-1 as a nucleocytoplasmic protein and a signaling fragment is also an attractive target, and in-cell NMR experiments can be devised to detect folding upon binding to different potential partner proteins co-expressed in the same cell.

4. MATERIALS AND METHODS

4.1 Production of the recombinant Jagged-1 intracellular region

Two different constructs of Jagged-1 intracellular region were produced. The first, J1_ic, is the 125 residue long protein corresponding to the cytoplasmic region of human Jagged-1 (residues 1094-1218 of JAG1_human):
RKRRKPGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVPIKDYENKNSK
MSKIRTHNSEVEEDDMDKHQQKARFAKQPAYTLVDREEKPPNGTPTKHPNW
TNKQDNRDLESAQSLNRMEYIV. The longer one, J1_tmic, differs from J1_ic by the additional 8 residue long hydrophobic fragment at its N-terminus. J1_tmic corresponds to the part of Jagged-1 that is supposed to be released into the cytoplasm upon regulated intramembrane cleavage, which occurs between L1085 and V1086 [51]. J1_tmic (residues 1086-1218 of JAG1_human) has the following amino acid sequence (the additional transmembrane segment is in bold and underlined): **VTAFYWAL**RKRRKPGSHTHSASEDNTTNNVR
EQLNQIKNPIEKHGANTVPIKDYENKNSKMSKIRTHNSEVEEDDMDKHQQKA
RFAKQPAYTLVDREEKPPNGTPTKHPNWTNKQDNRDLESAQSLNRMEYIV.
To avoid possible cross-linking, C1092 was mutated to alanine (highlighted in grey).

4.1.1 Gene synthesis of J1_ic

The synthetic gene encoding the cytoplasmic region of human Jagged-1 (JAG1_ic, SWISS-PROT code: JAG1_human, residues 1094-1218) was obtained in a two-step PCR (assembly and amplification). The gene was assembled by PCR from a set of oligonucleotides designed by DNAWorks [123], as described hereafter. The amino acid sequence of JAG1_ic (125 residues, plus an additional stop codon, X), was backtranslated using the E. coli Class II codon usage and the resulting DNA sequence was split into 16 partially overlapping oligonucleotides by DNAWorks v2.3. The input parameters for DNAWorks were: maximal oligonucleotide size, 40 bases; annealing temperature (T_m), 64°C; codon frequency threshold, 100%; default values were used for oligonucleotide concentration (2.00E-7 M), Na^+ concentration (50 mM) and Mg^{2+} concentration (2.0 mM). The synthetic oligonucleotides were purchased from Sigma-Genosys (0.05 μmol scale, desalted), dissolved in H_2O and mixed in equimolar amounts. The obtained stock solution, containing 100 mM of each

oligonucleotide, was further used for PCR. Assembly PCR was carried out on a Perkin Elmer PCR Amp 2400 using the Pfu polymerase (Promega), 20 nM of each oligonucleotide, and the following conditions: 5 min at 95°C during which Pfu was added (hot start), 25 cycles of amplification (30 s denaturation at 95°C, 30 s annealing at 58°C, 90 s elongation at 72°C), 10 min at 72°C for the final elongation.

Amplification of the assembled gene was achieved using the following forward and reverse primers (Sigma-Genosys, 0.05 µmol scale, purified by polyacrylamide gel electrophoresis): 5'-TAA TAG TAG CAT ATG AAA **CAC CAT CAC CAT CAC CAT** CGT AAA CGT CGT AAA CCG GGT AGC and 5'-TAG TAG GGA TCC **TCA TTA** AAC GAT GTA TTC CAT ACG GTT CAG GCT. The forward primer contains a *Nde* I restriction site (underlined) encoding the start methionine, a AAA triplet encoding lysine, and a six histidine tag (in bold). The reverse primer contains a *Bam*HI restriction site (underlined) and two consecutive stop codons (in bold). Amplification PCR was carried out using an aliquot of the assembly PCR mixture as template, the Pfu polymerase (Promega) and the following conditions: 5 min at 95°C during which Pfu was added (hot start), 25 cycles of amplification (30 s denaturation at 95°C, 30 s annealing at 62°C, 90 s elongation at 72°C), 10 min at 72°C for the final elongation. PCR products were analyzed by agarose gel electrophoresis (1% agar gel in TBE buffer), stained with EtBr, and visualized by UV.

The assembled and amplified synthetic gene, resulting in a sharp band of the correct size in the agarose gel, was purified, digested by *Nde*I/*Bam*HI, and directionally cloned into the pET-11a vector (Novagen). Positive clones were obtained by transformation in DH5α *E. coli* cells grown on LB plates in the presence of 50 µg/mL ampicillin. Plasmid constructs containing the synthetic gene were sequenced and the correct one was used for transformation of BL21(DE3) *E. coli* (Novagen). This construct encodes the J1_ic protein with the cleavable N-terminal his-tag (6 histidines).

4.1.2 Preparation of the J1_tmic gene construct

The DNA encoding J1_tmic (residues 1086–1218 of JAG1_human) was amplified by PCR from the above template plasmid containing the codon-optimized synthetic gene encoding J1_ic (residues 1094–1218). The following forward and

reverse primers [Sigma-Genosys (Cambridge, UK), purified by polyacrylamide gel electrophoresis] were used: 5′-TAA TAT TAG CAT ATG **GTG ACC GCT TTC TAT TGG GCG CTG** CGT AAA CGT CGT AAA CCG GGT AGC- 3′ and 5′-TAG TAG GGA TCC *TCA TTA* AAC GAT GTA TTC CAT ACG GTT CAG GCT-3′. The forward primer contains a *Nde*I restriction site encoding the start methionine (underlined) and 8 residues belonging to the putative transmembrane region (in bold). The reverse primer contains a *Bam*HI restriction site (underlined) and a double stop codon (in italics). The PCR product was purified, digested with *Nde*I and *Bam*HI and directionally cloned into a pET-11a vector (Novagen, Darmstadt, Germany), as described for the J1_ic construct.

4.1.3 J1_ic expression and purification

Typically, 100 mL of LB containing 50 μg/mL ampicillin were inoculated with the transformed BL21(DE3) clone, cells were grown at 37°C to a density of ~1 OD units, and induced with IPTG (1 mM) for 3 h. Cells were harvested by centrifugation and pellets frozen at -80°C. Cells were resuspended in the lysis buffer (20 mM sodium phosphate buffer, 0.5 M NaCl, 50 mM Chaps, 2% Tween 20, 1 mM DTT,1 10 mM imidazole, pH 7.4, containing one protease inhibitor cocktail tablet (Roche)) and sonicated.

After centrifugation, the supernatant was loaded on a Ni^{2+}–Sepharose HisTrap HP column (1 mL, Amersham Biosciences), the column washed with 20 mM sodium phosphate buffer, 0.5 M NaCl, 1 mM DTT, 10 mM imidazole, pH 7.4 and the protein eluted with a 10–500 mM imidazole gradient. All purification steps were carried out on a ÄKTA (Amersham Pharmacia) chromatography system (FPLC) controlled by UNICORN™ software and equipped with a Frac-950 fraction collector. Fractions were monitored by UV absorbance at 280 and 214 nm, conductivity, and pH.

Fractions containing the protein were analyzed by SDS-PAGE and LC-MS. The LC-MS apparatus was composed of a Gilson HPLC system coupled to a ESI-MS single quadrupole mass spectrometer (Perkin Elmer API-150EX). Samples were eluted with a 0-100 % linear gradient going from 0.1% TFA in water to 0.1% TFA in MeCN on a Zorbax SB300 C18 column (5 μm, 4.6x150 mm, Agilent Technologies, Inc) and analyzed in positive ion mode. The molecular mass of the protein was

calculated by deconvolution of the multicharge ion spectra using the BioMultiview software (Applied Biosystems).

The crude J1_ic material, obtained from the first purification step, was further purified by RP-HPLC on a Zorbax 300SB-CN column (9.4x250 mm, 5 μm, Agilent) using a 0–50% gradient of 0.1% TFA in H2O and 0.1% TFA in acetonitrile. Fractions were analyzed by SDS-PAGE and LC-MS. Samples containing J1_ic were frozen in liquid nitrogen (-196°C) and dried under vacuum in the presence of anhydrous potassium carbonate.

The N-terminal His_6-tag was enzymatically removed using a recombinant dipeptidyl peptidase I cloned from rat liver, and carrying a C-terminal His-tag (TAGzyme, Qiagen) using the protocol suggested by the supplier. The protein (1 mg) was dissolved in 500 μL of TAGzyme buffer (20 mM sodium phosphate, 150 mM NaCl, pH 7.0). TAGzyme (50 mU) was activated by the addition of cysteamine and added to the protein solution. The digestion mix was incubated for 1 h at 37°C and analyzed by SDS-PAGE and LC-MS. After completion of the His_6-tag removal, the digestion mix was loaded on a Ni^{2+}-Sepharose HisTrap HP column (1 mL, Amersham Biosciences), and eluted with a 10-500 mM imidazole gradient as previously described. The protein was further purified by RP-HPLC on a Zorbax 300SB-CN column, in the conditions described above. The identity and purity of the protein were checked by SDS-PAGE and LC-MS. The protein sample was freeze-dried and used for spectroscopical and other studies.

4.1.4 J1_ic purification in native conditions

The His_6-tagged J1_ic was purified on a HisTrap column as described above, desalted by gel-filtration on a 26 x 60 mm HiPrep Sephadex G-25 column (Amersham) (running buffer: 5 mM sodium phosphate, 1 mM DTT, pH 7.4) and further purified by ion-exchange chromatography on a 8 x 75 mm SP column (Shodex SP-825) using a 0-100% gradient from buffer A (5 mM sodium phosphate, 1 mM DTT, pH 7.4) to buffer B (5 mM sodium phosphate, 1 M NaCl, 1 mM DTT, pH 7.4). Fractions were concentrated and desalted (final buffer: 5 mM sodium phosphate, pH 7.2) by ultrafiltration on a Centricon 3000 (Amicon).

4.1.5 J1_tmic expression and purification

Bacteria expressing J1_tmic were grown in the same conditions, and cell harvesting was carried out using the same protocol described for J1_ic. Although J1_tmic does not contain a N-terminal his-tag, the purification protocol was the same as for J1_ic (HisTrap column followed by RP-HPLC).

For the preparation of the ^{15}N-labeled protein, cells were grown in M9 minimal medium (6 g·L^{-1} Na_2HPO_4, 3 g·L^{-1} KH_2PO_4, 0.5 g·L^{-1} NaCl, 0.12 g·L^{-1} $MgSO_4$, 0.01·L^{-1} $CaCl_2$, 0.5 g·L^{-1} $^{15}NH_4Cl$, 5 g·L^{-1} d-glucose) supplemented with 1.7 g·L^{-1} Yeast Nitrogen Base without amino acids and ammonium sulfate (Difco, Sparks, MD, USA) and containing 100 μg·L^{-1} ampicillin. Expression and purification of the labeled protein were carried out as described above for the unlabeled J1_tmic. The purified proteins were analyzed by LC-MS to confirm their identity, purity, and isotope incorporation.

4.2 Peptide synthesis

The following synthetic peptides corresponding to the C-terminal region of J1_tmic; two 6 residue long peptides (1213–1218 of JAG1_human) and five 24 residue long peptides (1195-1218 of JAG1_human) were prepared:

J1C6: RMEYIV

J1C6pY: RME**Y**IV

J1C24: Ac-NWTNKQDNRDLESAQ**S**LNRMEYIV

J1C24pS: Ac-NWTNKQDNRDLESAQ**S**LNRMEYIV

J1C24pSpY: Ac-NWTNKQDNRDLESAQ**S**LNRME**Y**IV

J1C24RQ: Ac-NWTNKQDNRDLESAQSLN***Q***MEYIV

All J1C24 peptides were acetylated at the N-terminus. Phosphorylated residues are underlined in bold while the mutated residue is in italic and bold (R1213Q in J1C24RQ).

Peptides were prepared by standard solid phase Fmoc methods on a NovaSyn TGT (Novabiochem) resin using a home-built automatic synthesizer based on a Gilson Aspec XL SPE. All amino acids were coupled using a 4 molar excess with respect to the resin load (AA/HCTU/DIPEA = 1/1/2), except phosphorylated amino acids where a double amount of DIPEA was used. After the synthesis, the amino terminal group was acetylated with 20% Ac_2O and 1.5 eq DIPEA for 30 min in DMF. The peptide-

resin was cleaved/deprotected in a mixture of DODT/H_2O/Thioanisole/Methyl Ethyl sulphide/Phenol/Hydroiodic acid/TFA 8/2.5/3/2/3/1.5/80 (v/v) for 4 h. Crude peptides were purified by semipreparative RP-HPLC on a Zorbax 300SB-C18 column (9.4 × 250 mm, 5 μm, Agilent) and freeze-dried. The identity of the peptides was checked by LC-MS and the purity (>95%) estimated from RP-HPLC.

For SPR measurements, J1C24, J1C24pS, J1C24pSpY and J1C24RQ were biotinylated through coupling of biotin-succinimide to the side chain amino group of K1199, followed by purification by RP-HPLC.

4.3 Production of the recombinant AF6 PDZ domain

4.3.1 Expression and purification

The Afadin PDZ domain (AF6_PDZ, residues 987-1078 of AFAD_HUMAN, isoform 5) was amplified by PCR from a cDNA clone (IMAGp998L1313190Q, ImaGenes GmBH, Germany). The following forward and reverse primers [Sigma-Genosys (Cambridge, UK), purified by polyacrylamide gel electrophoresis] were used: 5′-TAA TAA TAA CAT ATG AAA **CAC CAC CAC CAC CAT CAC** AAA GAA CCT GAA ATA ATC ACT GTG ACC CTA - 3′ and 5′-AAA TAA GGA TCC **TCA TTA** GGC ACC CTG CTT TGC TAC TTC CAG- 3′. The forward primer contains a *Nde* I restriction site (underlined) encoding the start methionine, a AAA triplet encoding lysine, and a six histidine tag (in bold). The reverse primer contains a *Bam*HI restriction site (underlined) and two consecutive stop codons (in bold).

The PCR product, containing a cleavable N-terminal His_6-tag, was cloned into a pET-11a vector (Novagen) and used to transform BL21(DE3) *E. coli* (Novagen) cells for expression. Unlabeled AF6_PDZ was prepared by growing bacteria in LB medium at 37°C to a density of ~1 OD units and inducing with 1 mM IPTG for 3 h. Cells were harvested by centrifugation, resuspended in the lysis buffer (20 mM sodium phosphate buffer, 0.5 M NaCl, 50 mM CHAPS, 2% Tween 20, 1 mM dl-dithiothreitol, 10 mM imidazole, 0.5 mM EDTA, pH 7.4, containing one protease inhibitor cocktail tablet (Roche)) and sonicated. After centrifugation, the supernatant was loaded on a Ni^{2+} Sepharose HisTrap HP column (5 mL, Amersham Biosciences), the column washed with a 20 mM sodium phosphate buffer containing 0.5 M NaCl, 1 mM dl-dithiothreitol, 10 mM imidazole, pH 7.4 and the protein was eluted with a 10–500 mM imidazole gradient. Fractions containing the pure protein were desalted using

a fast-desalting column HiPrep 26/10 (GE Healthcare) and freeze-dried. The N-terminal His_6-tag was removed enzymatically using a recombinant dipeptidyl peptidase, as described above for J1_ic. After completion of the His_6-tag removal, the digestion mix was purified on the HisTrap column and the pure protein desalted as described above.

The uniformly ^{15}N–labeled and ^{13}C,^{15}N–labeled AF6_PDZ were prepared growing *E. coli* cells in M9 minimum medium (described above) containing $^{15}NH_4Cl$ (0.5 mg/mL) and/or $^{13}C_6$-glucose (2.5 mg/mL) at 37°C to a density of ~1 OD units. After induction with 1 mM IPTG for 4.5 h, cells were harvested and the protein purified as described above. The identity, purity, and isotope incorporation of all proteins were verified by LC-MS. The amino acid sequence of the produced AF6_PDZ is:

*M*K<u>HHHHHH</u>*KEPEIITVTLKKQNGMGLSIVAAKGAGQDKLGIYVKSVVKGGA ADVDGRLAAGDQLLSVDGRSLVGLSQERAAELMTRTSSVVTLEVAKQGA.

The start methionie is in italic, N-terminal his-tag is in bold and the first residue K987 is marked with an asterisk.

4.4 Determination of protein and peptide concentrations

All protein/peptide concentrations were determined by UV absorbance at 280 nm (Abs_{280}), using Lambert-Beer's law ($Abs_{280} = \varepsilon \cdot c \cdot d$), where ε is the extinction coefficient at 280 nm, c is the protein molar concentration and d is a path-length of the cuvette. For each protein/peptide, the extinction coefficient was calculated using the ProtParam tool (http://expasy.ch/protparam) of the Swiss Institute of Bioinformatics:

ε_{280} (J1_ic) = 9530 M^{-1} cm^{-1}

ε_{280} (J1_tmic) = 16500 M^{-1} cm^{-1}

ε_{280} (J1C6) = 1490 M^{-1} cm^{-1}

ε_{280} (J1C24, J1C24pS, J1C24pSpY and J1C24RQ) = 6990 M^{-1} cm^{-1}

ε_{280} (AF6_PDZ; with or without his-tag) = 1490 M^{-1} cm^{-1}

For J1C6 and AF6_PDZ, which contain only one tyrosine and no tryptophan residues, concentration calculated from the extinction coefficient may be affected by a 10% error. For them, an additional amino acid analysis was carried out to check the validity of the estimated ε_{280} and to calibrate the absorbance readings (real concentration is 85% of the measured one).

4.5 Size exclusion chromatography

The freeze-dried protein was dissolved in the elution buffer (Tris/HCl 50 mM, 100 mM KCl, pH 7.4), loaded onto a Sephacryl S-200 column (GE Healthcare) and eluted in the same elution buffer. The apparent molecular mass of J1_tmic was estimated from a calibration carried out with the following molecular standards: lactate dehydrogenase (147 kDa), bovine serum albumin (67 kDa), carbonic anhydrase (29 kDa) and horse myoglobin (17 kDa). Stokes radii of native (RSN) and fully unfolded (RSU) proteins of known molecular mass (m) were determined according to the equations: log(RSN)¼)(0.254 ± 0.002) + (0.369 ± 0.001) log(m), and log(RSU)¼) (0.543 ± 0.004) + (0.502 ± 0.001) log(m) [131].

4.6 Limited proteolysis

The freeze-dried material (J1_ic) (~150 µg) was dissolved in Tris buffer (0.1 M, pH 7.4) containing $CaCl_2$ (10 mM) and subjected to proteolysis at room temperature by addition of chymotrypsin (sequencing grade, Roche, 1:100 w/w) dissolved in Tris buffer [127]. Aliquots of the reaction mixture were quenched with HCOOH, and analyzed by LC–MS on a Gilson HPLC system coupled to an ESI-MS single quadrupole mass spectrometer (Applied Biosystems API-150EX). Separation of peptide fragments was achieved with a C18 column (150 x 1 mm, 5 µm, Phenomenex) using a 0–40% gradient of 0.1% HCOOH in H2O and 0.1% HCOOH in acetonitrile at a flow of 35 µL/min. Comparison between the fragments expected from chymotrypsin cleavage and experimentally found fragments was carried out using the BioMultiView software (Applied Biosystems). The identity of candidate peptide fragments was confirmed by deconvolution of the multicharge ion spectra.

4.7 Preparation of phospholipid vesicles and micelles

The synthetic phospholipids DMPG (1,2-dimyristoyl-sn-glycero-3-[phospho-rac-(1-glycerol)] sodium salt), DMPS (1,2-dimyristoyl-sn-glycero-3-[phospho-L-serine] sodium salt) or DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine); (Avanti, Alabaster, AL, USA), were dissolved in $CHCl_3$/CH_3OH (2:1, v/v) in round-bottomed flasks and the solvent evaporated to obtain a thin lipid film. After drying under vacuum to remove the residual solvent, lipids were hydrated in 5 mM Tris/HCl buffer, pH 7.4, to obtain a 10 mM lipid suspension which was sonicated to clarity at

37°C in a high intensity bath sonicator (Branson 3200, Branson Sonic Power Co., Danbury, CT, USA).

Micelles were prepared directly by dissolving SDS (Sodium Dodecyl Sulphate) or lysophospholipids (LMPG, 1-Myristoyl-2-Hydroxy-*sn*-Glycero-3-[Phospho-*rac*-(1-glycerol)] sodium salt and LMPC, 1-Myristoyl-2-Hydroxy-*sn*-Glycero-3-Phosphocholine) (Avanti Polar Lipids Inc., Alabaster, AL, USA) in buffer solutions.

4.8 Circular dichroism

Samples for CD spectroscopy were prepared dissolving the freeze-dried protein in 5 mM Tris⁄HCl buffer, pH 7.4. The protein concentrations used were in the range 7 μM-12 μM. CD spectra were recorded at 25°C or 37°C on a Jasco J-810 spectropolarimeter (JASCO International Co., Tokyo, Japan) using jacketed quartz cuvettes of 1 mm pathlength. Five scans were acquired for each spectrum in the range 190–250 nm at a scan rate of 20 $nm \cdot min^{-1}$. Mean residue ellipticity (MRE = $deg \cdot cm^2 \cdot dmol^{-1} \cdot residue^{-1}$) was calculated from the baseline-corrected spectrum. A quantitative estimation of secondary structure content was carried out using SELCON3, CONTINLL, and CDSSTR, all run from the DichroWeb server (www.cryst.bbk.ac.uk/cdweb/html/home/html) [124]. Helical content was also estimated from the mean residue ellipticity at 222 nm according to the formula $[\alpha] = -100 \cdot MRE_{222}/40000 \cdot (1-2.57/N)$, where N is the number of peptide bonds [236].

4.9 Fluorescence spectroscopy

Samples prepared for CD were also used for fluorescence spectroscopy. Fluorescence spectra were recorded at 298 K on a Jobin-Yvon FluoroMax-3 spectrofluorometer equipped with a Peltier temperature control apparatus using 1x0.2 cm pathlength quartz cuvettes. Excitation was set at 295 nm to minimize the contribution from Tyrosine and Phenylalanine residues to the overall fluorescence. Spectra were recorded between 300 and 450 nm. Fluorescence anisotropy was measured at the maximum of emission (355 nm) using the same excitation wavelength. Measurements were corrected for the background and averaged.

4.10 NMR spectroscopy

4.10.1 NMR of the Jagged-1 intracellular region

NMR of unlabeled J1_ic. The sample for NMR spectroscopy was prepared dissolving the freeze-dried material in H_2O/D_2O (90/10, v/v) containing 2 mM EDTA-d_{16} and 20 μM DSS and adjusting the pH to 6.6 with small aliquots of 0.1 N KOH, for a final sample concentration of ~0.9 mM. Spectra were recorded at 298 K on a Bruker Avance 500 operating at a ^1H frequency of 500.13 MHz and equipped with a three-axis gradient probe. A 2D TOCSY spectrum was acquired using a 50 ms mixing time, with solvent suppression obtained by excitation sculpting [237]. The acquisition was performed over a spectral width of 7002.8 Hz in both dimensions, with matrix size of 2048 points in t_2 and 512 points in t_1. Data were transformed using X-WinNMR (Bruker). Chemical shifts were referenced to internal DSS.

NMR of ^{15}N-labeled J1_tmic. Protein samples for NMR spectroscopy were prepared dissolving the freeze-dried material in H2O⁄D2O (90:10, v⁄v) and adjusting the pH to 7.0 with small aliquots of 0.1 N NaOH, for a final protein concentration of ~0.5 mM. The sample containing SDS was prepared by dissolving solid SDS sodium salt in the NMR sample, for a final SDS concentration of 50 mM. Additional spectra were recorded at pH 5.5. Spectra were recorded at 303 K on a Bruker spectrometer (Bruker Biospin, Rheinstetten, Germany) operating at a ^1H frequency of 600.13 MHz and equipped with a $^1H/^{13}C/^{15}N$ triple resonance Z-axis gradient probe. Transmitter frequencies in the ^1H and ^{15}N dimensions were set on the water line and at 118.0 p.p.m., respectively. HSQC and HSQC-TOCSY experiments were carried out in phase sensitive mode using echo⁄anti-echo-TPPI gradient selection and ^{15}N decoupling during acquisition. HSQC spectra were acquired with 1 K complex points, 256 t_1 experiments, 32 scans per increment, over a spectral width of 13 and 28 p.p.m. in the ^1H and ^{15}N dimensions, respectively. HSQC-TOCSY spectra were acquired with the same parameters, but with 128 scans per t_1 increment and a 40 ms DIPSI mixing time. Data were transformed using X-WinNMR (Bruker) and analyzed using CARA (http://www.nmr.ch). ^1H chemical shifts were referenced to internal DSS (8 μM).

For in-cell NMR experiments [238], [239], 200 mL of *E. coli* culture was grown in M9 medium containing $^{15}NH_4Cl$ as the only nitrogen source, as described above. The culture was split; in one sample expression was induced with isopropyl

thio-β-d-galactoside, and the other was used as control. Cells were centrifuged at ~500 *g* in a Sorvall RC5B centrifuge (Sorvall Instruments Inc., Newton, CT, USA) using a GSA rotor. The supernatant was discarded, and the pellet was gently resuspended in 50 mL of cold PBS (phosphate buffered saline; 10 $g \cdot L^{-1}$ NaCl, 0.25 $g \cdot L^{-1}$ KCl, 0.25 $g \cdot L^{-1}$ KH_2PO_4, 3.6 $g \cdot L^{-1}$ $Na_2HPO_4 \cdot 12H_2O$, pH 7.2). After an additional centrifugation step, the pellet was gently resuspended in 500 μL PBS, D_2O (55 μL) was added, and a standard NMR tube was filled with the *E. coli* slurry. After NMR analysis, the slurry was recovered from the NMR tube, centrifuged for 2 min at ~14000 *g* in a Millipore MC-13 microcentrifuge (Amicon Bioseparations Inc., Beverly, MA, USA) and the clear supernatant subjected to further NMR analysis. HSQC spectra on the induced sample, on the control sample, and on the supernatant were acquired in identical conditions at 303 K with 1 K complex points, 128 t_1 experiments, 32 scans per increment, over a spectral width of 13 and 26 p.p.m. in the ^1H and ^{15}N dimensions, respectively, for a total experiment time of ~1 h for each HSQC. A sample of freeze-dried, purified protein dissolved in PBS was used to acquire a reference spectrum.

NMR of ^{15}N-labeled J1_tmic bound to lysophospolipid micelles. NMR data were acquired at 298 K on a Bruker Avance 500 spectrometer, operating at a ^1H frequency of 500.13 MHz, equipped with a TXI triple resonance cryoprobe. HSQC spectra were acquired placing the carrier frequency on the water resonance in the ^1H dimension and at 122 ppm in the ^{15}N dimension, with a spectral width of 16 and 40 ppm in the ^1H and ^{15}N dimensions, respectively. Data were processed using XwinNMR and analyzed with CARA. Preliminary ^{15}N relaxation data were obtained using five delays: 0.010, 0.150, 0.540, 1.00, 2.50 s for T_1 experiments and 17, 68, 153, 204, 271 ms for T_2 experiments. Peak intensities were evaluated from peak height at the maximum, and the uncertainty on the peak intensity was set as $2.5*\sigma_{noise}$, where σ_{noise} is the standard deviation of the base plane noise in each spectrum. Signal decay was fitted with a two-parameter exponential function using a Levenberg-Marquardt non-linear least squares fitting routine, as implemented in Curvefit. Uncertainties in the fitted longitudinal (R_1) and transverse (R_2) relaxation rates were set to be equal to the error in the fitted parameters.

4.10.2 NMR of AF6_PDZ

All NMR samples were prepared dissolving the freeze-dried proteins or peptides in phosphate buffer (50 mM Na-phosphate, 50 mM KCl, 1 mM EDTA, 10% v/v D_2O, pH 5.8). NMR experiments were carried out at 298 K on Bruker AVANCE 600 and 800 MHz spectrometers equipped with 5-mm triple-resonance cryoprobes. NMR data were processed using the Azara suite of programs (provided by Wayne Boucher and the Department of Biochemistry, University of Cambridge, UK) or X-WinNMR (Bruker). Spectra were analyzed using the CCPN program ANALYSIS (http://www.ccpn.ac.uk) and CARA (http://www.nmr.ch).

Resonance assignments of AF6_PDZ. For assignment purposes, the AF6_PDZ protein concentration was 0.5 mM and 0.8 mM for the $^{13}C,^{15}N$-labeled and ^{15}N-labeled samples, respectively. Backbone resonance assignments of AF6_PDZ were obtained from TOCSY-HSQC and NOESY-HSQC spectra acquired on the ^{15}N-labeled protein and confirmed by CBCANH and CBCA(CO)NH spectra acquired on the $^{13}C,^{15}N$-labeled AF6_PDZ. Assignments of additional methyl groups were obtained from a 3D HCCH-TOCSY spectrum using the Cα and Cβ resonances assigned by above CBCANH and CBCA(CO)NH spectra. All above spectra were acquired in a phase-sensitive mode using echo/anti-echo-TPPI gradient selection and ^{15}N decoupling during acquisition.

Chemical shift mapping. For chemical shift perturbation experiments, aliquots of AF6_PDZ (1 mM) and peptide stock solutions (1-5 mM) were mixed to give different samples with the wanted peptide:PDZ ratio and a constant final concentration of 0.1 and 0.2 mM for the $^{13}C,^{15}N$-labeled and ^{15}N-labeled AF6_PDZ, respectively. Typically, NMR spectra (^{1}H-^{15}N-HSQC and ^{1}H-^{13}C-HSQC) were recorded at titration points corresponding to a peptide:PDZ ratio of 1, 2, 4, 6, 8, and 10. ^{1}H-^{15}N-HSQC and ^{1}H-^{13}C-HSQC experiments were carried out in phase-sensitive mode with decoupling during acquisition. Both HSQC spectra were acquired placing the carrier frequency on the water resonance in the ^{1}H dimension and at 117.5 ppm in the ^{15}N dimension or 30.9 ppm in the ^{13}C dimension. Spectral width in the ^{1}H, ^{15}N, and ^{13}C dimensions were 13, 40 and 32 ppm, respectively.

Chemical shift mapping of the $^{13}C,^{15}N$-labeled AF6_PDZ titrated with J1_tmic C-terminal peptides was based on ^{1}H-^{15}N-HSQC spectra [140], [205]. Combined chemical shift perturbation (CSP) was calculated using the following formula:

$CSP = sqrt [(\Delta\delta_H)^2 + \alpha_N(\Delta\delta_N)^2]$ using a $\Delta\delta_N$ scaling factor (α_N) of 0.102.
CSP values were plotted onto the first model of the solution structure of AF6_PDZ domain complexed with a LFSTEV peptide (PDB: 2AIN) [240].

Dissociation constants (K_d) were calculated plotting normalized CSP values against the molar ratios X (peptide:PDZ) and fitting by nonlinear regression analysis according to the formula [241]:

$CSP/CSP_{max} = 0.5*\{1+X+K_d/[P] - [(1 + X + K_d/[P])^2 - 4X]^{1/2}\}$

where [P] is a total concentrations of AF6_PDZ. For selected alanine, valine, and leucine residues belonging to the GMGL motif, to the βB-strand and to αA- and αB-helices additional CSP values were extracted from the methyl region of ^1H-^{13}C-HSQC spectra and K_d values calculated as described above.

NMR of J1_tmic. ^{15}N-labeled J1_tmic (100 μM) was dissolved in phosphate buffer (50 mM Na-phosphate, 50 mM KCl, 1 mM EDTA, 10% v/v D_2O, pH 5.8) and ^{15}N-HSQC spectra were acquired for the J1_tmic alone and in the presence of unlabeled AF6_PDZ (600 μM). Recorded ^1H-^{15}N-HSQC spectra were processed and analyzed as described above for AF6_PDZ.

4.11 Surface Plasmon Resonance

Interactions of AF6_PDZ with selected peptides were assessed using a Biacore T100 apparatus (Biacore, General Electrics Healthcare). Preliminary experiments were carried out by immobilising AF6_PDZ on a CM5 sensor chip and injecting the different peptides. Dissociations constants, calculated using this method, were however higher than expected and not very reproducible, probably due to the limited solubility of some of the peptides. This problem was overcome using the reciprocal immobilization strategy.

When possible, the peptides were biotinylated and immobilized on the streptavidin-coated sensor chip (SA sensor chip) through non-covalent interactions. This approach gave better results, but suffered from a significant degree of non-specific binding to the control flow cell without the peptide, possibly because of charge-charge interactions between streptavidin (calculated pI = 6.1) and AF6_PDZ (calculated pI = 9.4). The following protocol, employing avidin (calculated pI = 9.7) instead of streptavidin, gave the best results. The CM5 sensor chip was primed with 1 x HBS-EP running buffer freshly diluted from a 10x stock solution (0.1 M Hepes, 1.5

M NaCl, 30 mM EDTA, 0.5 % surfactant P20). ImmunoPure® Avidin from hen egg white (Pierce) was immobilized on the surface of two flow cells through the amine coupling method using 0.05 M Nhydroxysuccinimide, 0.2 M N-ethyl-N'-(3-diethylaminopropyl) carbodiimide and ethanolamine (pH 8.5) (Biacore, General Electrics Healthcare) following the protocols recommended by the producer. The avidin immobilization level was between 2000-3000 response units (RU) for all peptides studied. Biotinylated peptides were then injected over the test cell, while the reference cell consisted of only avidin and served to subtract nonspecific binding of AF6_PDZ. Final levels of immobilized peptides were 330, 444, 349 and 338 RU for J1C24, J1C24pS, J1C24pSpY and J1C24RQ, respectively. AF6_PDZ was prepared as a stock solution in 1 x HBS-EP buffer. It was injected across the sensor chip at a flow rate 5 μL/min for 1 min and dissociation was monitored for 60 s.

The surface was regenerated between injections with 30 s pulse of 0.5 % SDS and stabilized with a flow of running buffer at the same flow rate for 150 s. All experiments were performed at 25°C. Obtained sensorgrams were evaluated with the Biacore T100 Evaluation Software. Dissociation constants (K_d) were estimated from the binding levels at the end of the injection, which for most injections of AF6_PDZ reached steady-state levels. The K_d values were determined by fitting the following equation to the obtained data: $R_{eq} = C*R_{max}/(K_d+C)$ where R_{eq} is steady state binding level, C is the concentration of AF6_PDZ and R_{max} is the AF6_PDZ binding capacity of the surface. The K_ds were determined from 3 (J1C24, J1C24pS), 4 (J1C24pSpY) or 7 (J1C24RQ) independent titrations.

5. BIOPHYSICAL METHODS

5.1 Circular Dichroism

Circular Dichroism (CD) is a type of absorption spectroscopy that allows the investigation of the secondary and tertiary structure of proteins and peptides in solution. The CD phenomenon arises from the interaction between polarized light and chromophores - functional groups giving rise to electronic transitions - that are in an asymmetrical environment. CD can in principle be measured at any λ of the electromagnetic spectrum, but it is most commonly measured in the UV-visible region of the spectrum, associated with electronic transitions. When a chromophore is hit with right- or left-circularly polarized light, a differential absorption of the two beams is measured and plotted as a function of the wavelength.

CD of proteins is largely dictated by the strong transitions arising from the peptide bond chromophore in the far-UV (240–180 nm) and, to a lesser extent, by the weaker transitions arising from aromatic residues (Phe, Tyr, Trp) and disulfide bonds in the near-UV (320-260 nm) [242]. Two transitions in the peptide group (n-π* and π–π*) are responsible for the positive and negative bands observed in the spectra in the far-UV CD. The intensity and energy of these transitions depend on the dihedral angles adopted by the polypeptide chain. Hence, different types of regular secondary structure elements give rise to distinct types of far-UV CD spectra. Far-UV CD spectra of random coil polypeptides display a negative band at 195 nm and a weaker additional band at ~212 nm, which can appear either as a weakly positive band or as a negative shoulder. β-sheet polypeptides display a positive band at 196 nm and a negative band at 218 nm. The CD spectrum of an α-helix is characterized by a strong positive band at 192 nm and by two typical negative bands, the first at 222 nm and the second at 209 nm. The estimation of the protein secondary structure from its far-UV CD spectrum is based on the fact that the absorbance is additive, and thus the spectrum can be represented by a linear combination of individual reference spectra of pure secondary structure [243]. These "pure secondary structure" spectra can in turn be derived either from the analysis of CD spectra obtained for a dataset of proteins of known structure, or from CD spectra of synthetic polypeptides like poly-L-lysine in its α-helical, β-sheet, and random coil conformations. Several mathematical methods have been developed for the deconvolution of far-UV CD spectra: multilinear

regression, singular value decomposition, ridge regression, convex constraint analysis, neural network, and the self-consistent method. The output generally consists of a list of secondary structural fractions and a calculated spectrum, which can be compared to the experimental spectrum in order to evaluate the quality of the deconvolution. Results are quite accurate (>95%) for the α-helix, but less accurate in the estimation of other types of secondary structures, like β-sheets (60-90%) and turns (40-50%). Factors that can affect the estimation of secondary structure content include the wavelength range used in data acquisition, the accuracy in the determination of the protein concentration, and the presence of aromatic residues. In fact, the side chains of aromatic residues (the phenyl group of Phe, the phenolic group of Tyr, and the indole group of Trp) also have adsorption bands in the far-UV, which overlap with the contribution from the peptide bond chromophore.

Near-UV CD, on the other hand, can give information on the presence of tertiary structure in a protein, as they reflect the symmetry of the environment of aromatic residues. Each aromatic amino acid type has a specific spectrum in the near-UV CD: phenylalanine residues show an absorption band in the 250-270 nm region, while tyrosine and tryptophan in the 270-290 and 280-300 nm regions, respectively. Intense near-UV CD spectra with a fine structure are expected in proteins with highly ordered tertiary structure, as they are reflective of the asymmetric environment around aromatic residues. Conversely, the lack of ordered tertiary structure usually leads to less intense and less complex spectra in the near-UV region.

The principal advantages of CD derive from the speed and simplicity of the technique. Compared with X-ray crystallography and NMR, CD measurements can be carried out rapidly. Good quality spectra in the far- and near-UV can each be obtained within 30 min. Far-UV CD studies, in particular, require only small amounts of material (~1 mL of 1-10 μM solution) and since the technique is non-destructive, it is usually possible to recover most of the sample and perform multiple experiments. Caution should be taken in the choice of the buffer as, ideally, it should be transparent in the far-UV range. For this reason, low ionic strength phosphate is usually the buffer of choice. CD is best suited to monitor changes in secondary structure associated with binding to metals, micelles or liposomes, binding to other proteins or to DNA, or changes in temperature (thermal denaturation), pH, or solvent. The main limitation of

CD is that it gives only an estimate of the percentage of a particular secondary structure present, without telling where it is located in the sequence.

5.2 Fluorescence spectroscopy

Fluorescence spectroscopy is a technique widely used to investigate interactions and structural transformations of biological macromolecules [244]. It takes advantage of the light emission from molecules in electronically excited state. Absorption of light causes the excitation of an electron from a ground to an excited state. After its excitation, the electron rapidly relaxes from the higher to the lowest vibrational level of the excited electronic state. The decay to the ground state can occur by a number of mechanisms: internal conversion (heat), quenching (external conversion), by emission of a photon (fluorescence), or by intersystem crossing (phosphorescence). In fluorescence, the emission wavelength is significantly longer (lower energy) than in excitation. This red-shift is proportional to the energy difference between the higher and the lowest vibrational levels of the excited electronic state. In biochemistry, fluorophores can be divided in two classes, extrinsic and intrinsic. In proteins, the intrinsic fluorescent residues are restricted to phenylalanine, tyrosine, and tryptophan. The advantage compared to extrinsic probes is that the protein itself is not changed. The absorption maxima of Phe, Tyr, and Trp occur at 257, 274, and 280 nm, respectively, with emission spectra at 282, 303, and 350 nm. The absorption of Phe and Tyr residues is weaker by 28 and 4 times respectively, compared to that of Trp and the fluorescence quantum yields are lower. As a result, the fluorescence of proteins is usually dominated by the indole group of Trp. When a long excitation wavelength is used (~295 nm) the absorption, therefore the contribution to overall fluorescence of Tyr and Phe becomes negligible.

The Trp emission spectrum is very sensitive to its local environment. For example, the wavelength of emission maximum (λ_{max}) can very from ~350 nm for a completely solvent exposed Trp to 320 nm or less for a Trp buried in a hydrophobic environment. This means that the value of λ_{max} can give some basic information as to whether the given protein is compact under the given conditions. Moreover, the emission intensity (I_{max}) of tryptophan depends on the polarity of the surrounding solvent. The increase of the tryptophan I_{max} and the blue-shift of its λ_{max} is a consequence of the fluorophore's transfer from an aqueous to a non-polar

environment. The monitoring of I_{max} and λ_{max} can be used to confirm the conformational changes or the interaction involvement of the protein region in the tryptophan vicinity. However, the wavelength of the emitted light is usually a better indication of the environment of the fluorophore, because the intensity magnitude can vary in a somewhat unpredictable manner. For example, comparing to the unfolded state, the tryptophan I_{max} may be either increased or decreased by the protein folding.

As tryptophan is a relatively rare amino acid (at ~1% mol), many proteins contain only one or a few tryptophan residues which makes fluorescent spectra of proteins relatively simple. In proteins with more tryptophans, the emission spectrum is the sum of all individual spectra and appears as a single peak. Already in a protein containing two tryptophans, it is impossible to separate two spectra since the peak half-intensity width (50-60 nm) is larger than the maximum difference in λ_{max} (~30 nm) two tryptophans can show.

Furthermore, tryptophan displays a high anisotropy, the extent of which is sensitive to the extent of motion during the excited state. Samples exhibiting non-zero anisotropies display polarized emission upon excitation with polarized light. Experimentally, the sample is excited with vertically polarized light and two emission polarizers, one parallel to the excited light, the other perpendicular, measure the parallel ($I_{||}$) and the perpendicular ($I_{\perp}$) intensity. Anisotropy is defined as the difference between $I_{||}$ and $I_{\perp}$ normalized for the total emitted intensity ($I_{TOT} = I_x + I_y + I_y = I_{||} + 2I_{\perp}$):

$$r = (I_{||} - I_{\perp}) / (I_{||} + 2I_{\perp})$$

Anisotropy can be also expressed using the angle (θ) between the emission and excitation dipoles. For a population of molecules in a random, immobilized orientation and assuming that the absorption and emission transition moments are parallel, anisotropy is given by:

$$r = 1/2 \ (3\langle\cos^2\theta\rangle - 1)$$

The maximum of anisotropy ($r = 1$) is measured when the emission is completely polarized ($I_{\perp} = 0$) or ($\cos^2\theta = 1$). However, when polarized light interacts with a population of molecules, the probability of absorption depends on the orientation of the molecule, and from this probability distribution it is possible to calculate that the maximum value for $\langle\cos^2\theta\rangle$ is 3/5, which gives a maximum value for anisotropy of $r_0 = 0.4$, still assuming that absorption and emission moments are collinear. For dilute

protein solutions, anisotropy is mainly determined by the rotational motion of the fluorophore, which in turn depends on the size and shape of the macromolecule, and on the extent of aggregation. The effect of molecular tumbling on anisotropy is given by the Perrin equation:

$$r_0/r = 1 + t/t_s$$

where t_s is the rotational correlation time (assuming a spherical object) and t the fluorescence lifetime. The increase of the fluorophore tumbling velocity (the decrease of its rotational correlation time) results as a decrease of its anisotropy. Fluorescence anisotropy is commonly used for measuring the interactions between molecules. Upon binding of the "tryptophan protein" to the partner molecule, a more stable complex is formed, which tumbles more slowly and thus increases the measured anisotropy. The anisotropy increase is more evident if the partner molecule has a large MW, like a micelle or a liposome.

Although apparently trivial, the measurement of anisotropy can be affected by a series of experimental problems. Turbidity of protein samples can induce scattering of the incident light, which will reduce the measured anisotropy. A possible effect of turbidity on the sample can be evaluated through measurements at different concentrations. A more serious problem can arise from scattered light reaching the detector. As scattered light is strongly polarized, even a small amount can significantly alter the anisotropy measurement. Anisotropy values equal to or larger than 0.4 should be looked with suspicion and carefully evaluated. Additional problems can arise from misalignment of the polarizers and background signals.

5.3 Surface plasmon resonance

Surface plasmon resonance (SPR) is another commonly used technique for studying macromolecular interactions [245]. This technique provides a direct and rapid determination of association and dissociation rates of binding process, does not require any labeling of macromolecules, and only small amounts of sample is used in the assay. When polarized light passes through a prism on a sensor chip with a thin metal film on top, the light will be reflected by the metal film acting as a mirror. At a certain angle of incidence (resonance angle or SPR angle), the intensity of the reflected light passes through a minimum. This intensity dip is caused by the excitation of surface plasmons which can be considered as propagating electron

density waves occurring at the interface between metal (usually gold) and dielectric. The SPR angle is dependant upon the refractive indices of media at both side of the metal. While the refractive index at the prism side is not changing, the refractive index in the immediate vicinity of the metal surface (within a few hundred nanometers) will change when accumulated mass adsorb on it. Therefore, the optical conditions are changing and the shift of the SPR angle is suited to provide information on the kinetics of the adsorption on the surface.

In practice, a ligand of interest (ex. a protein_L) is immobilized on a dextran polymer, and a solution of analyte (ex. a protein_A) is flowed through a cell, one wall of which is composed of this polymer. If protein_A interacts with the immobilized ligand, it is retained on the polymer surface. Hence, the increased amount of protein near the polymer changes the SPR angle of impinging light. Since all proteins have the same refractive index and since there is a linear correlation between resonance angle shift and amount of absorbed protein on the surface, one can measure changes in protein concentration at the sensor chip due to the protein-protein binding. The shift of the resonance angle can be monitored over time and if the observed changes are due to a biomolecular interaction, the kinetics of the binding can be studied in real time. The SPR signal is expressed in response units (RU) (1RU corresponds to 1 pg of protein / mm^2) and is plotted against time (a sensorgram). Determination of a binding constant requires measurement of two parameters from the sensorgram. The k_{on} constant can be determined from the slope of the RU increase, which is measured as a function of time by passing a buffer containing the analyte over the immobilized ligand. Once the RU value is stabilized, the buffer lacking the analyte is injected on to the sensor and the non-specifically bound components are flushed off. Also in this step, dissociation of the analyte starts, enabling the determination of the dissociation rate (k_{off}). This procedure is repeated using different concentrations of the analyte. The signal baseline should be constant in all recorded sensorgrams and therefore the system requires the regeneration of the sensor chip surface after each measurement. The dissociation constant (K_d) is then calculated as a ratio k_{off} / k_{on}. An alternative determination of Kd can be made by using the steady-state RU values at different protein concentrations. In SPR experiments it is of vital importance to have the ligand properly anchored to the sensor surface, without any loss of its binding activity during regeneration and measurement steps [246] . Several SPR sensor chips have functional groups for covalent immobilization attached to the dextran which is attached to the

surface of the gold chip. One of the most commonly used sensor chips (CM5) is carboxymethylated dextran, which enables an easy covalent coupling of the ligand macromolecule via amino groups. The dextran layer enables high flexibility and, with little steric hindrance, immobilized molecules behave as in a solution [247]. A wide range of different sensor chips, with different kinds of immobilization methods, allows studies of various biological and biochemical interactions (protein–protein, protein–nucleic acid, protein–carbohydrate, protein–membrane, protein-drugs etc.).

A major problem encountered when studying biomolecular interactions through SPR may be represented by improductive coupling of the biomolecule to the chip, with consequent masking of the binding site, and non specific interactions between the matrix and the immobilized biomolecule or the flowing ligand. These problems often need to be addressed using different sensor chip types and different immobilization strategies in order to find the optimal experimental conditions.

5.4 NMR spectroscopy

5.4.1 Physical background

High resolution nuclear magnetic resonance (NMR) spectroscopy is currently the only available method for determining the three-dimensional structure of proteins in solution at atomic resolution. In addition, applying specific NMR techniques, it is also possible to study protein-ligand interactions, protein folding and protein dynamics [248]. The phenomenon of magnetic resonance results from the interaction of the magnetic moment of an atomic nucleus with an external magnetic field. The cause of this is the nuclear spin angular momentum, which is a particular quantum mechanical property of atom nuclei. In a static magnetic field the nuclear spin behaves like a magnetic dipole which precesses around it. This field causes the energy level to split into a favourable low-energy and an unfavourable high-energy level. For a nucleus with a spin of 1/2, as for example ^{1}H, ^{13}C, ^{15}N, ^{19}F or ^{31}P, these two energy levels correspond to the magnetic spin quantum numbers 1/2 and -1/2. The energy difference between two levels is proportional to the applied magnetic field (up to 20 Tesla in instruments of the latest generations) and the proportionality constant is known as gyromagnetic ratio, a property of individual nuclei. The spin of each NMR-active nucleus precesses at its own resonance frequency, which is dependent on the chemical nature of the nucleus and on its local environment. The NMR experiment

measures a large ensemble of spins derived from a population of molecules. The sum of the dipole moments of identical spins is called magnetization. At the equilibrium, the magnetization vector of each nucleus has only a longitudinal component (along the static field). This is because the spin precession phases are not correlated, meaning that for each vector of the dipole moment pointing in one direction of the transverse plane a corresponding vector can be found which points into the opposite direction. In this state there is no NMR signal, since only transverse magnetization is detectable in the receiver coil. If a radio-frequency field (RF pulse), with a frequency matching the energy difference of the two levels, is applied perpendicularly to the static magnetic field, transitions from the lower level to the upper are induced and resonance absorption can be observed. This creates a "coherence" state, in which phases of the spins are partially correlated, resulting in a non zero transverse magnetization and giving rise to a signal in the detection coil. The RF pulse contains many frequencies in a broad band around the fixed work frequency and thus excites the resonances of all spins in a sample at the same time. The emitted signal is a superposition of all excited frequencies and its evolution in time is recorded. The differences in resonance frequencies are referred to as chemical shifts and give the possibility to distinguish between nuclei. The chemical shift is the difference between resonance frequency of the nucleus and a standard, relative to the standard itself. Traditionally, in NMR spectra the chemical shift is not expressed as a frequency, but in ppm (part per million):

$\delta = [(\nu_1 - \nu_{ref}) / \nu_{ref}] * 10^6$

where ν_1 is the resonance frequency of the observed nucleus and ν_{ref} is the resonance frequency of the nucleus of a reference compound. For ^1H NMR the reference is usually TMS (tetramethylsilane) or DSS (2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt). The chemical shift phenomenon arises because motion of electrons generates secondary magnetic fields that, at the location of a nucleus, can enhance or oppose the main static magnetic field, thereby perturbing the energy levels of the spins. Thus, chemical shifts are sensitive to the protein conformation, and can be used to derive several structural parameters. The magnetization does not precess infinitely in the transverse plane but turns back to the equilibrium state. This process is called relaxation and consists of longitudinal and transverse relaxation, with T1 and T2 time-constants respectively. While T1 is the time of the thermal equilibrium re-establishment, T2 is the time of the transverse magnetization dephasing. The T2

constant characterizes the exponential decay of the signal in the receiver coil and depends on the homogeneity of the static magnetic field, on the spin-spin dipolar interactions, and on the overall tumbling time of the molecule. The transverse relaxation determines the line width of the signal. A slow overall tumbling results in a shorter T2 and thus the spectra of large proteins display very broad lines (line-broadening of signal).

Another important effect in NMR spectroscopy is the interaction between spins. The scalar coupling (J-coupling) is the interaction of neighboring spins, mediated by the electron spins of the chemical bonds. The interaction between two spins can occur also through space. This phenomenon is called the 'nuclear Overhauser effect' (NOE) and is a consequence of dipole-dipole interactions of two spins close in space. Both of the above spin couplings are detectable only between non-equivalent nuclei (with different chemical shifts). While J-coupling depends on the angle and the number of bonds between the coupling nuclei, NOE depends on the space distance between the nuclei. Therefore, both contain important structural parameters of the molecule.

5.4.2 Protein NMR

Proteins contain a limited number of nucleus types (H, C, N, O and S) and only NMR-active nuclei can be exploited to gather chemical and structural information from NMR experiments. For instance, the most abundant natural isotopes of carbon (^{12}C), oxygen (^{16}O) and sulfur (^{32}S) are not NMR-active. The most abundant nitrogen isotope (^{14}N) is NMR-active, but hardly detectable. Eventually, the most widely exploited nucleus in the study of unlabeled proteins is ^{1}H. The main problem in the interpretation of protein NMR spectra is represented by the resonance assignment, in other words, to know which resonance corresponds to which specific atom. The dispersion of signals caused by chemical shifts and the possibility of magnetization transfer through spin-spin interaction are the prerequisite for site specific resonance assignments, which are the base for structure determination. One-dimensional ^{1}H spectra of proteins are very crowded and signal overlap is a serious problem. This problem was addressed designing pulse sequences that yield a two-dimensional ^{1}H-NMR spectrum, thus expanding the information in a second frequency dimension. By varying the number, the duration, the power, the phase, and the timing of the applied RF pulses within a pulse sequence, the evolution of the

magnetization in the sample can be manipulated in such a way that the NMR experiment can be tailored to extract specific information about the spin systems. The diagonal in a 2D ^{1}H-NMR spectrum corresponds to a normal one-dimensional NMR spectrum [249]. The off-diagonal peaks (cross-peaks) result from interactions between hydrogen atoms. These cross-peaks can reveal interactions occurring either through bonds or through space. COSY (Correlation Spectroscopy), DQF-COSY (Double Quantum Filtered COSY), and TOCSY (Total Correlation Spectroscopy) belong to the first class, whereas NOESY (Nuclear Overhauser Effect Spectroscopy) and ROESY (Rotating frame Overhauser Effect Spectroscopy) belong to the second. COSY experiments give cross-peaks between hydrogen atoms that are covalently connected through one or two bonds and display a measurable spin-spin coupling (J).

As each amino acid type has a specific set of covalently connected hydrogen atoms that will give a specific combination of cross-peaks, from the COSY spectrum it is possible to identify the spin system of each residue in the protein. TOCSY experiments represent a more powerful version of the COSY pulse sequence and correlate all protons of a spin system (of one amino acid). NOESY spectra, on the other hand, give rise to peaks between pairs of hydrogen atoms that are close together in space even if they belong to residues that are quite distant in the primary sequence. NOESY spectra therefore contain the information on the three-dimensional structure of the protein.

The resonance assignment can then be divided in two steps: in the first step, sets of resonances belonging to the same residue are identified and the spin system of that residue, i.e. the amino acid type, is assigned. COSY and TOCSY spectra are used to this purpose. In the second step, sequence specific information is obtained, i.e. the residue number of a specific spin system is assigned. NOESY spectra are used to this purpose, taking advantage of the fact that residues that are one next to the other in the sequence are necessarily close in space, regardless of the secondary structure and the three-dimensional structure of the protein.

While nearly complete assignments are necessary for the calculation of the three-dimensional structure, NMR spectra can be also used in a qualitative way, providing nevertheless a very useful mean to determine if a protein is correctly folded, if it is only partially folded, or if it is unfolded in solution, in nearly physiological conditions of temperature, pH and ionic strength. Depending on the chemical type, the ^{1}H chemical shifts of a protein fall in different regions of the spectrum: 8-9 ppm

(amide NH), 7-8 ppm (CH, aromatic), 4-5 ppm (CHα, aliphatic), 0-3 ppm (CH, CH_2, CH_3, aliphatic). These values, however, have been derived for random coil peptides, and depend not only on the residue type, but also on the secondary and tertiary structure of the protein.

For a short, random coil polypeptide dissolved in H_2O, nearly all the atoms experience the same environment, which is essentially given by the solvent. On the contrary, in a globular protein with a solvent exposed surface and a hydrophobic, well packed core, residues will experience very different environments, leading to a high dispersion in the chemical shifts. Indeed, the high dispersion in the chemical shift of amide protons is commonly used as a sign of globular structure. It has also been shown that there is a correlation between the chemical shift of the CHα protons and the secondary structure type of that specific residue. Finally, the proximity of methyl groups to aromatic rings in the hydrophobic core makes them sensitive to ring current effects, which can induce strong deviations from random coil chemical shift values.

Even though homonuclear spectra are useful, modern NMR spectroscopy relies on the use of heteronuclear (^{15}N and ^{13}C) multidimensional experiments, which facilitate the assignment and structure determination, especially of larger proteins. The natural abundance of the ^{15}N and ^{13}C isotopes is only 0.2 % and 1 %, respectively. To be able to benefit from the carbon and nitrogen chemical shifts to assign and the hundreds of protein ^{1}H-resonances, the isotopes ^{15}N and ^{13}C are enriched to almost 100 % in the target protein. This is achieved by growing the bacteria expressing the protein on media containing $^{15}NH_4Cl$ and ^{13}C-glucose as the sole nitrogen and carbon sources, respectively. In such uniformly labeled protein, also the J-coupling between protons and nitrogen or carbon and between nitrogen and carbon can be exploited for NMR spectroscopy.

One of the most widely used experiment in protein NMR is the HSQC (Heteronuclear Single Quantum Correlation), which allows correlating each proton to either its directly bound nitrogen (^{1}H–^{15}N HSQC) or carbon (^{1}H–^{13}C-HSQC). The ^{1}H–^{15}N HSQC is fundamental to protein NMR studies because detects only protons covalently bonded to ^{15}N atoms. For ^{15}N-labeled proteins it gives information about every amide proton on the backbone as well as side-chain amide protons. Using a variety of three-dimensional spectra with uniformly ^{15}N- and ^{13}C-labeled protein, each of the amide resonances that appear on the ^{15}N-HSQC spectrum can be assigned to its corresponding position in the protein's primary sequence. Obtaining "assignments"

for the $^{1}H–^{15}N$ HSQC spectrum is the first critical step for translating NMR data into biochemically relevant information.

The standard approach for obtaining assignments of NH cross-peaks is based on the use of triple resonance spectra (^{1}H, ^{15}N and ^{13}C) that provide selective chemical shift correlations. For instance, the HNCACB experiment correlates the NH cross-peak with the Cα and Cβ from the same residue (intra-residue correlation) and from the previous residue (inter-residue correlation) [248]. Therefore, each amide residue (except Gly) should display four peaks, corresponding to the two Cα and Cβ signals. The analysis of a 3D experiment is usually reduced to the analysis of a series of 2D spectra (strips), in which each strip corresponds to the relevant portion of a 2D plane. The assignment process consists of matching the intra-residue Cα and Cβ signals from one strip with the inter-residue Cα and Cβ from another strip. In this manner, segments of connecting residues are linked together. To complement the information provided by the HNCACB, the CBCACONH is used. Magnetization transfer in this experiment goes through the C'. Therefore, it correlates exclusively the amide nitrogen and proton with the Cα and Cβ from the previous residue. The two spectra are generally analyzed together and displayed as alternating strips.

Protein structure determination by NMR involves resonance assignment, listing of structural restraints, and structure calculation and validation. The most important restraints derive from NOE cross-peak assignments yielding proton-proton distances (< 6°A) and from J-coupling constants giving dihedral angle restraints. Applying the experimental restraints together with geometric parameters known from small molecules, an ensemble of protein conformations can be obtained. All structures within the ensemble fulfill the structural restraints, representing the NMR solution structure. The r.m.s. (root mean square) deviation between these structures is used to assess how well the structure calculations have converged. Dihedral angles are validated by their agreement with the most favored angle combinations in the Ramachandran plot.

Beside the protein three-dimensional structure, the knowledge of the protein dynamics is also crucial for the understanding of protein function. Proteins possess a variety of motions such as bond vibrations, side-chain rotations, segmental motions, domain movements etc. Protein dynamics are extremely complex and difficult to analyze, because a variety of motions take place in the same molecule and at the same

time [250]. Since nuclear spin relaxation mechanisms are strictly correlated with local motion, overall and internal protein motions can be accurately monitored by nuclear spin relaxation measurements. Nuclear spin relaxation experiments in solutions of uniformly ^{15}N labeled proteins are a popular approach for studying global protein motions and internal protein backbone mobility [251]. Typical experiments involved in ^{15}N relaxation studies include measurements of the transverse (R2=1/T2) and longitudinal (R1=1/T1) relaxation rates of the ^{15}N nuclear spin and of the ^{1}H-^{15}N cross-relaxation rate (^{1}H-^{15}N NOE). Typically, a series of several 2D (^{1}H–^{15}N HSQC) spectra are recorded with various relaxation delays, ranging from very short to the longest delays, which usually correspond to 1.5–2 relaxation times. By measuring NH peak intensities in these spectra, the NMR signal decay can be determined, which further gives relaxation rates for a ^{15}N nucleus of a particular residue. R2, R1 and NOE depend on how fast the NH vector tumbles in solution. In a rigid protein tumbling isotropically, all residues would be expected to display identical values of R2, R1, and NOE [252], [253]. The evidence for the existence of the protein internal flexibility can be obtained from local differences in the magnitude of the relaxation parameters.

5.4.3 Mapping protein-ligand interactions by NMR

NMR can provide a well of information on protein-ligand interactions, including structure, dynamics, binding kinetics and affinity. Of the NMR parameters, chemical shifts are the ones that have been used most extensively in detecting protein ligand interactions (the chemical shift perturbation approach) [204], [205], [140]. Since the chemical shift of a nucleus depends on its chemical environment, it will be affected by any intra- or intermolecular event, including binding of a ligand. For this reason, the chemical shifts of protons (and other nuclei) at the binding interface can be used to monitor binding and identify the interacting structural motifs. The observed chemical shift changes may be caused either by direct interaction or by conformational changes. In the chemical shift perturbation experiments, two-dimensional ^{1}H–^{15}N HSQC or ^{1}H–^{13}C HSQC spectra are recorded in presence and absence of the ligand. By comparing the chemical perturbations in the spectrum after addition of the ligand, binding can be detected. Only peaks of the isotopically labeled protein are observed.

The use of $^{1}H–^{15}N$ HSQC spectra allows detection of changes in the amide cross-peaks of the backbone (NH) and side chains of Gln and Asn (NH_2). Chemical shift perturbation (csp) is calculated as the difference in NH chemical shift between the free and bound protein (described in methods). If the backbone NH assignments are known, protein residues involved in binding can be determined, and if the protein 3D structure is also known, the binding site can be mapped to a specific protein region.

The use of $^{1}H–^{13}C$ HSQC spectra yields information on chemical shift changes in all side chains (CH, CH_2 and CH_3 cross-peaks). In the $^{1}H–^{13}C$ HSQC, methyl resonances are mostly used, firstly because of their signal intensity (there are 3 hydrogen nucleus of the same chemical shift), and secondly because their spectral region is well defined (0.5-2 ppm in the ^{1}H dimension, and 8-30 ppm in the ^{13}C dimension). The use of $^{1}H–^{15}N$ HSQC is normally preferred because it requires neither the relatively expensive ^{13}C enrichment nor the often lengthy process of side chain assignment.

Chemical shift of backbone NHs, but not of CH_3s, may be influenced not only by ligand binding but also by hydrogen bonding, exchange with the solvent, and subtle structural changes. For this reason, ^{13}C chemical shift perturbation may be more suitable to monitor binding, especially when the protein-ligand interaction is mainly dictated by hydrophobic contacts.

The chemical shift perturbation approach is probably the most straightforward and widely used NMR technique for determining binding interfaces. It can be applied to a variety of molecular systems, such as protein-protein, protein-peptide, protein-nucleic acids, protein-carbohydrates and protein-drug interactions, in a wide range of binding affinities (K_d values from nM to mM). Nonetheless, there are also some weak points which should be taken into account. It can be difficult to distinguish chemical shift perturbations directly due to binding from those associated with small structural changes in the protein and even cascade effects occurring far from the binding site [142], [143]. Line shape is determined by the binding kinetics, and, while slow exchange on the NMR time scale can lead to spectra complicated by the presence of two sets of resonances, arising from the free and bound form of the protein, intermediate exchange can lead to line broadening and eventually disappearance of cross-peaks. Finally, it should be also remembered that csp experiments alone do not provide any information on the bound ligand.

Several alternative NMR techniques can also provide information about the protein-ligand interface. In cross-saturation or saturation transfer (STD) experiments [254], the observed protein is $^{2}H,^{15}N$-labeled with its amide deuterons exchanged back to protons, whereas the other 'donating' partner protein is unlabeled. Saturation of the unlabeled ligand leads by cross-relaxation mechanisms to signal attenuation (typically monitored by ^{15}N-HSQC spectra) of those residues in the labeled protein that are at the binding interface. Cross-saturation experiments are believed to give a more reliable picture of the interface than CSP data, which can suffer from 'false positives' due to conformational changes. However, protein deuteration is a prerequisite for these experiments.

To structural information of the bound ligand can be provided by transferred NOE experiments (tr-NOE) [207]. These experiments are usually applied to systems where small molecular weight ligands bind to large proteins with moderate affinities and suitable k_{off} rates. The chemical exchange of the ligand between its free and bound states results in a transfer of information about the structure of the ligand in its bound conformation to the NMR spectrum of the free ligand. The latter is much easier to observe because of the narrow NMR lines widths and the higher concentration. Thus, the resonance line shapes resemble those of the unbound ligand in solution, but the cross-relaxation measured from the NOE spectrum is predominantly determined by internuclear distances within the ligand in the bound state. The tr-NOE spectroscopy is apparently straightforward from the technical point of view, as it implies the acquisition of standard 2D NOESY or ROESY spectra on samples that do not have to be isotopically labeled. These experiments are very attractive because they allow for a detailed structural analysis of high-molecular-weight complexes that are not suitable for direct study by NMR due to line broadening. On the other hand, the molar excess of the ligand required for these experiments may lead to the non-specific binding, which is the main weak point of this method [255].

CSP experiments can be used not only qualitatively to map a binding site, but also quantitatively to measure binding constants [256]. In this case, a titration of the protein with the ligand is carried out, and an NMR spectrum is acquired at each titration point. CSP values can be plotted vs. ligand concentration or fractional saturation, and residue specific K_d values can be estimated fitting data with a curve. Although this method is very powerful, in that it provides residue specific information, it is applied to a wide variety of ligands at the equilibrium, in solution,

without any labeling, it has also some disadvantages. It is very time consuming, it requires fairly large amounts of materials because of the relatively low sensitivity of NMR, and an estimate of the K_d is required in advance. Moreover, exchange processes should be taken into consideration. On the NMR time scale, rates of exchange between the free and complexed forms are broadly classified as slow, intermediate and fast. In the slow exchange rate limit, two distinct resonances are detectable, one associated with the free form, the other with the bound form of the protein. Upon titration, the intensity of the peak associated with the bound form increases while the intensity of the peak associated with the free form decreases. In the fast exchange, one single resonance can be detected, at a chemical shift value that is the population weighted average of the resonances in the free and bound form. In this case, upon addition of a ligand, the resonance will shift towards the frequency corresponding to the bound form. In the intermediate exchange rate, the lines corresponding to the free and bound form are broadened, and may even become undetectable.

5.4.4 In-cell NMR

By in-cell NMR experiments, proteins and other macromolecules can be investigated directly in their natural environment [196], [197], [198]. There are three main prerequisites for investigating proteins in the cellular environment using in-cell NMR spectroscopy. First, the NMR signals of the target protein must be distinguished from the background of the NMR resonances arising from all other proteins, nucleic acids, and other cellular components. Second, the protein must tumble freely inside the cell, and its rotational correlation time must be within a range that allows the observation of narrow resonance lines. Third, cells have to survive the conditions inside the NMR tube without significant changes of their metabolic state at least for the time period of the experiment.

To distinguish NMR resonances, the target protein must be labeled with the NMR active isotopes, usually ^{15}N and ^{13}C. This can be achieved either by overexpression of the target protein in a isotopically labeled medium, by direct injection of the purified labeled protein into cells, or by delivery of the labeled protein into cells using cell-penetrating peptides as a fusion tag [257].

Bacterial cells (especially E. *coli* strains) have been the most widely used system for the overexpression of the target protein [196]. Overexpression in the isotopically labeled medium, however, causes labeling not only of the target protein, but also of endogenous proteins and other cellular components. To reduce the background signal, bacterial can be first grown in an unlabeled medium and transferred into the labeled medium only during the induction time. The use of drugs that suppress expression of all endogenous proteins, like rifampicin, was also proposed as an alternative to reduce the background signals [196]. However, the crucial factor is the expression level of the target protein, whereas suppression of the background signal has a minor impact on spectral quality. While E. coli are perfectly well suited for biophysical studies or to study prokaryotic proteins, a eukaryotic system is preferred when studying human proteins. *Xenopus laevis* oocytes were the first eukaryotic cells to be used for in-cell NMR, due to their very large size that allows for direct injection of the labeled protein [200]. In this way, the background due to cellular components is completely removed, and the cellular concentration of the target protein can be controlled. Recently, other eukaryotic expression systems, such as yeast and insect cells, have also been used to express exogenous proteins for structural studies [198].

Limitations in in-cell NMR can arise from excessive line broadening. This can be due to the formation of large complexes, to unspecific binding to other macromolecules or cellular structures, or to localization in a very viscous cellular compartment. The intracellular viscosity (especially in nucleus) is higher than in commonly used buffers, and cannot be manipulated in order to increase the spectral quality. Instead, a direct comparison of protein line widths in buffer and in in-cell NMR experiments (in the absence of protein binding) can be used for a qualitative estimation of the intracellular viscosity. Poor cell survival, leakage from the cell, and poor homogeneity of the sample can represent additional problems. By decreasing the cell concentration in the NMR tube, both the homogeneity and the cell survival increase, but at the same time the signal intensity decreases. Cells can be kept alive for a longer time modifying NMR tubes to allow for the constant exchange of the medium to supply fresh nutrients and oxygen.

In-cell NMR has been used to address several biophysical and biological issues, for example to investigate the effect of molecular crowding on the structure of proteins that behave as intrinsically disordered in solution, to detect folding of

disordered regions upon binding to their molecular target, to map protein-protein interaction surfaces, to monitor post-translational modifications such as phosphorylation in real time. Recently, the 3D structure determination of a protein in its cellular environment has been reported [258]. In-cell NMR has also the potential to become a useful tool for drug discovery, combining the screening power of *in vitro* NMR experiments with the capability to select from a pool of potential drugs only those molecules that not only bind to a certain macromolecule but also that can penetrate the cellular membrane and reach the target inside the cell.

5.5 Membrane mimicking models

Biological membranes, and in particular the plasma membrane of human cells, are very complex systems made of a lipid bilayer composed of a variety of lipids with different lengths in the hydrocarbon tail and different head groups, and populated by a number of integral, transmembrane, membrane-anchored, and proximal membrane proteins [259]. Phospholipids constitute the main component of cell membranes and consist of an elongated double-chain hydrophobic tail and a hydrophilic head. The phospholipid interfacial layer is ~15 Å thick, is made by the polar groups of the fatty acid esters, the negatively charged group of the phosphate moiety, and by the phospholipid head group, which in turn determines the overall net charge of the phospholipid. While phospholipids are the major lipid component of most cell membranes, the exact composition varies from one cell type to another. There are many different classes of lipids bearing zwitterionic, anionic, or cationic head groups. Additionally, each phospholipid type can have considerable variations in the acyl chain composition, including chain length and degree of saturation. The length of the acyl chains vary in natural lipids between C14 and C18. The distribution of lipids between the two leaflets of plasma membranes is asymmetric. The leaflet facing the cytoplasm (the inner leaflet) consists primarily of the negatively charged phosphatidylserine (PS) and of the zwitterionic phosphatidylethanolamine (PE). The outer leaflet consists mostly of the zwitterionic phosphatidylcholine (PC).

The urge to study integral membrane proteins or fragments thereof lead very early to the development of several molecular systems that were mainly designed to mimic the hydrophobic environment experienced by membrane proteins integrated in the membrane lipid bilayer. These systems include SDS and DPC micelles,

lysophospholipid micelles, liposomes, bicelles and nanodiscs. The use of these membrane mimicking models has been recently extended to the study of membrane-associated and interfacial proteins. While mimicking the complexity of biological membranes (**Figure 2.21**) remains a serious challenge, these artificial systems proved to be quite effective in biophysical studies.

5.5.1 Micelles

One of the simplest membrane models is given by micelles, which are easy to prepare and suitable for different spectroscopic techniques. Micelles are formed in aqueous solution from amphipathic molecules (detergents, lysophospholipids, etc) which position their polar groups towards the surrounding aqueous medium whereas their hydrophobic chains are oriented toward the inside of the micelle. Detergent micelles such as sodium docecyl sulfate (SDS) or dodecyl phosphocholine (DPC) represent a commonly used and often reasonable model for biological membranes. SDS and DPC micelles consist of about ~60 molecules, and are negatively charged (SDS) or neutral (DPC). Micelle preparation is straightforward, and in most cases is achieved simply by dispersing a given amount of a suitable detergent in a buffer. Formation of micelle is spontaneous when the detergent concentration is equal or above its specific critical micelle concentration (2–8 mM for SDS, depending on the buffer ion strength). Micelles are approximately spherical in shape with diameters of about 3 to 5 nm, which is close to the thickness of phospholipid bilayers (3.7 to 5.3 nm, dependent on the type of lipid and its physical state). Owing to their small size and high surface curvature, detergent micelles are characterized by low lipid packing densities comparing to biological membranes.

Micelles can be also prepared from lysophospholipids, phospholipids with a single alkyl chain. Like detergent micelles, lysophospholipid micelles are prepared in a straightforward manner and are relatively stable in time in a wide range of temperature and pH conditions. Lysophospholipid micelles are biologically more relevant than detergent micelles, having a head groups identical to those found in phospholipids of biological membranes. Moreover, lysophospholipid micelles have higher aggregation numbers, ~120 molecules/micelle and thus have higher lipid packing densities compared to SDS or DPC micelles [260].

However, micelles are not ideal membrane mimicking systems as they have a monolayer structure and a high curvature, a loose packing of the polar head groups, and tighter chain entanglement in the center. These features therefore differ significantly from the biological membrane morphology.

5.5.2 Liposomes

Liposomes are aqueous compartments enclosed by phospholipid bilayer membranes. The liposome bilayer is composed of two monolayers of lipid molecules with their hydrophobic tails facing each other and their polar (hydrophilic) surfaces facing the aqueous solvent. Liposome bilayers can exist in different phases, depending on the lipid structure and the temperature. The two most important phases are the gel phase and the liquid–crystalline phase [261]. In the gel phase, the bilayer is well packed while in the liquid–crystalline phase (fluid phase) the chains are more disordered and more loosely packed. The fluid phase is biologically more relevant. Liposomes are classified by whether the sequestered aqueous volume is enclosed by on (unilamellar) or more (multilamellar) bilayers. Unilamellar liposomes, used as membrane models in structural biology, are further divided into small unilamellar vesicles (SUV) and large unilamellar vesicles (LUV) [262]. SUVs are almost spherical, like SDS micelles, but are much bigger (25-100 nm) and thus with smaller curvature. While SUVs are good models of the cell membrane, they are stable only in a narrow pH range and above the critical temperature, they tend to fuse to yield larger particles, and they have a statistical size distribution, much alike polymers, which can be limited only to some extent by extrusion through membranes of known pore diameter.

Phospholipid vesicles are suitable for CD and fluorescence studies, but usually not for high-resolution NMR studies, due to their limited stability and large size.

5.5.3 Bicelles

Bicelles are disc-shaped, mixed micelles composed of a bilayer of long-chain phospholipids (12-18 carbons) such as dimyristoyl phosphatidylcholine (DMPC) surrounded by a rim of short-chain detergents (6–8 carbons) such as dihexanoyl phosphatidylcholine (DHPC) [263], [264]. The short chain detergents play the role of stabilizing the edges of the bicelle. In terms of lipid packing density, conformation and dynamics, the bicelle DMPC bilayer is very similar to the liquid crystalline state

of the liposome DMPC bilayer. In contrast to liposomes, bicelles do not have an aqueous compartment enclosed inside. The molar ratio of DMPC to DHPC determines the diameter of the bicelle. Bicelles may be doped with phosphatidylserine (PS) or phosphatidyglyerol (PG), making them negatively charged [139]. Compared to micelles, bicelles have higher lipid packing density, lower radius of curvature. The lipid bilayer of bicelles is considered to mimic the biological membrane better than micelles. At the same time, bicelles are sufficiently small to give rise to narrow lines in solution NMR studies [265], [266], [267]. On the other hand, preparation of bicelles is somewhat tricky and it may be difficult to ascertain whether the interaction of the bound protein with the bicelle is occurring with the tightly packed DMPC bilayer, with the looser packed DHPC rim, or both.

5.5.4 Nanodiscs

Nanodiscs consist of a phospholipid bilayer surrounded by an amphiphilic protein that defines their size [268]. The lipid core of the nanodisc is made of ~150 dipalmitoylphosphatidylcholine (DPPC) and palmitoyloleoyl- phosphatidylcholine (POPC) molecules, and is surrounded by a scaffold protein, usually a derivative of apolipoprotein A-I lacking the globular N-terminal domain and including only the helical region. The overall structure is similar to that of discoidal high density lipoprotein particles involved in cholesterol transport. In principle, both the lipid composition and the size of the nanodisc can be varied, with the length of the scaffold protein surrounding the lipid core determining its size, and they have a nearly flat surface. On the other hand, their complex preparation protocol has so far limited their applications. It has been recently shown that nanodiscs are suitable not only for solid state and magic angle spinning NMR techniques but also for NMR solution studies of membrane proteins [269].

Abbreviations

ADAM, A Disintegrin and Metalloproteinase

AF6, ALL-1 Fusion partner on chromosome 6 (a smaller splicing variant of Afadin) in this thesis used as a synonymous of Afadin

AF6_PDZ, the Afadin PDZ domain

AGS, Alagille Syndrome

ANK, Ankyrin repeats

AP1, Activator protein 1

BTD, β-Trefoil Domain

CD, Circular Dichroism

CSL, CBF1/RBP, Suppressor of Hairless and LAG-1

CM5, Carboxymethylated Dextran (sensor chip for SPR)

COSY, Correlation Spectroscopy

CSP, Chemical Shift Perturbation

Dlg, Disks large protein

DLL1-4, Delta-like 1-4

DHPC, Dihexanoyl Phosphatidylcholine

DMPC, 1,2-Dimyristoyl-sn-glycero-3-Phosphocholine

DMPG, 1,2-Dimyristoyl-sn-glycero-3-[Phospho-*rac*-(1-Glycerol)] sodium salt

DMPS, 1,2-Dimyristoyl-sn-glycero-3-[Phospho-L-Serine] sodium salt

DPC, Dodecyl Phosphocholine

DSL, Delta/Serrate/Lag-2 domain

DSS, 2,2-Dimethyl-2-Silapentane-5-Sulfonic acid, sodium salt

EDTA, Ethylene Diamine Tetraacetic Acid

EGF, Epidermal Growth Factor repeat

EHBA, Extrahepatic Biliary Atresia

FPLC, Fast Protein Liquid Chromatography

GRIP, Glutamate Receptor Interacting Protein

HD, Heterodimerization Domain of Notch

HES, Hairy/Enhancer of Split

HSQC, Heteronuclear Single-Quantum Correlation

IMAC, Immobilized Metal Ion Affinity Chromatography

IDP, Intrinsically Disordered Protein

J1_ic, Jagged-1 intracellular region (1094-1218 of human Jagged-1)

J1_tmic, Jagged-1 intracellular region (1086-1218 of human Jagged-1)

LC–MS, Liquid Chromatography coupled to Mass Spectrometry

LMPC, 1-Myristoyl-2-Hydroxy-*sn*-glycero-3-Phosphocholine

LMPG, 1-Myristoyl-2-Hydroxy-*sn*-glycero-3-[Phospho-*rac*-(1-glycerol)

MAGI, MAGUK with inverted domain arrangement

MAGUK, Membrane Associated Guanylate Kinases

Mib, Mind bomb (E3 ligase)

MW, Molecular Weight

N_ec, Notch extracellular region

Neur, Neuralized (E3 ligase)

N_ic, Notch intracellular region

NLS, Nuclear Localization Signals

NMDA, *N*-methyl *D*-aspartate (NMDA receptor)

NMR, Nuclear Magnetic Resonance

NOE, Nuclear Overhauser Effect

NOESY, Nuclear Overhauser Effect Spectroscopy

NRR, Negative Regulatory Region (of N_ec)

NSC, Neuronal Stem Cells

N_tm, Notch transmembrane region

PBS, Phosphate Buffered Saline

PDB, Protein Data Bank

PDZ, PSD-95, Disc-large, and ZO-1 domain

PIP, Phosphatidylinositol Phosphate

PSD-95, Postsynaptic Density protein 95

RAM, RBP-Jkappa-Associated Module

RF, Radio-Frequency field (NMR)

RHR-C, C-terminal Rel-Homology Region

RHR-N, N-terminal Rel-homology Region

RIP, Regulated Intramembrane Proteolysis

RKE, Rat Kidney Epithelial cells

RMSD, Root Mean Square Deviation

RP-HPLC, Reverse Phase-High Performance Liquid Chromatography

SDS, Sodium Dodecyl Sulfate

SH3, Src Homology 3 domain

SMART, Simple Modular Architecture Research Tool

SOP, Sensory Organ Precursors

SPR, Surface Plasmon Resonance

SUV, Small Unilamellar Vesicles

T-ALL, T cell Acute Lymphoblastic Leukemia

TFE, 2,2,2-Trifluoroethanol

TOCSY, Total Correlation Spectroscopy

TOF, Tetralogy Of Fallot

tr-NOE, transferred NOE

References

1. Weinmaster, G., *Notch signal transduction: a real rip and more.* Curr. Opin. Genet. Dev., 2000. **10**(4): p. 363-9.

2. Kadesch, T., *Notch signaling: the demise of elegant simplicity.* Curr. Opin. Genet. Dev., 2004. **14**(5): p. 506-12.

3. Brou, C., et al., *A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE.* Mol. Cell, 2000. **5**(2): p. 207-16.

4. Lyman, D. and M.W. Young, *Further evidence for function of the Drosophila Notch protein as a transmembrane receptor.* Proc. Natl. Acad. Sci. U. S. A., 1993. **90**(21): p. 10395-9.

5. Rebay, I., et al., *Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor.* Cell, 1991. **67**(4): p. 687-99.

6. Brown, M.S., et al., *Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans.* Cell, 2000. **100**(4): p. 391-8.

7. Medina, M. and C.G. Dotti, *RIPped out by presenilin-dependent gamma-secretase.* Cell. Signal., 2003. **15**(9): p. 829-41.

8. Fortini, M.E. and S. Artavanis-Tsakonas, *The suppressor of hairless protein participates in notch receptor signaling.* Cell, 1994. **79**(2): p. 273-82.

9. Maier, M.M. and M. Gessler, *Comparative analysis of the human and mouse Hey1 promoter: Hey genes are new Notch target genes.* Biochem. Biophys. Res. Commun., 2000. **275**(2): p. 652-60.

10. Fischer, A. and M. Gessler, *Delta-Notch--and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors.* Nucleic Acids Res., 2007. **35**(14): p. 4583-96.

11. Kageyama, R., et al., *Roles of bHLH genes in neural stem cell differentiation.* Exp. Cell Res., 2005. **306**(2): p. 343-8.

12. Chen, H., et al., *Conservation of the Drosophila lateral inhibition pathway in human lung cancer: a hairy-related protein (HES-1) directly represses achaete-scute homolog-1 expression.* Proc. Natl. Acad. Sci. U. S. A., 1997. **94**(10): p. 5355-60.

13. Castella, P., et al., *HES-1 repression of differentiation and proliferation in PC12 cells: role for the helix 3-helix 4 domain in transcription repression.* Mol. Cell. Biol., 2000. **20**(16): p. 6170-83.

14. Radtke, F. and K. Raj, *The role of Notch in tumorigenesis: oncogene or tumour suppressor?* Nat. Rev. Cancer, 2003. **3**(10): p. 756-67.

15. Artavanis-Tsakonas, S., M.D. Rand, and R.J. Lake, *Notch signaling: cell fate control and signal integration in development.* Science, 1999. **284**(5415): p. 770-6.

16. Greenwald, I., *LIN-12/Notch signaling: lessons from worms and flies.* Genes Dev., 1998. **12**(12): p. 1751-62.

17. Bolos, V., J. Grego-Bessa, and J.L. de la Pompa, *Notch signaling in development and cancer.* Endocr. Rev., 2007. **28**(3): p. 339-63.

18. Sullivan, D.C. and R. Bicknell, *New molecular pathways in angiogenesis.* Br. J. Cancer, 2003. **89**(2): p. 228-31.

19. Allman, D., et al., *An invitation to T and more: notch signaling in lymphopoiesis.* Cell, 2002. **109**: p. S1-11.

20. Anderson, A.C., E.A. Robey, and Y.H. Huang, *Notch signaling in lymphocyte development.* Curr. Opin. Genet. Dev., 2001. **11**(5): p. 554-60.

21. Screpanti, I., et al., *Notch, a unifying target in T-cell acute lymphoblastic leukemia?* Trends Mol. Med., 2003. **9**(1): p. 30-5.

22. Weng, A.P., et al., *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia.* Science, 2004. **306**(5694): p. 269-71.

23. Harper, J.A., et al., *Notch signaling in development and disease.* Clin. Genet., 2003. **64**(6): p. 461-72.

24. Gridley, T., *Notch signaling and inherited disease syndromes.* Hum. Mol. Genet., 2003. **12**(1): p. R9-13.

25. Piccoli, D.A. and N.B. Spinner, *Alagille syndrome and the Jagged1 gene.* Semin. Liver Dis., 2001. **21**(4): p. 525-34.

26. Le Caignec, C., et al., *Familial deafness, congenital heart defects, and posterior embryotoxon caused by cysteine substitution in the first epidermal-growth-factor-like domain of jagged 1.* Am. J. Hum. Genet., 2002. **71**(1): p. 180-6.

27. Kamath, B.M., et al., *Vascular anomalies in Alagille syndrome: a significant cause of morbidity and mortality.* Circulation, 2004. **109**(11): p. 1354-8.

28. Kohsaka, T., et al., *The significance of human jagged 1 mutations detected in severe cases of extrahepatic biliary atresia.* Hepatology, 2002. **36**(4 Pt 1): p. 904-12.

29. Eldadah, Z.A., et al., *Familial Tetralogy of Fallot caused by mutation in the jagged1 gene.* Hum. Mol. Genet., 2001. **10**(2): p. 163-9.

30. Bulman, M.P., et al., *Mutations in the human delta homologue, DLL3, cause axial skeletal defects in spondylocostal dysostosis.* Nat. Genet., 2000. **24**(4): p. 438-41.

31. Spinner, N.B., *CADASIL: Notch signaling defect or protein accumulation problem?* J. Clin. Invest., 2000. **105**(5): p. 561-2.

32. Haines, N. and K.D. Irvine, *Glycosylation regulates Notch signalling.* Nat. Rev. Mol. Cell Biol., 2003. **4**(10): p. 786-97.

33. Shi, S. and P. Stanley, *Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(9): p. 5234-9.

34. Nichols, J.T., et al., *DSL ligand endocytosis physically dissociates Notch1 heterodimers before activating proteolysis can occur.* J. Cell Biol., 2007. **176**(4): p. 445-58.

35. Klueg, K.M. and M.A. Muskavitch, *Ligand-receptor interactions and trans-endocytosis of Delta, Serrate and Notch: members of the Notch signalling pathway in Drosophila.* J. Cell Sci., 1999. **112 (Pt 19)**: p. 3289-97.

36. Nichols, J.T., A. Miyamoto, and G. Weinmaster, *Notch signaling--constantly on the move.* Traffic, 2007. **8**(8): p. 959-69.

37. Le Borgne, R., *Regulation of Notch signalling by endocytosis and endosomal sorting.* Curr. Opin. Cell Biol., 2006. **18**(2): p. 213-22.

38. Hori, K., et al., *Drosophila deltex mediates suppressor of Hairless-independent and late-endosomal activation of Notch signaling.* Development, 2004. **131**(22): p. 5527-37.

39. Bray, S.J., *Notch signalling: a simple pathway becomes complex.* Nat. Rev. Mol. Cell Biol., 2006. **7**(9): p. 678-89.

40. Ronchini, C. and A.J. Capobianco, *Induction of cyclin D1 transcription and CDK2 activity by Notch(ic): implication for cell cycle disruption in transformation by Notch(ic).* Mol. Cell. Biol., 2001. **21**(17): p. 5925-34.

41. Rangarajan, A., et al., *Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation.* EMBO J., 2001. **20**(13): p. 3427-36.

42. de La Coste, A. and A.A. Freitas, *Notch signaling: distinct ligands induce specific signals during lymphocyte development and maturation.* Immunol. Lett., 2006. **102**(1): p. 1-9.

43. Hrabe de Angelis, M., J. McIntyre, 2nd, and A. Gossler, *Maintenance of somite borders in mice requires the Delta homologue DII1*. Nature, 1997. **386**(6626): p. 717-21.

44. Xue, Y., et al., *Embryonic lethality and vascular defects in mice lacking the Notch ligand Jagged1*. Hum. Mol. Genet., 1999. **8**(5): p. 723-30.

45. Gale, N.W., et al., *Haploinsufficiency of delta-like 4 ligand results in embryonic lethality due to major defects in arterial and vascular development.* Proc. Natl. Acad. Sci. U. S. A., 2004. **101**(45): p. 15949-54.

46. Fleming, R.J., Y. Gu, and N.A. Hukriede, *Serrate-mediated activation of Notch is specifically blocked by the product of the gene fringe in the dorsal compartment of the Drosophila wing imaginal disc.* Development, 1997. **124**(15): p. 2973-81.

47. Fleming, R.J., *Structural conservation of Notch receptors and ligands.* Semin. Cell Dev. Biol., 1998. **9**(6): p. 599-607.

48. Pintar, A., et al., *The intracellular region of Notch ligands: does the tail make the difference?* Biol. Direct, 2007. **2**: p. 19.

49. Koutelou, E., et al., *Neuralized-like 1 (Neurl1) targeted to the plasma membrane by N-myristoylation regulates the Notch ligand Jagged1*. J. Biol. Chem., 2008. **283**(7): p. 3846-53.

50. Takeuchi, T., Y. Adachi, and Y. Ohtsuki, *Skeletrophin, a novel ubiquitin ligase to the intracellular region of Jagged-2, is aberrantly expressed in multiple myeloma.* Am. J. Pathol., 2005. **166**(6): p. 1817-26.

51. LaVoie, M.J. and D.J. Selkoe, *The Notch ligands, Jagged and Delta, are sequentially processed by alpha-secretase and presenilin/gamma-secretase and release signaling fragments.* J. Biol. Chem., 2003. **278**(36): p. 34427-37.

52. Kolev, V., et al., *The intracellular domain of Notch ligand Delta1 induces cell growth arrest.* FEBS Lett., 2005. **579**(25): p. 5798-5802.

53. Ikeuchi, T. and S.S. Sisodia, *The Notch ligands, Delta1 and Jagged2, are substrates for presenilin-dependent "gamma-secretase" cleavage.* J. Biol. Chem., 2003. **278**(10): p. 7751-4.

54. Six, E., et al., *The Notch ligand Delta1 is sequentially cleaved by an ADAM protease and gamma-secretase.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(13): p. 7638-43.

55. Hiratochi, M., et al., *The Delta intracellular domain mediates TGF-beta/Activin signaling through binding to Smads and has an important bi-directional function in the Notch-Delta signaling pathway.* Nucleic Acids Res., 2007. **35**(3): p. 912-22.

56. Hock, B., et al., *PDZ-domain-mediated interaction of the Eph-related receptor tyrosine kinase EphB3 and the ras-binding protein AF6 depends on the kinase activity of the receptor.* Proc. Natl. Acad. Sci. U. S. A., 1998. **95**(17): p. 9779-84.

57. Ascano, J.M., L.J. Beverly, and A.J. Capobianco, *The C-terminal PDZ-ligand of JAGGED1 is essential for cellular transformation.* J. Biol. Chem., 2003. **278**(10): p. 8771-9.

58. Six, E.M., et al., *The notch ligand Delta1 recruits Dlg1 at cell-cell contacts and regulates cell migration.* J. Biol. Chem., 2004. **279**(53): p. 55818-26.

59. Wright, G.J., et al., *Delta proteins and MAGI proteins: an interaction of Notch ligands with intracellular scaffolding molecules and its significance for zebrafish development.* Development, 2004. **131**(22): p. 5659-69.

60. Mizuhara, E., et al., *MAGI1 recruits Dll1 to cadherin-based adherens junctions and stabilizes it on the cell surface.* J. Biol. Chem., 2005. **280**(28): p. 26499-507.

61. Mandai, K., et al., *Afadin: A novel actin filament-binding protein with one PDZ domain localized at cadherin-based cell-to-cell adherens junction.* J. Cell Biol., 1997. **139**(2): p. 517-28.

62. Hsueh, Y.P., et al., *Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2.* Nature, 2000. **404**(6775): p. 298-302.

63. Lee, J.H., et al., *Interactions with p300 enhance transcriptional activation by the PDZ-domain coactivator Bridge-1.* J. Endocrinol., 2005. **187**(2): p. 283-92.

64. Nakata, A., et al., *GRIP1tau, a novel PDZ domain-containing transcriptional activator, cooperates with the testis-specific transcription elongation factor SII-T1.* Genes Cells, 2004. **9**(11): p. 1125-35.

65. Pfister, S., et al., *Interaction of the MAGUK family member Acvrinp1 and the cytoplasmic domain of the Notch ligand Delta1.* J. Mol. Biol., 2003. **333**(2): p. 229-35.

66. Quilliam, L.A., et al., *M-Ras/R-Ras3, a transforming ras protein regulated by Sos1, GRF1, and p120 Ras GTPase-activating protein, interacts with the putative Ras effector AF6.* J. Biol. Chem., 1999. **274**(34): p. 23850-7.

67. Doyle, D.A., et al., *Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ.* Cell, 1996. **85**(7): p. 1067-76.

68. Songyang, Z., et al., *Recognition of unique carboxyl-terminal motifs by distinct PDZ domains.* Science, 1997. **275**(5296): p. 73-7.

69. Niethammer, M., et al., *CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90.* Neuron, 1998. **20**(4): p. 693-707.

70. Zhang, Y., et al., *Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain.* Protein Sci., 2007. **16**(8): p. 1738-50.

71. Daniels, D.L., et al., *Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition.* Nat. Struct. Biol., 1998. **5**(4): p. 317-25.

72. Stricker, N.L., et al., *PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences.* Nat. Biotechnol., 1997. **15**(4): p. 336-42.

73. Tochio, H., et al., *Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide.* Nat. Struct. Biol., 1999. **6**(5): p. 417-21.

74. Maximov, A., T.C. Sudhof, and I. Bezprozvanny, *Association of neuronal calcium channels with modular adaptor proteins.* J. Biol. Chem., 1999. **274**(35): p. 24453-6.

75. Zhou, H., et al., *Solution structure of AF-6 PDZ domain and its interaction with the C-terminal peptides from Neurexin and Bcr.* J. Biol. Chem., 2005. **280**(14): p. 13841-7.

76. Jaulin-Bastard, F., et al., *The ERBB2/HER2 receptor differentially interacts with ERBIN and PICK1 PSD-95/DLG/ZO-1 domain proteins.* J. Biol. Chem., 2001. **276**(18): p. 15256-63.

77. Karthikeyan, S., T. Leung, and J.A. Ladias, *Structural determinants of the Na+/H+ exchanger regulatory factor interaction with the beta 2 adrenergic and platelet-derived growth factor receptors.* J. Biol. Chem., 2002. **277**(21): p. 18973-8.

78. Kang, B.S., et al., *PDZ tandem of human syntenin: crystal structure and functional properties.* Structure, 2003. **11**(4): p. 459-68.

79. Basdevant, N., H. Weinstein, and M. Ceruso, *Thermodynamic basis for promiscuity and selectivity in protein-protein interactions: PDZ domains, a case study.* J. Am. Chem. Soc., 2006. **128**(39): p. 12766-77.

80. Gerek, Z.N., O. Keskin, and S.B. Ozkan, *Identification of specificity and promiscuity of PDZ domain interactions through their dynamic behavior.* Proteins, 2009. **77**(4): p. 796-811.

81. Cohen, N.A., et al., *Binding of the inward rectifier K+ channel Kir 2.3 to PSD-95 is regulated by protein kinase A phosphorylation.* Neuron, 1996. **17**(4): p. 759-67.

82. Chung, H.J., et al., *Regulation of the NMDA receptor complex and trafficking by activity-dependent phosphorylation of the NR2B subunit PDZ ligand.* J. Neurosci., 2004. **24**(45): p. 10248-59.

83. Matsuda, S., S. Mikawa, and H. Hirai, *Phosphorylation of serine-880 in GluR2 by protein kinase C prevents its C terminus from binding with glutamate receptor-interacting protein.* J. Neurochem., 1999. **73**(4): p. 1765-8.

84. Hegedus, T., et al., *C-terminal phosphorylation of MRP2 modulates its interaction with PDZ proteins.* Biochem. Biophys. Res. Commun., 2003. **302**(3): p. 454-61.

85. Chung, H.J., et al., *Phosphorylation of the AMPA receptor subunit GluR2 differentially regulates its interaction with PDZ domain-containing proteins.* J. Neurosci., 2000. **20**(19): p. 7258-67.

86. Christopherson, K.S., et al., *PSD-95 assembles a ternary complex with the N-methyl-D-aspartic acid receptor and a bivalent neuronal NO synthase PDZ domain.* J. Biol. Chem., 1999. **274**(39): p. 27467-73.

87. Im, Y.J., et al., *Crystal structure of GRIP1 PDZ6-peptide complex reveals the structural basis for class II PDZ target recognition and PDZ domain-mediated multimerization.* J. Biol. Chem., 2003. **278**(10): p. 8501-7.

88. Zimmermann, P., et al., *PIP(2)-PDZ domain binding controls the association of syntenin with the plasma membrane.* Mol. Cell, 2002. **9**(6): p. 1215-25.

89. Kachel, N., et al., *Structure determination and ligand interactions of the PDZ2b domain of PTP-Bas (hPTP1E): splicing-induced modulation of ligand specificity.* J. Mol. Biol., 2003. **334**(1): p. 143-55.

90. Gomperts, S.N., *Clustering membrane proteins: It's all coming together with the PSD-95/SAP90 protein family.* Cell, 1996. **84**(5): p. 659-62.

91. Kim, E. and M. Sheng, *PDZ domain proteins of synapses.* Nat. Rev. Neurosci., 2004. **5**(10): p. 771-81.

92. Takai, Y. and H. Nakanishi, *Nectin and afadin: novel organizers of intercellular junctions.* J. Cell Sci., 2003. **116**(Pt 1): p. 17-27.

93. Miyoshi, J. and Y. Takai, *Molecular perspective on tight-junction assembly and epithelial polarity.* Adv. Drug Deliv. Rev., 2005. **57**(6): p. 815-55.

94. Prasad, R., et al., *Cloning of the ALL-1 fusion partner, the AF-6 gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation.* Cancer Res., 1993. **53**(23): p. 5624-8.

95. Boettner, B., et al., *The junctional multidomain protein AF-6 is a binding partner of the Rap1A GTPase and associates with the actin cytoskeletal regulator profilin.* Proc. Natl. Acad. Sci. U. S. A., 2000. **97**(16): p. 9064-9.

96. Kouichi, T., *Two Cell Adhesion Molecules, Nectin and Cadherin, Interact through Their Cytoplasmic Domain–associated Proteins.* The Journal of Cell Biology, 2000. **150** p. 1161–1175.

97. Radziwill, G., et al., *The Bcr kinase downregulates Ras signaling by phosphorylating AF-6 and binding to its PDZ domain.* Mol. Cell. Biol., 2003. **23**(13): p. 4663-72.

98. Ebnet, K., et al., *Junctional adhesion molecule interacts with the PDZ domain-containing proteins AF-6 and ZO-1.* J. Biol. Chem., 2000. **275**(36): p. 27979-88.

99. Su, L., et al., *AF-6 controls integrin-mediated cell adhesion by regulating Rap1 activation through the specific recruitment of Rap1GTP and SPA-1.* J. Biol. Chem., 2003. **278**(17): p. 15232-8.

100. Buchert, M., et al., *The junction-associated protein AF-6 interacts and clusters with specific Eph receptor tyrosine kinases at specialized sites of cell-cell contact in the brain.* J. Cell Biol., 1999. **144**(2): p. 361-71.

101. Bezprozvanny, I. and A. Maximov, *Classification of PDZ domains.* FEBS Lett., 2001. **509**(3): p. 457-62.

102. Hambleton, S., et al., *Structural and functional properties of the human notch-1 ligand binding region.* Structure, 2004. **12**(12): p. 2173-83.

103. Cordle, J., et al., *A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition.* Nat. Struct. Mol. Biol., 2008. **15**(8): p. 849-57.

104. Pintar, A., et al., *Exon 6 of human JAG1 encodes a conserved structural unit.* BMC Struct. Biol., 2009. **9**: p. 43.

105. Shimizu, K., et al., *Mouse jagged1 physically interacts with notch2 and other notch receptors. Assessment by quantitative methods.* J. Biol. Chem., 1999. **274**(46): p. 32961-9.

106. Gordon, W.R., et al., *Structural basis for autoinhibition of Notch.* Nat. Struct. Mol. Biol., 2007. **14**(4): p. 295-300.

107. Nam, Y., et al., *Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes.* Cell, 2006. **124**(5): p. 973-83.

108. Wilson, J.J. and R.A. Kovall, *Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA.* Cell, 2006. **124**(5): p. 985-96.

109. Nam, Y., et al., *Structural requirements for assembly of the CSL.intracellular Notch1.Mastermind-like 1 transcriptional activation complex.* J. Biol. Chem., 2003. **278**(23): p. 21232-9.

110. Mosavi, L.K., et al., *The ankyrin repeat as molecular architecture for protein recognition.* Protein Sci., 2004. **13**(6): p. 1435-48.

111. Bertagna, A., et al., *The effects of conformational heterogeneity on the binding of the Notch intracellular domain to effector proteins: a case of biologically tuned disorder.* Biochem. Soc. Trans., 2008. **36**(Pt 2): p. 157-66.

112. Petcherski, A.G. and J. Kimble, *Mastermind is a putative activator for Notch.* Curr. Biol., 2000. **10**(13): p. R471-3.

113. Fryer, C.J., et al., *Mastermind mediates chromatin-specific transcription and turnover of the Notch enhancer complex.* Genes Dev., 2002. **16**(11): p. 1397-411.

114. Friedmann, D.R., J.J. Wilson, and R.A. Kovall, *RAM-induced allostery facilitates assembly of a notch pathway active transcription complex.* J. Biol. Chem., 2008. **283**(21): p. 14781-91.

115. Radivojac, P., et al., *Intrinsic disorder and functional proteomics.* Biophys J., 2007. **92**(5): p. 1439-56.

116. Dunker, A.K., et al., *Intrinsically disordered protein.* J. Mol. Graph. Model., 2001. **19**(1): p. 26-59.

117. Iakoucheva, L.M., et al., *Intrinsic disorder in cell-signaling and cancer-associated proteins.* J. Mol. Biol., 2002. **323**(3): p. 573-84.

118. Tompa, P., *The interplay between structure and function in intrinsically unstructured proteins.* FEBS Lett., 2005. **579**(15): p. 3346-54.

119. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat. Rev. Mol. Cell. Biol., 2005. **6**(3): p. 197-208.

120. Gustafsson, C., S. Govindarajan, and J. Minshull, *Codon bias and heterologous protein expression.* Trends Biotechnol., 2004. **22**(7): p. 346-53.

121. Medigue, C., et al., *Evidence for horizontal gene transfer in Escherichia coli speciation.* J. Mol. Biol., 1991. **222**(4): p. 851-6.

122. Zhang, S.P., G. Zubay, and E. Goldman, *Low-usage codons in Escherichia coli, yeast, fruit fly and primates.* Gene, 1991. **105**(1): p. 61-72.

123. Hoover, D.M. and J. Lubkowski, *DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis.* Nucleic Acids Res., 2002. **30**(10): p. e43.

124. Lobley, A., L. Whitmore, and B.A. Wallace, *DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra.* Bioinformatics, 2002. **18**(1): p. 211-2.

125. Sreerama, N., S.Y. Venyaminov, and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis.* Anal. Biochem., 2000. **287**(2): p. 243-51.

126. Sreerama, N. and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set.* Anal. Biochem., 2000. **287**(2): p. 252-60.

127. Fontana, A., et al., *Probing the conformational state of apomyoglobin by limited proteolysis.* J. Mol. Biol., 1997. **266**(2): p. 223-30.

128. Arnold, U., et al., *Thermal unfolding and proteolytic susceptibility of ribonuclease A.* Eur. J. Biochem., 1996. **237**(3): p. 862-9.

129. Dyson, H.J. and P.E. Wright, *Unfolded proteins and protein folding studied by NMR.* Chem. Rev., 2004. **104**(8): p. 3607-22.

130. Dedmon, M.M., et al., *FlgM gains structure in living cells.* Proc. Natl. Acad. Sci. U. S. A., 2002. **99**(20): p. 12681-4.

131. Uversky, V.N., *Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule.* Biochemistry, 1993. **32**(48): p. 13288-98.

132. Jones, D.T., *Protein secondary structure prediction based on position-specific scoring matrices.* J. Mol. Biol., 1999. **292**(2): p. 195-202.

133. Cuff, J.A. and G.J. Barton, *Application of multiple sequence alignment profiles to improve protein secondary structure prediction.* Proteins, 2000. **40**(3): p. 502-11.

134. Pollastri, G., et al., *Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles.* Proteins, 2002. **47**(2): p. 228-35.

135. Roccatano, D., et al., *Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study.* Proc. Natl. Acad. Sci. U. S. A., 2002. **99**(19): p. 12179-84.

136. Tiriveedhi, V. and P. Butko, *A fluorescence spectroscopy study on the interactions of the TAT-PTD peptide with model lipid membranes.* Biochemistry, 2007. **46**(12): p. 3888-95.

137. Ladokhin, A.S. and S.H. White, *Interfacial folding and membrane insertion of a designed helical peptide.* Biochemistry, 2004. **43**(19): p. 5782-91.

138. van Duffelen, M., L.R. Chrin, and C.L. Berger, *Nucleotide dependent intrinsic fluorescence changes of W29 and W36 in smooth muscle myosin.* Biophys J., 2004. **87**(3): p. 1767-75.

139. Struppe, J., J.A. Whiles, and R.R. Vold, *Acidic phospholipid bicelles: a versatile model membrane system.* Biophys J., 2000. **78**(1): p. 281-9.

140. Schumann, F.H., et al., *Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions.* J. Biomol. NMR, 2007. **39**(4): p. 275-89.

141. Geyer, M., et al., *Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling.* Nat. Struct. Biol., 1997. **4**(9): p. 694-9.

142. Lockless, S.W. and R. Ranganathan, *Evolutionarily conserved pathways of energetic connectivity in protein families.* Science, 1999. **286**(5438): p. 295-9.

143. Fuentes, E.J., C.J. Der, and A.L. Lee, *Ligand-dependent dynamics and intramolecular signaling in a PDZ domain.* J. Mol. Biol., 2004. **335**(4): p. 1105-15.

144. Wiedemann, U., et al., *Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides.* J. Mol. Biol., 2004. **343**(3): p. 703-18.

145. Remaut, H. and G. Waksman, *Protein-protein interaction through beta-strand addition.* Trends Biochem. Sci., 2006. **31**(8): p. 436-44.

146. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Curr. Opin. Struct. Biol., 2009. **19**(1): p. 31-8.

147. Nourry, C., S.G. Grant, and J.P. Borg, *PDZ domain proteins: plug and play!* Sci. STKE, 2003. **2003**(179): p. RE7.

148. Ikenoue, T., et al., *PTEN acetylation modulates its interaction with PDZ domain.* Cancer Res., 2008. **68**(17): p. 6908-12.

149. Purbey, P.K., et al., *Acetylation-dependent interaction of SATB1 and CtBP1 mediates transcriptional repression by SATB1.* Mol. Cell. Biol., 2009. **29**(5): p. 1321-37.

150. Xue, Y., et al., *GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy.* Mol. Cell. Proteomics, 2008. **7**(9): p. 1598-608.

151. Harris, B.Z., B.J. Hillier, and W.A. Lim, *Energetic determinants of internal motif recognition by PDZ domains.* Biochemistry, 2001. **40**(20): p. 5921-30.

152. White, S.H. and W.C. Wimley, *Membrane protein folding and stability: physical principles.* Annu. Rev. Biophys. Biomol. Struct., 1999. **28**: p. 319-65.

153. Olivotto, M., et al., *Electric fields at the plasma membrane level: a neglected element in the mechanisms of cell signalling.* Bioessays, 1996. **18**(6): p. 495-504.

154. Slavik, J., *Intracellular pH topography: determination by a fluorescent probe.* FEBS Lett., 1983. **156**(2): p. 227-30.

155. Chandra, S., et al., *A broken alpha -helix in folded alpha -Synuclein.* J. Biol. Chem., 2003. **278**(17): p. 15313-8.

156. Spooner, W.A., *Phospholipid phase transitions as revealed by NMR.* Chem. Phys. Lipids, 1991. **57**: p. 195–211.

157. Arora, A., *Biophysical approaches to membrane protein structure determination.* Curr. Opin. Struct. Biol. 2001. **11**: p. 540 –547.

158. Xu, C., et al., *Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif.* Cell, 2008. **135**(4): p. 702-13.

159. Sigalov, A.B., et al., *Lipid-Binding Activity of Intrinsically Unstructured Cytoplasmic Domains of Multichain Immune Recognition Receptor Signaling Subunits.* Biochemistry, 2006. **45**(51): p. 15731-15739.

160. Murray, D., et al., *Electrostatics and the membrane association of Src: theory and experiment.* Biochemistry, 1998. **37**(8): p. 2145-59.

161. Leventis, R. and J.R. Silvius, *Lipid-binding characteristics of the polybasic carboxy-terminal sequence of K-ras4B.* Biochemistry, 1998. **37**(20): p. 7640-8.

162. Michaelson, D., et al., *Differential localization of Rho GTPases in live cells: regulation by hypervariable regions and RhoGDI binding.* J. Cell Biol., 2001. **152**(1): p. 111-26.

163. Nomikos, M., et al., *Binding of phosphoinositide-specific phospholipase C-zeta (PLC-zeta) to phospholipid membranes: potential role of an unstructured cluster of basic residues.* J. Biol. Chem., 2007. **282**(22): p. 16644-53.

164. Yeung, T., et al., *Receptor activation alters inner surface potential during phagocytosis.* Science, 2006. **313**(5785): p. 347-51.

165. Lange, A., et al., *Classical nuclear localization signals: definition, function, and interaction with importin alpha.* J. Biol. Chem., 2007. **282**(8): p. 5101-5.

166. Swierczynski, S.L. and P.J. Blackshear, *Myristoylation-dependent and electrostatic interactions exert independent effects on the membrane association of the myristoylated alanine-rich protein kinase C substrate protein in intact cells.* J. Biol. Chem., 1996. **271**(38): p. 23424-30.

167. Bivona, T.G., et al., *PKC regulates a farnesyl-electrostatic switch on K-Ras that promotes its association with Bcl-XL on mitochondria and induces apoptosis.* Mol. Cell, 2006. **21**(4): p. 481-93.

168. Perisic, O., et al., *Crystal structure of a calcium-phospholipid binding domain from cytosolic phospholipase A2.* J. Biol. Chem., 1998. **273**(3): p. 1596-604.

169. Gitlin, I., J.D. Carbeck, and G.M. Whitesides, *Why are proteins charged? Networks of charge-charge interactions in proteins measured by charge ladders and capillary electrophoresis.* Angew. Chem. Int. Ed. Engl., 2006. **45**(19): p. 3022-60.

170. Sprong, H., P. van der Sluijs, and G. van Meer, *How proteins move lipids and lipids move proteins.* Nat. Rev. Mol. Cell Biol., 2001. **2**(7): p. 504-13.

171. Downes, C.P., A. Gray, and J.M. Lucocq, *Probing phosphoinositide functions in signaling and membrane trafficking.* Trends Cell Biol., 2005. **15**(5): p. 259-68.

172. Barre, P. and D. Eliezer, *Folding of the repeat domain of tau upon binding to lipid surfaces.* J. Mol. Biol., 2006. **362**(2): p. 312-26.

173. Eliezer, D., et al., *Conformational properties of alpha-synuclein in its free and lipid-associated states.* J. Mol. Biol., 2001. **307**(4): p. 1061-73.

174. Ulmer, T.S. and A. Bax, *Comparison of structure and dynamics of micelle-bound human alpha-synuclein and Parkinson disease variants.* J. Biol. Chem., 2005. **280**(52): p. 43179-87.

175. Koo, B.K., et al., *Mind bomb 1 is essential for generating functional Notch ligands to activate Notch.* Development, 2005. **132**(15): p. 3459-70.

176. Simanek, E.E., et al., *Glycosylation of threonine of the repeating unit of RNA polymerase II with β-linked N-acetylglucosame leads to a turn-like structure.* J. Am. Chem. Soc., 1998. **120**: p. 11567-11575.

177. Wells, L., K. Vosseller, and G.W. Hart, *Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc.* Science, 2001. **291**(5512): p. 2376-8.

178. Shoemaker, B.A., J.J. Portman, and P.G. Wolynes, *Speeding molecular recognition by using the folding funnel: the fly-casting mechanism.* Proc. Natl. Acad. Sci. U. S. A., 2000. **97**(16): p. 8868-73.

179. Zeev-Ben-Mordehai, T., et al., *The intracellular domain of the Drosophila cholinesterase-like neural adhesion protein, gliotactin, is natively unfolded.* Proteins, 2003. **53**(3): p. 758-67.

180. Aivazian, D. and L.J. Stern, *Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition.* Nat. Struct. Biol., 2000. **7**(11): p. 1023-6.

181. Duchardt, E., et al., *Structure Induction of the T-Cell Receptor zeta-Chain upon Lipid Binding Investigated by NMR Spectroscopy.* Chembiochem, 2007. **8**(7): p. 820-827.

182. Sigalov, A., D. Aivazian, and L. Stern, *Homooligomerization of the cytoplasmic domain of the T cell receptor zeta chain and of other proteins containing the immunoreceptor tyrosine-based activation motif.* Biochemistry, 2004. **43**(7): p. 2049-61.

183. Kim, P.W., et al., *A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8.* Science, 2003. **301**(5640): p. 1725-8.

184. Tompa, P., *Intrinsically unstructured proteins.* Trends Biochem. Sci., 2002. **27**(10): p. 527-33.

185. Uversky, V.N., *What does it mean to be natively unfolded?* Eur. J. Biochem., 2002. **269**(1): p. 2-12.

186. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins, 2000. **41**(3): p. 415-27.

187. Uversky, V.N., *Natively unfolded proteins: a point where biology waits for physics.* Protein Sci., 2002. **11**(4): p. 739-56.

188. Sonnichsen, F.D., et al., *Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide.* Biochemistry, 1992. **31**(37): p. 8790-8.

189. Yang, J.J., M. Pikeathly, and S.E. Radford, *Far-UV circular dichroism reveals a conformational switch in a peptide fragment from the beta-sheet of hen lysozyme.* Biochemistry, 1994. **33**(23): p. 7345-53.

190. Najbar, L.V., et al., *CD and NMR determination of the solution structure of a peptide corresponding to T4 lysozyme residues 38-51.* Biochim. Biophys. Acta, 1995. **1250**(2): p. 163-70.

191. Martenson, R.E., J.Y. Park, and A.L. Stone, *Low-ultraviolet circular dichroism spectroscopy of sequential peptides 1-63, 64-95, 96-128, and 129-168 derived from myelin basic protein of rabbit.* Biochemistry, 1985. **24**(26): p. 7689-95.

192. Wang, J., R.S. Hodges, and B.D. Sykes, *Effect of trifluoroethanol on the solution structure and flexibility of desmopressin: a two-dimensional NMR study.* Int. J. Pept. Protein Res., 1995. **45**(5): p. 471-81.

193. Mouillon, J.M., P. Gustafsson, and P. Harryson, *Structural investigation of disordered stress proteins. Comparison of full-length dehydrins with isolated peptides of their conserved segments.* Plant Physiol., 2006. **141**(2): p. 638-50.

194. Semisotnov, G.V., *Protein globularization during folding. A study by synchrotron small-angle X-ray scattering.* J. Mol. Biol., 1996 **262**: p. 559-574.

195. Haris, P., *FTIR spectroscopic characterization of protein structure in acqueous and non-acqueous media.* J. Mol. Catal. B-Enzym. 1999 **7**: p. 207–221

196. Serber, Z. and V. Dotsch, *In-cell NMR spectroscopy.* Biochemistry, 2001. **40**(48): p. 14317-23.

197. Reckel, S., F. Lohr, and V. Dotsch, *In-cell NMR spectroscopy.* Chembiochem, 2005. **6**(9): p. 1601-6.

198. Selenko, P. and G. Wagner, *Looking into live cells with in-cell NMR spectroscopy.* J. Struct. Biol., 2007. **158**(2): p. 244-53.

199. McNulty, B.C., G.B. Young, and G.J. Pielak, *Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder.* J. Mol. Biol., 2006. **355**(5): p. 893-7.

200. Selenko, P., et al., *Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes.* Proc. Natl. Acad. Sci. U. S. A., 2006. **103**(32): p. 11904-9.

201. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm.* J. Mol. Biol., 1999. **293**(2): p. 321-31.

202. Cortese, M.S., V.N. Uversky, and A.K. Dunker, *Intrinsic disorder in scaffold proteins: getting more from less.* Prog. Biophys. Mol. Biol., 2008. **98**(1): p. 85-106.

203. Harrison, S.C., *Peptide-surface association: the case of PDZ and PTB domains.* Cell, 1996. **86**(3): p. 341-3.

204. Zuiderweg, E.R., *Mapping protein-protein interactions in solution by NMR spectroscopy.* Biochemistry, 2002. **41**(1): p. 1-7.

205. Rajagopal, P., et al., *Demonstration of protein-protein interaction specificity by NMR chemical shift mapping.* Protein Sci., 1997. **6**(12): p. 2624-7.

206. Vaynberg, J. and J. Qin, *Weak protein-protein interactions as probed by NMR spectroscopy.* Trends Biotechnol., 2006. **24**(1): p. 22-7.

207. Sykes, B.D., *Determination of the conformations of bound peptides using NMR-transferred nOe techniques.* Curr. Opin. Biotechnol., 1993. **4**(4): p. 392-6.
208. Piserchio, A., et al., *The PDZ1 domain of SAP90. Characterization of structure and binding.* J. Biol. Chem., 2002. **277**(9): p. 6967-73.

209. Spera, S., *Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment..* J Biomol. NMR 1991. **2**: p. 155-65.

210. Pascal, S.M., *Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide.* Cell, 1994. **77(3)**

211. Niv, M.Y. and H. Weinstein, *A flexible docking procedure for the exploration of peptide binding selectivity to known structures and homology models of PDZ domains.* J. Am. Chem. Soc., 2005. **127**(40): p. 14072-9.

212. Breeze, A.L., *Isotope-filtered NMR methods for the study of biomolecular structure and interactions.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2000. **36**: p. 323–372.

213. Horovitz, A., *Double-mutant cycles: a powerful tool for analyzing protein structure and function.* Fold. Des., 1996. **1**(6): p. R121-6.

214. Cooper, A. and D.T. Dryden, *Allostery without conformational change. A plausible model.* Eur. Biophys. J., 1984. **11**(2): p. 103-9.

215. Karthikeyan, S., et al., *Crystal structure of the PDZ1 domain of human Na(+)/H(+) exchanger regulatory factor provides insights into the mechanism of carboxyl-terminal leucine recognition by class I PDZ domains.* J. Mol. Biol., 2001. **308**(5): p. 963-73.

216. Birrane, G., J. Chung, and J.A. Ladias, *Novel mode of ligand recognition by the Erbin PDZ domain.* J. Biol. Chem., 2003. **278**(3): p. 1399-402.

217. Appleton, B.A., et al., *Comparative structural analysis of the Erbin PDZ domain and the first PDZ domain of ZO-1. Insights into determinants of PDZ domain specificity.* J. Biol. Chem., 2006. **281**(31): p. 22312-20.

218. Hung, A.Y. and M. Sheng, *PDZ domains: structural modules for protein complex assembly.* J. Biol. Chem., 2002. **277**(8): p. 5699-702.

219. Kurakin, A., et al., *The PDZ domain as a complex adaptive system.* PLoS One, 2007. **2**(9): p. e953.

220. Cao, T.T., et al., *A kinase-regulated PDZ-domain interaction controls endocytic sorting of the beta2-adrenergic receptor.* Nature, 1999. **401**(6750): p. 286-90.

221. Chetkovich, D.M., et al., *Phosphorylation of the postsynaptic density-95 (PSD-95)/discs large/zona occludens-1 binding site of stargazin regulates binding to PSD-95 and synaptic targeting of AMPA receptors.* J. Neurosci., 2002. **22**(14): p. 5791-6.

222. Tian, Q.B., et al., *Interaction of LDL receptor-related protein 4 (LRP4) with postsynaptic scaffold proteins via its C-terminal PDZ domain-binding motif, and its regulation by Ca/calmodulin-dependent protein kinase II.* Eur. J. Neurosci., 2006. **23**(11): p. 2864-76.

223. Sulka, B., et al., *Tyrosine dephosphorylation of the syndecan-1 PDZ binding domain regulates syntenin-1 recruitment.* J. Biol. Chem., 2009. **284**(16): p. 10659-71.

224. Boisguerin, P., et al., *Characterization of a putative phosphorylation switch: adaptation of SPOT synthesis to analyze PDZ domain regulation mechanisms.* Chembiochem, 2007. **8**(18): p. 2302-7.

225. Mann, M., et al., *Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome.* Trends. Biotechnol., 2002. **20**(6): p. 261-8.

226. Blom, N., et al., *Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence.* Proteomics, 2004. **4**(6): p. 1633-49.

227. Xue, Y., et al., *GPS: a comprehensive www server for phosphorylation sites prediction.* Nucleic Acids Res., 2005. **33**(Web Server issue): p. W184-7.

228. Xue, Y., et al., *PPSP: prediction of PK-specific phosphorylation site with Bayesian decision theory.* BMC Bioinformatics, 2006. **7**: p. 163.

229. Obenauer, J.C., L.C. Cantley, and M.B. Yaffe, *Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs.* Nucleic Acids Res., 2003. **31**(13): p. 3635-41.

230. Huang, H.D., et al., *KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites.* Nucleic Acids Res., 2005. **33**(Web Server issue): p. W226-9.

231. Kim, J.H., et al., *Prediction of phosphorylation sites using SVMs.* Bioinformatics, 2004. **20**(17): p. 3179-84.

232. Brinkworth, R.I., R.A. Breinl, and B. Kobe, *Structural basis and prediction of substrate specificity in protein serine/threonine kinases.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(1): p. 74-9.

233. Li, T., F. Li, and X. Zhang, *Prediction of kinase-specific phosphorylation sites with sequence features by a log-odds ratio approach.* Proteins, 2008. **70**(2): p. 404-14.

234. Neuberger, G., G. Schneider, and F. Eisenhaber, *pkaPS: prediction of protein kinase A phosphorylation sites with the simplified kinase-substrate binding model.* Biol. Direct, 2007. **2**: p. 1.

235. Kalay, E., et al., *A novel D458V mutation in the SANS PDZ binding motif causes atypical Usher syndrome.* J. Mol. Med., 2005. **83**(12): p. 1025-32.

236. Chen, Y.H., J.T. Yang, and K.H. Chau, *Determination of the helix and beta form of proteins in aqueous solution by circular dichroism.* Biochemistry, 1974. **13**(16): p. 3350-9.

237. Hwang, T.L. and A.J. Shaka, *Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients.* J. Magn. Reson., 1995. **112**: p. 275-279.

238. Serber, Z., et al., *Evaluation of parameters critical to observing proteins inside living Escherichia coli by in-cell NMR spectroscopy.* J. Am. Chem. Soc., 2001. **123**(37): p. 8895-901.

239. Serber, Z., et al., *In-cell NMR spectroscopy.* Methods Enzymol., 2005. **394**: p. 17-41.

240. Chen, Q., et al., *Solution structure and backbone dynamics of the AF-6 PDZ domain/Bcr peptide complex.* Protein Sci., 2007. **16**(6): p. 1053-62.

241. Wintjens, R., et al., *1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides.* J. Biol. Chem., 2001. **276**(27): p. 25150-6.

242. Woody, R.W., *Circular dichroism.* Methods Enzymol., 1995 p. 24634-71.

243. Sreerama, N. and R.W. Woody, *Computation and analysis of protein circular dichroism spectra.* Methods Enzymol., 2004. **383**: p. 318-51.

244. Lakowicz, J., *Principles of Fluorescence Spectroscopy*. Second Edition ed. 1999, New York: Kluwer Academic/Plenum Publishers.

245. Schasfoort, R.B.M.a.T., A.J. , *Handbook of Surface Plasmon Resonance*. 2008, Cambridge: RSCPublishing.

246. Jonsson, U., et al., *Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology.* Biotechniques, 1991. **11**(5): p. 620-7.

247. Day, Y.S., et al., *Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods.* Protein Sci., 2002. **11**(5): p. 1017-25.

248. Cavanagh, J., *Protein NMR Spectroscopy: Principles and Practice* 1996, San Diego: Academic press.

249. Wuthrich, K., *NMR of proteins and nucleic acids*. First ed. 1986, New York: Wiley.

250. Forsen, S. and J. Kordel, *Biomolecular structure and dynamics--experiment and theory.* J. Pharm. Biomed. Anal., 1996. **14**(3): p. 233-6.

251. Ishima, R. and D.A. Torchia, *Protein dynamics from NMR.* Nat. Struct. Biol., 2000. **7**(9): p. 740-3.

252. Wand, A.J., *Dynamic activation of protein function: a view emerging from NMR spectroscopy.* Nat Struct Biol, 2001. **8**(11): p. 926-31.

253. Kay, L.E., *Protein dynamics from NMR.* Nat. Struct. Biol., 1998. **5 Suppl**: p. 513-7.

254. Takahashi, H., et al., *A novel NMR method for determining the interfaces of large protein-protein complexes.* Nat. Struct. Biol., 2000. **7**(3): p. 220-3.

255. Post, C.B., *Exchange-transferred NOE spectroscopy and bound ligand structure determination.* Curr. Opin. Struct. Biol., 2003. **13**(5): p. 581-8.

256. Pellecchia, M., *Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions.* Chem. Biol., 2005. **12**(9): p. 961-71.

257. Inomata, K., et al., *High-resolution multi-dimensional NMR spectroscopy of proteins in human cells.* Nature, 2009. **458**(7234): p. 106-9.

258. Sakakibara, D., et al., *Protein structure determination in living cells by in-cell NMR spectroscopy.* Nature, 2009. **458**(7234): p. 102-5.

259. Lehninger, A.L., *Principles of Biochemistry*. 1997, New York: Worth Publishers, Inc., NY, USA,. 240-297.

260. Peric, M., M. Alves, and B.L. Bales, *Combining precision spin-probe partitioning with time-resolved fluorescence quenching to study micelles. Application to micelles of pure lysomyristoylphosphatidylcholine (LMPC) and LMPC mixed with sodium dodecyl sulfate.* Chem. Phys. Lipids, 2006. **142**(1-2): p. 1-13.

261. Gershfeld, N.L., *The critical unilamellar lipid state: a perspective for membrane bilayer assembly.* Biochim. Biophys. Acta, 1989. **988**(3): p. 335-50.

262. Woodle, M.C. and D. Papahadjopoulos, *Liposome preparation and size characterization.* Methods Enzymol., 1989. **171**: p. 193-217.

263. Sanders, C.R., 2nd and G.C. Landis, *Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies.* Biochemistry, 1995. **34**(12): p. 4030-40.

264. Seddon, A.M., P. Curnow, and P.J. Booth, *Membrane proteins, lipids and detergents: not just a soap opera.* Biochim. Biophys. Acta, 2004. **1666**(1-2): p. 105-17.

265. Glover, K.J., et al., *Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules.* Biophys J., 2001. **81**(4): p. 2163-71.

266. Prosser, R.S., et al., *Current applications of bicelles in NMR studies of membrane-associated amphiphiles and proteins.* Biochemistry, 2006. **45**(28): p. 8453-65.

267. Marcotte, I., *Bicelles as Model Membranes for Solidand Solution-State NMR Studies of Membrane Peptides and Proteins.* Concepts in Magnetic Resonance Part A, 2005. **24A**: p. 17-37.

268. Borch, J. and T. Hamann, *The nanodisc: a novel tool for membrane protein studies.* Biol. Chem., 2009. **390**(8): p. 805-14.

269. Gluck, J.M., et al., *Integral membrane proteins in nanodiscs can be studied by solution NMR spectroscopy.* J. Am. Chem. Soc., 2009. **131**(34): p. 12060-1.

Supplementary material

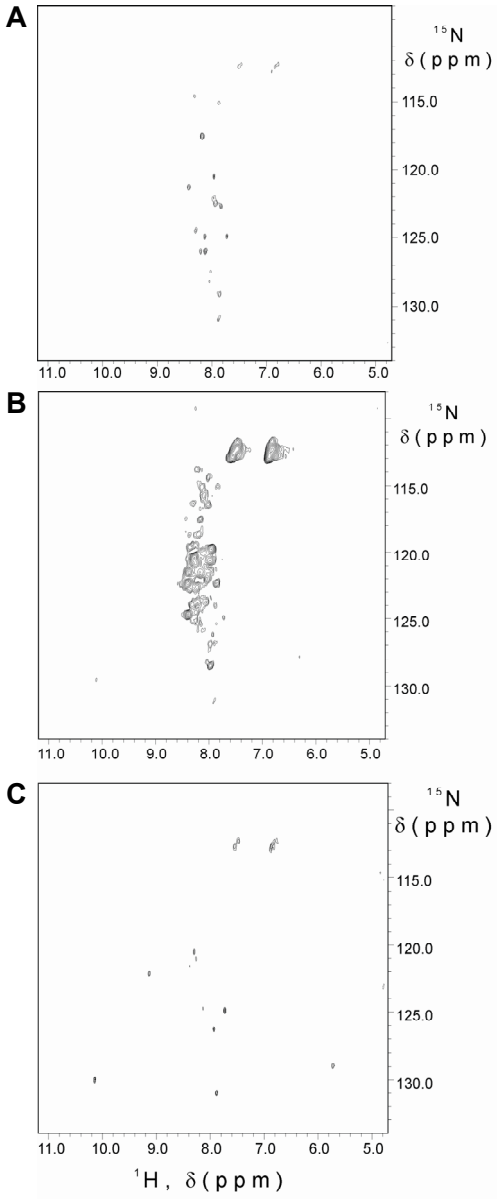


Figure S1. In-cell NMR. ^{1}H-^{15}N HSQC spectra of the *E. coli* slurry ((**A**), not induced; (**B**), induced) and of (**C**) the supernatant of the induced culture. All spectra were recorded in identical conditions and plotted at the same contour level.

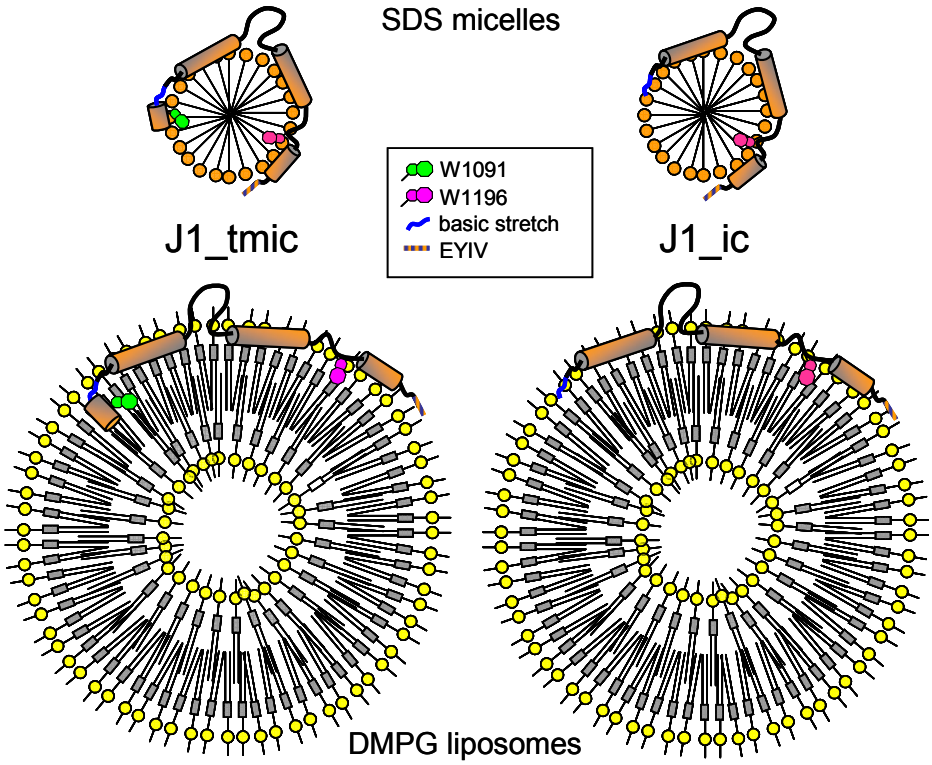


Figure S2. Mode of interaction with micelles and liposomes. Proposed model of the interaction of J1_ic and J1_tmic with micelles and liposomes. The two proteins are sketched as cartoons, with α-helices shown as orange cylinders. The content and the localization of α-helices in the proteins are hypothetical.

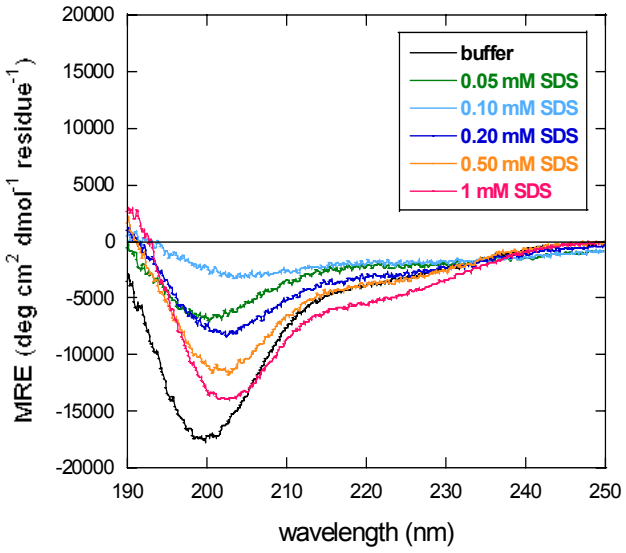


Figure S3. Circular dichroism in the presence of SDS. Far-UV CD spectra of J1_tmic (7.5 μM) in 5 mM Tris⁄HCl buffer, pH 7.4 at different concentrations of SDS (from 0.05 mM to 1 mM).

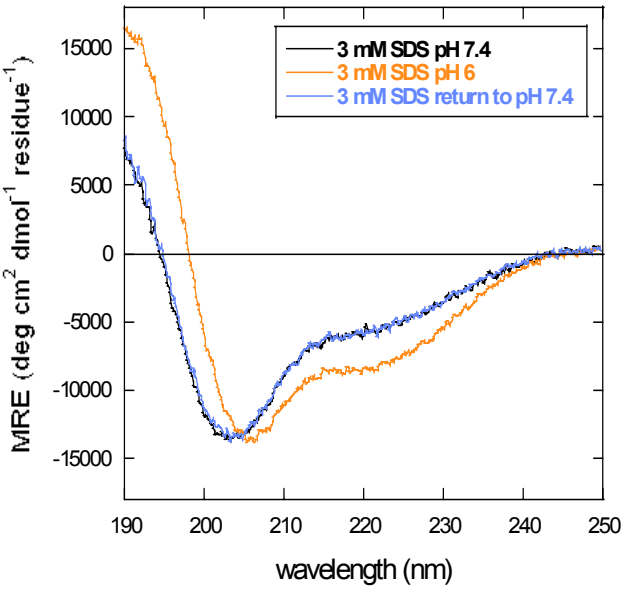


Figure S4. Circular dichroism in the presence of SDS – reversible pH effect. Far-UV CD spectra of J1_tmic (7.5 µM in 5 mM Tris⁄HCl buffer) in the presence of SDS (3 mM) at pH 7.4, at pH 6.0 and back to pH 7.4.

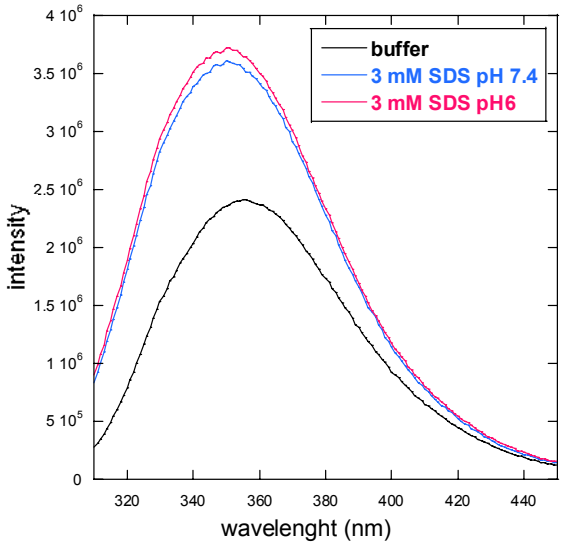


Figure S5. Fluorescence spectroscopy in the presence of SDS – pH effect. Tryptophan fluorescence emission spectra of J1_tmic (7.5 µM) in 5mM Tris⁄HCl buffer, pH 7.4, in the presence of SDS (3 mM) at pH 7.4 and at pH 6 (excitation wavelength was set to 295 nm).

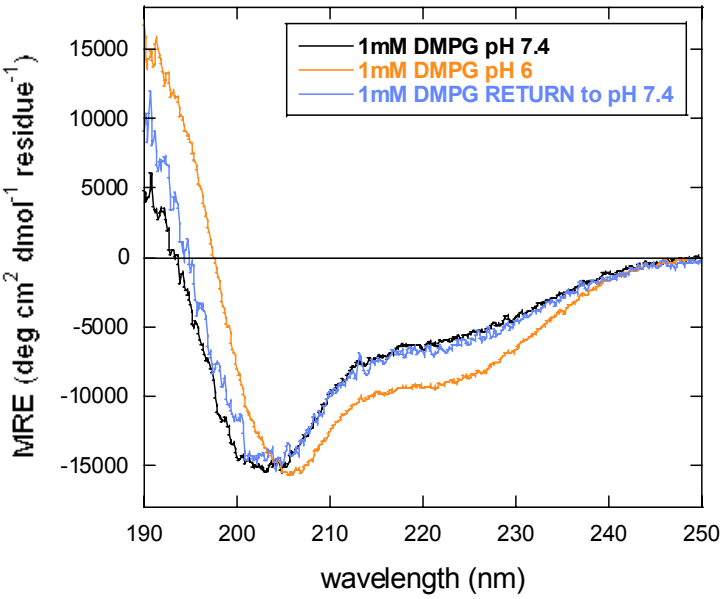


Figure S6. Circular dichroism in the presence of DMPG liposomes – reversible pH effect. Far-UV CD spectra of J1_tmic (7.5 µM in 5 mM Tris HCl buffer) in the presence of LMPG (1 mM) at pH 7.4, at pH 6.0 and back to pH 7.4.

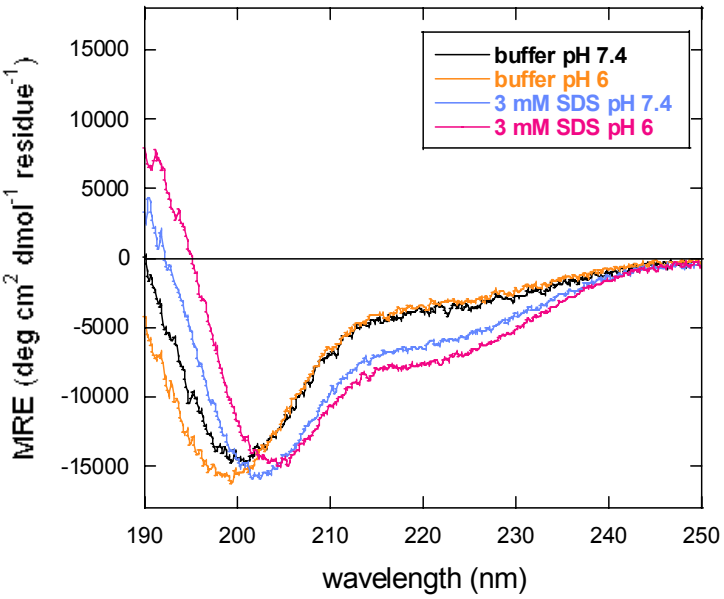


Figure S7. CD spectra of J1_ic in the presence of SDS. Far-UV CD spectra of J1_ic (7 µM) in 5 mM Tris⁄HCl buffer, at different pH values (7.4 and 6.0) in buffer alone and in the presence of SDS (3 mM).

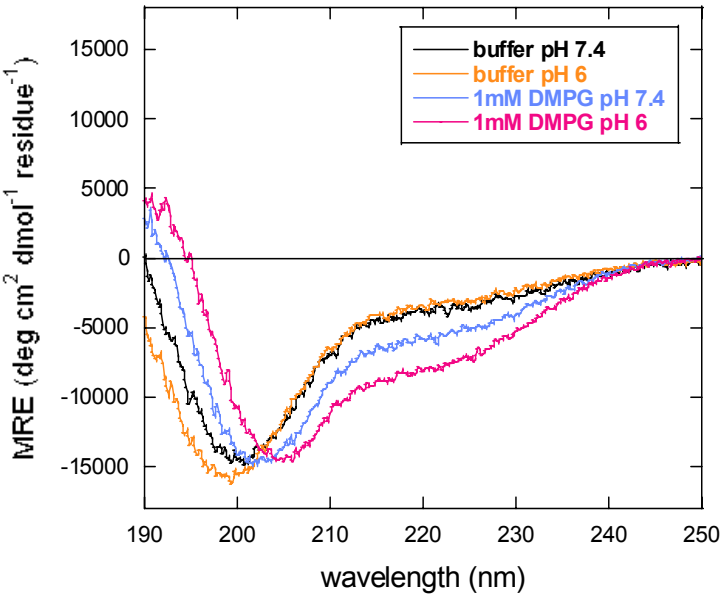


Figure S8. CD spectra of J1_ic in the presence of DMPG. Far-UV CD spectra of J1_ic (7 μM) in 5 mM Tris⁄HCl buffer, at different pH values (7.4 and 6.0) in buffer alone and in the presence of SDS (1 mM).

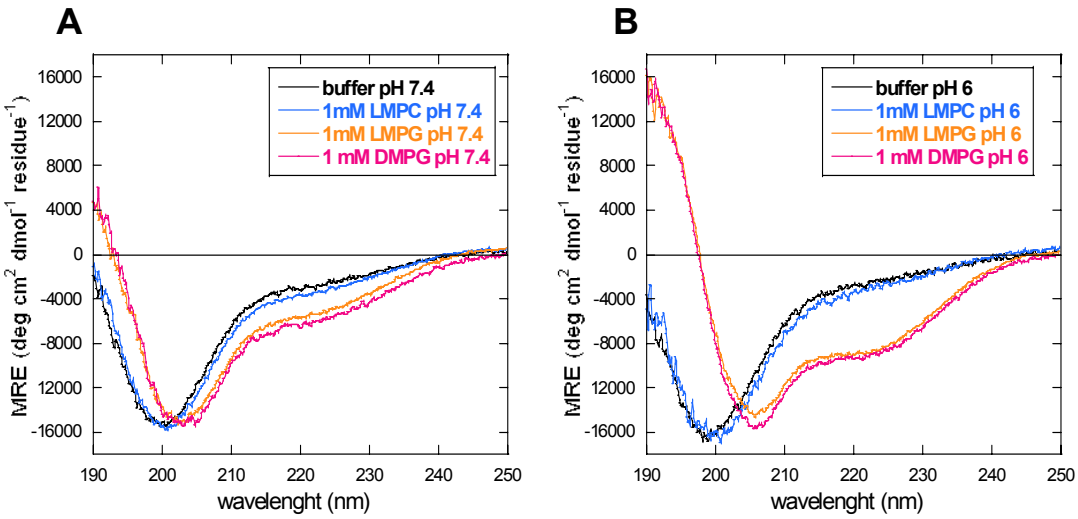


Figure S9. Circular dichroism in the presence of phospholipid micelles. Far-UV CD spectra of J1_tmic (7.5 μM) in the presence of lysophospholipid micelles (1 mM LMPC or 1 mM LMPG) and phospholipid vesicles (1mM DMPG) at (**A**) pH 7.4 and (**B**) pH 6.0.

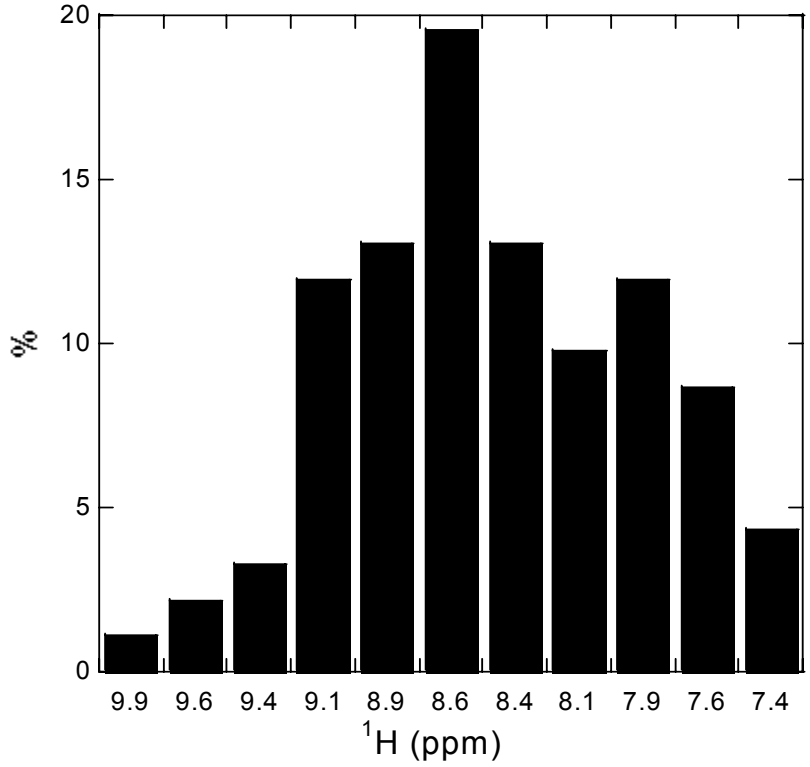


Figure S10. NMR of AF6_PDZ. Chemical shift dispersion of detectable ^{1}HNs of AF6_PDZ. Chemical shift values were estimated from the ^{1}H-^{15}N HSQC spectrum and referenced to DSS.

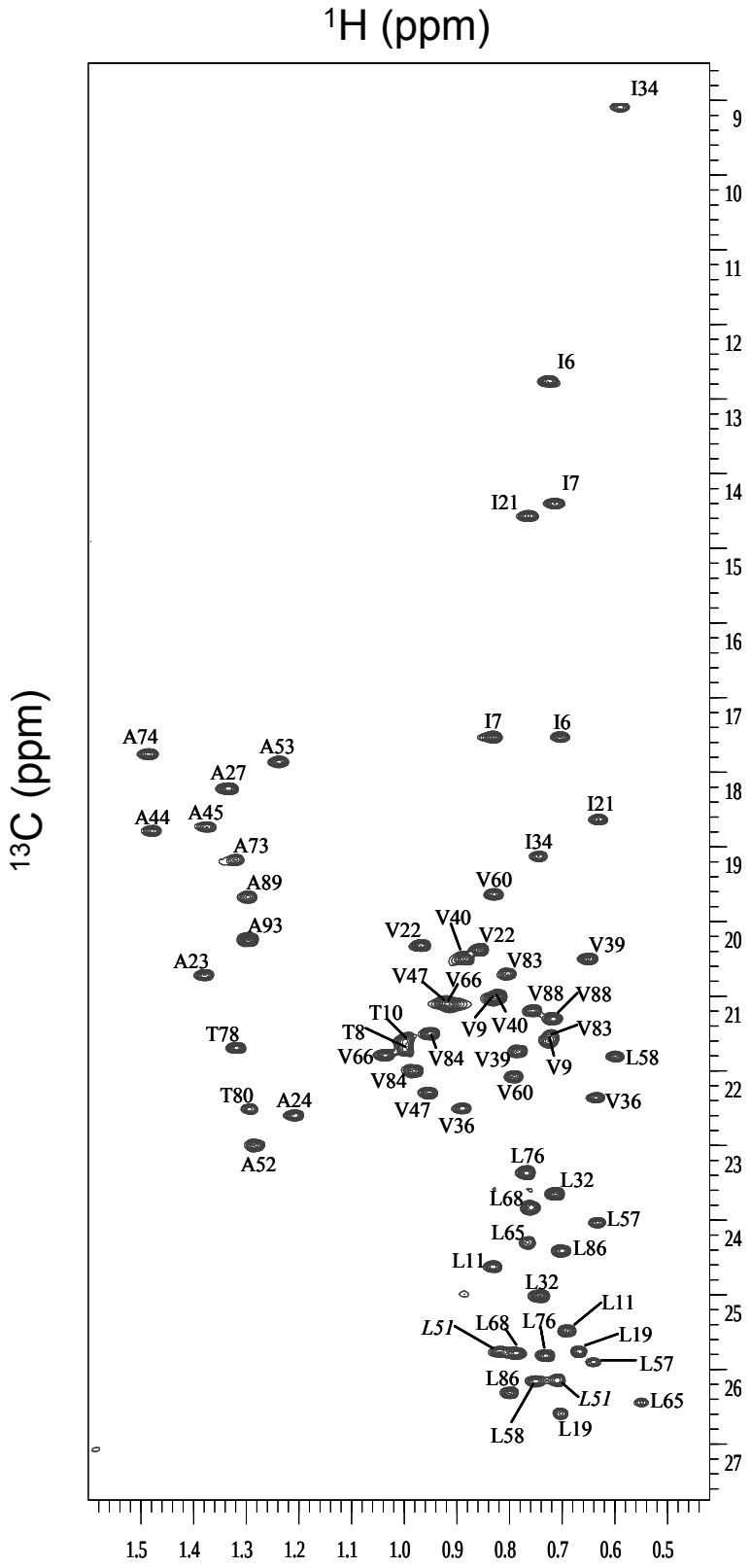


Figure S11. NMR of AF6_PDZ. Methyl region of the ^{1}H-^{13}C HSQC spectrum of AF6_PDZ, including the isoleucine $C^{\delta 1}(H^{\delta 1})_3$ cross- peaks.

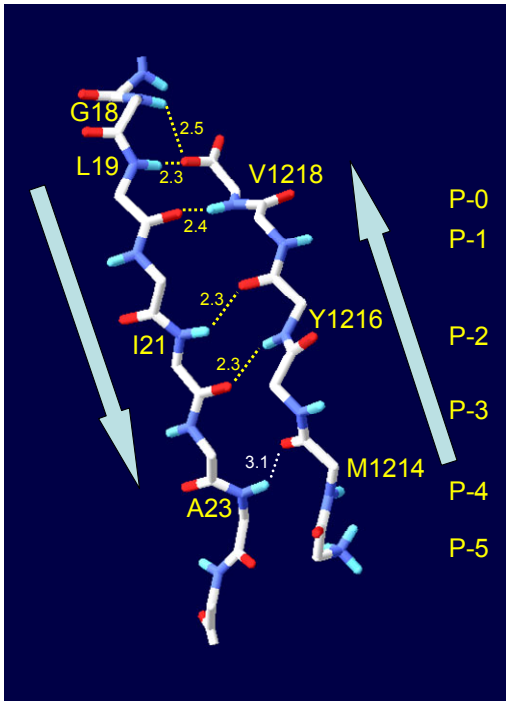


Figure S12. Homology model of J1C6 bound to AF6_PDZ. The model of the RMEYIV peptide (J1C6) was built using Modeller. The NMR-derived structure of the complex between afadin PDZ domain and the LFSTEV peptide was used as as template (PDB: 2AIN). Only the βB strand of the PDZ and the backbone of the peptide are shown. Calculated hydrogen bonds are shown as yellow dashed lines (distance in Å). The short distance between the HN of A23 and the CO of M1214 is shown as a white dashed line.

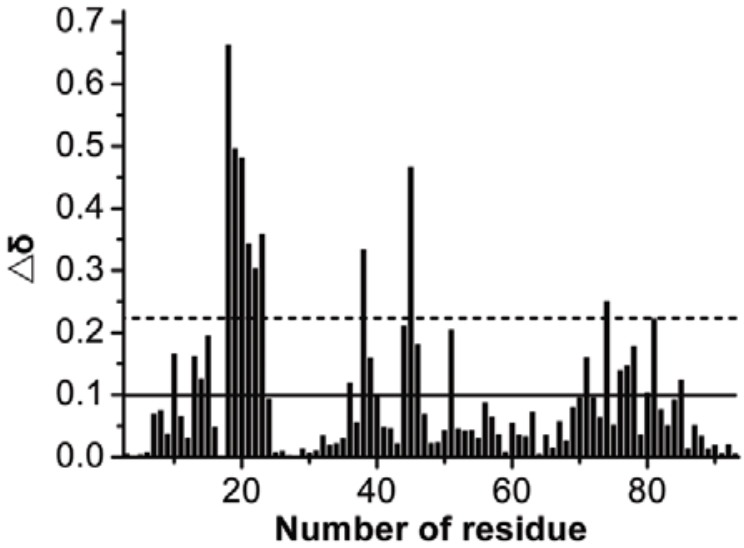


Figure S13. Chemical shift mapping. CSP plot of backbone amides of AF6 PDZ domain in the presence of the C-terminal peptide of Neurexin (sequence: LFSTEV). Adapted from (2005JBC_Zhou).

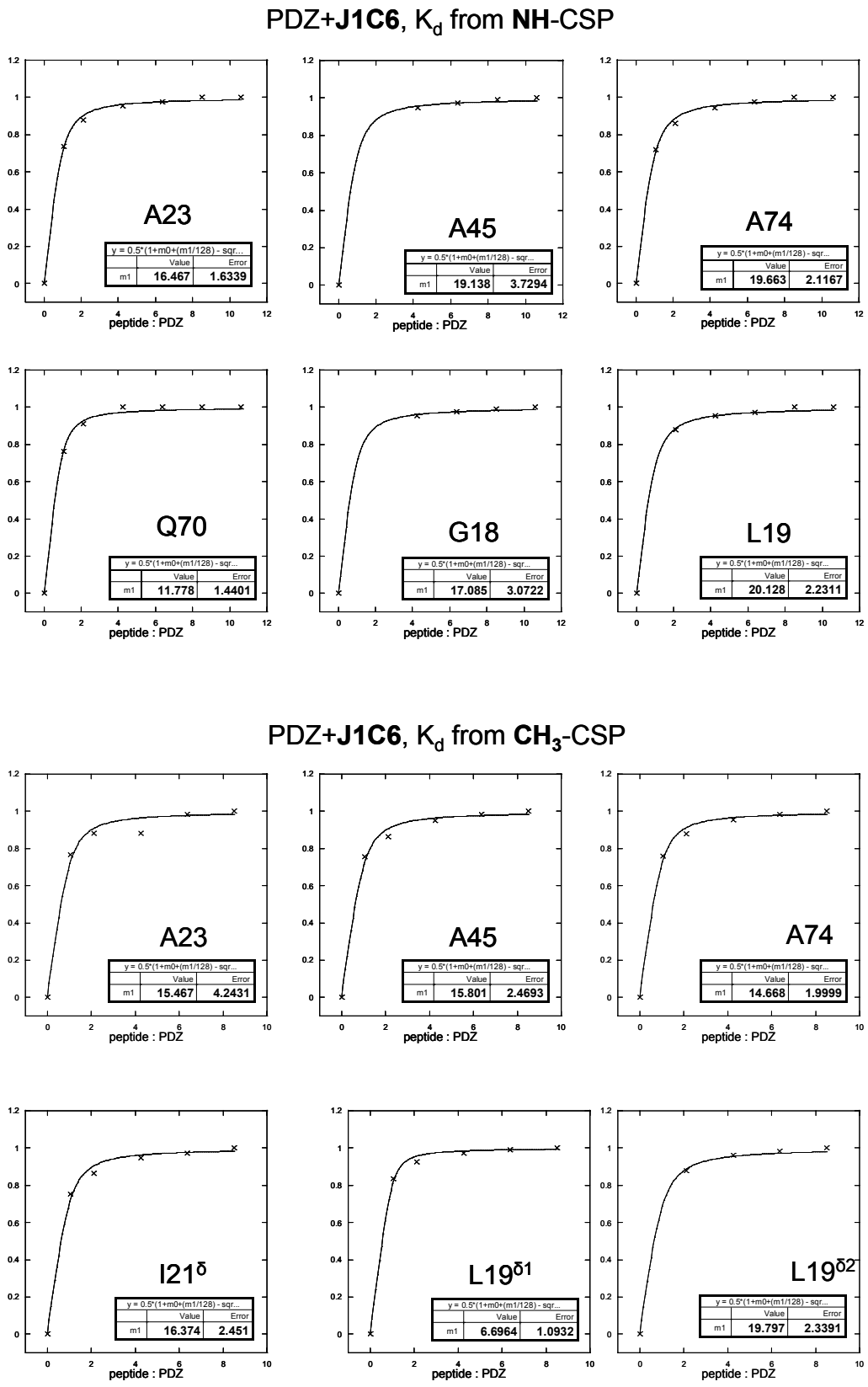
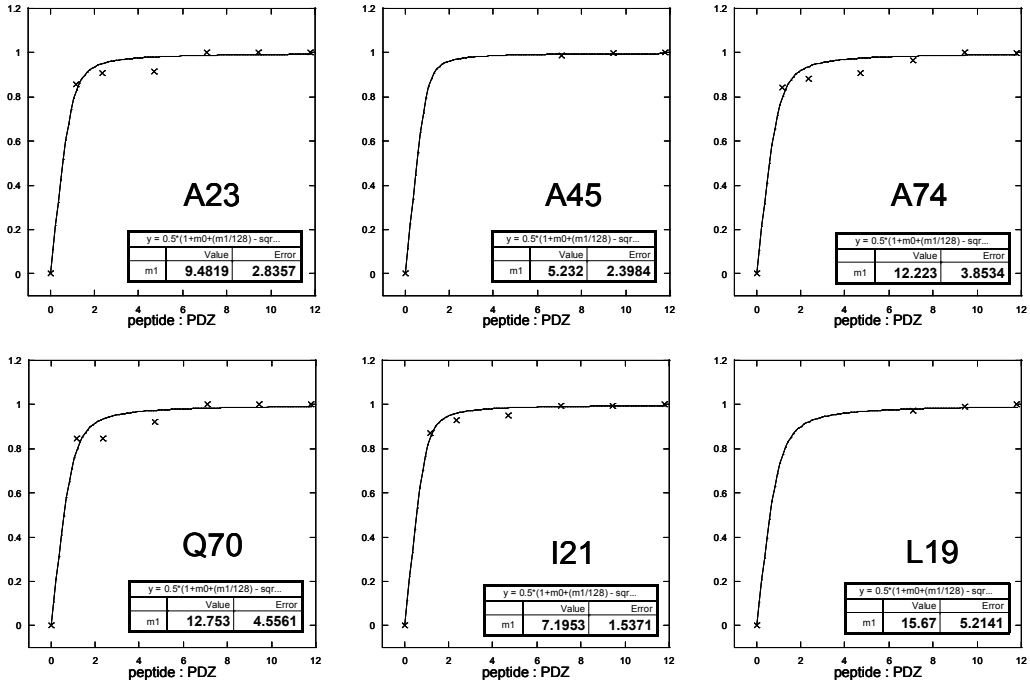


Figure S14. Determination of dissociation constants from NMR

PDZ+**J1C24**, K_d from **NH**-CSP



PDZ+**J1C24**, K_d from **CH_3**-CSP

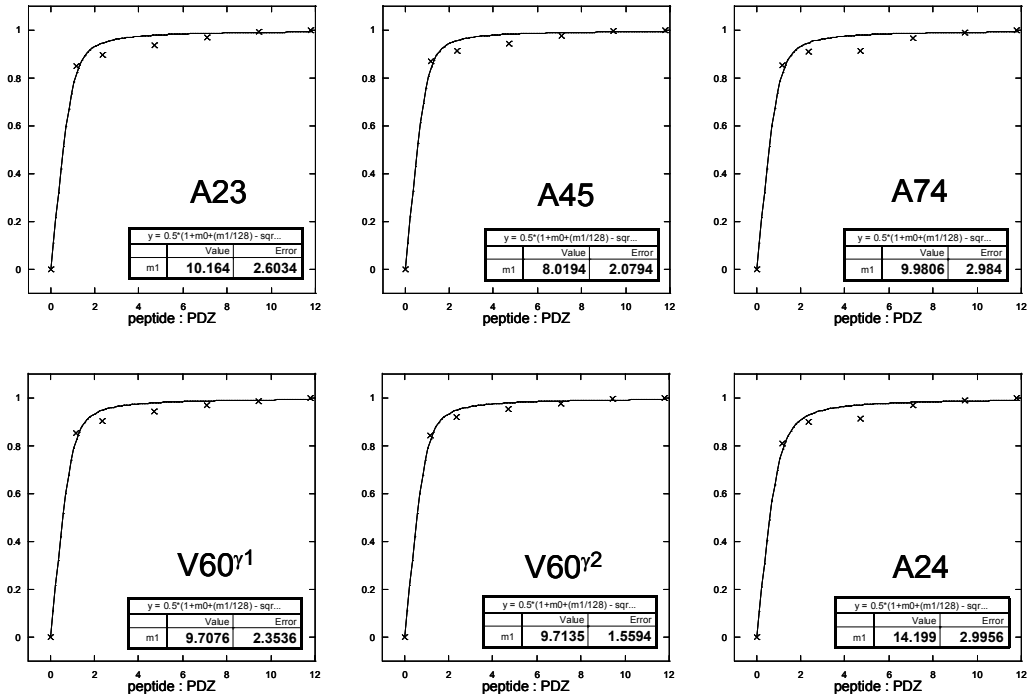
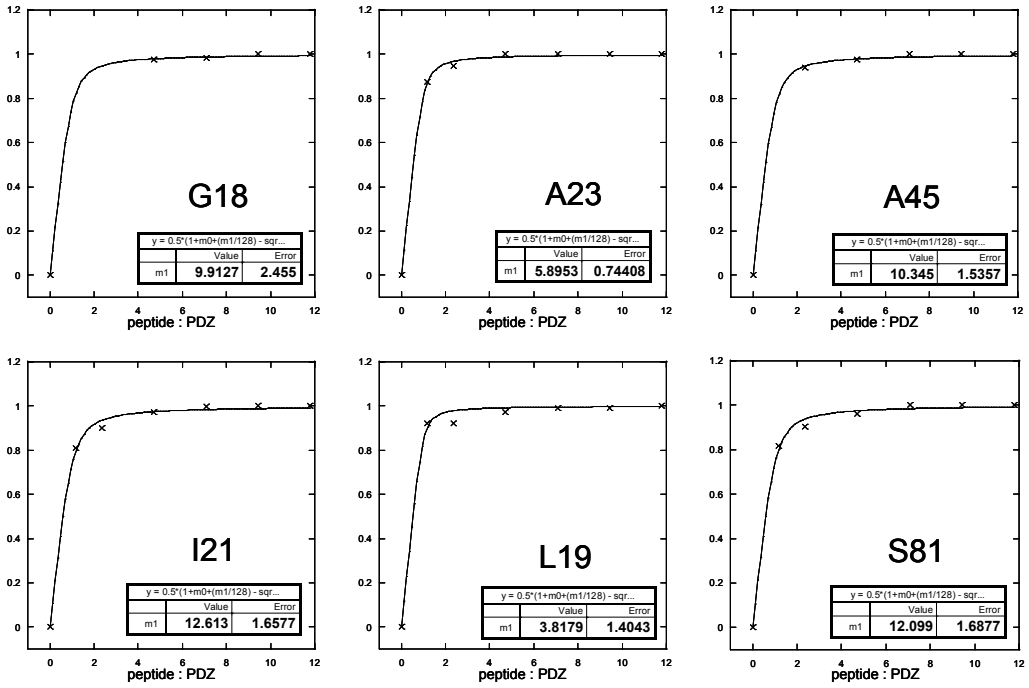


Figure S14. Determination of dissociation constants from NMR (continued)

PDZ+**J1C24pS**, K_d from **NH**-CSP



PDZ+**J1C24pS**, K_d from **CH_3**-CSP

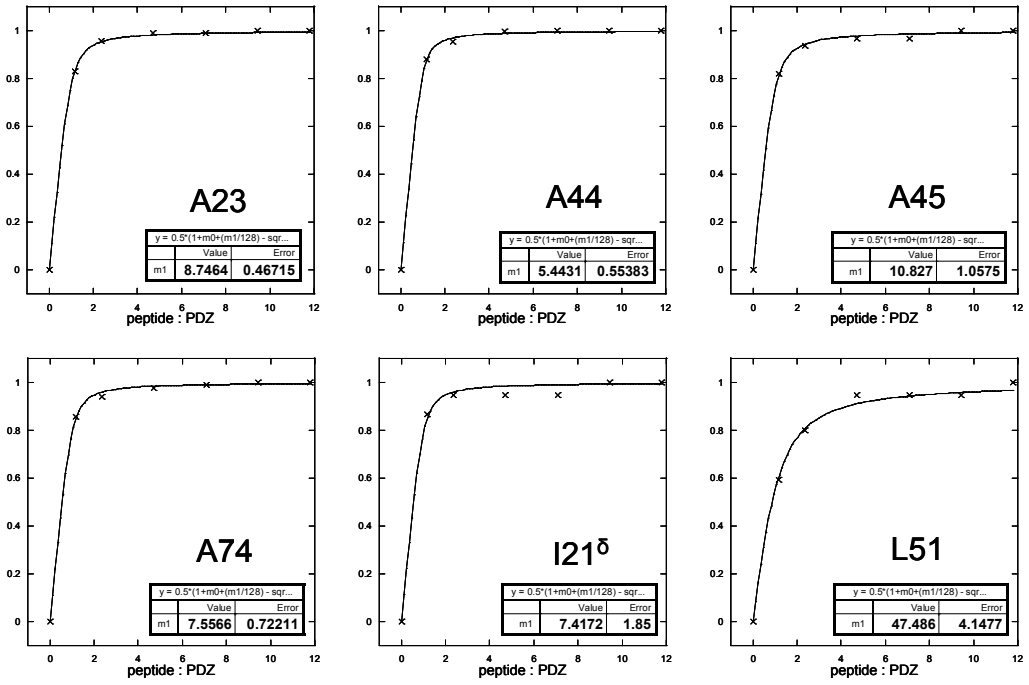
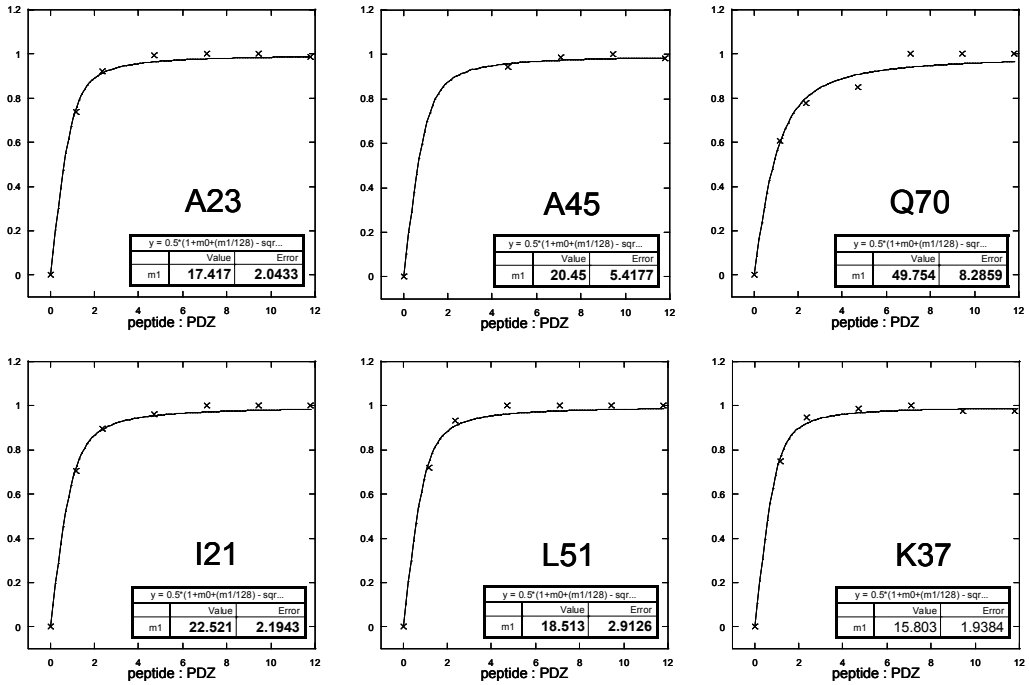


Figure S14. Determination of dissociation constants from NMR (continued)

PDZ+**J1C24pSpY**, K_d from **NH**-CSP



PDZ+**J1C24pSpY**, K_d from **CH_3**-CSP

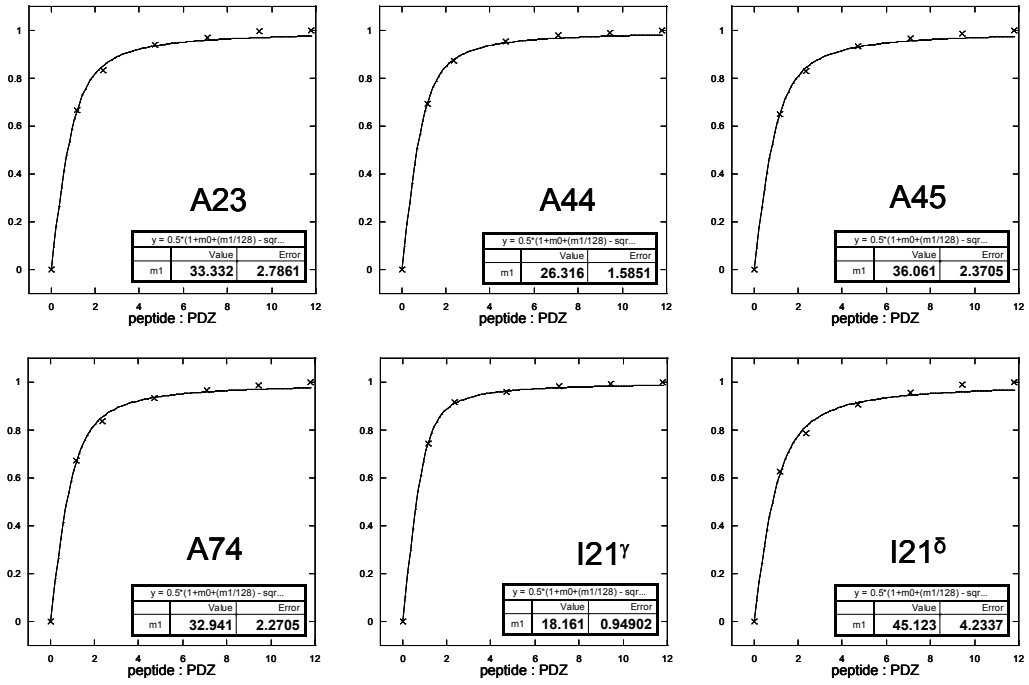
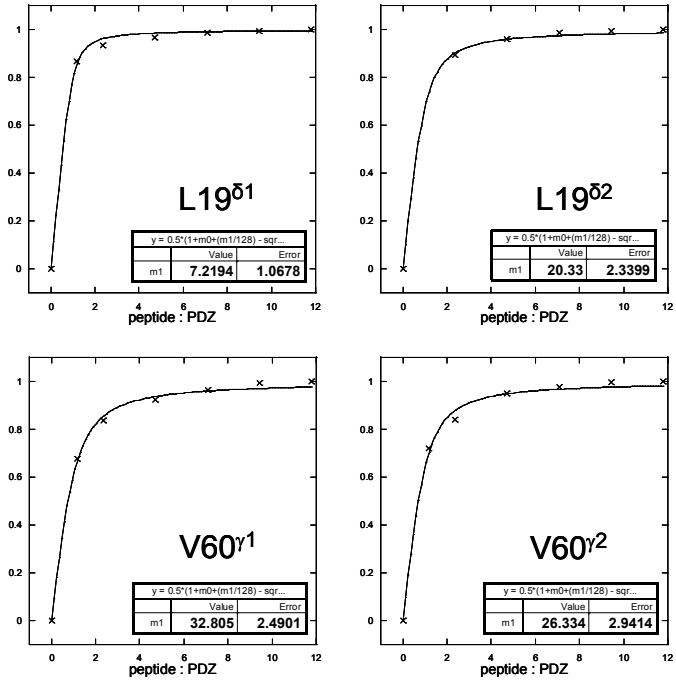


Figure S14. Determination of dissociation constants from NMR (continued)

PDZ+**J1C24pSpY**, K_d from **CH_3**-CSP



PDZ+J**1C24RQ**, K_d from **NH**-CSP

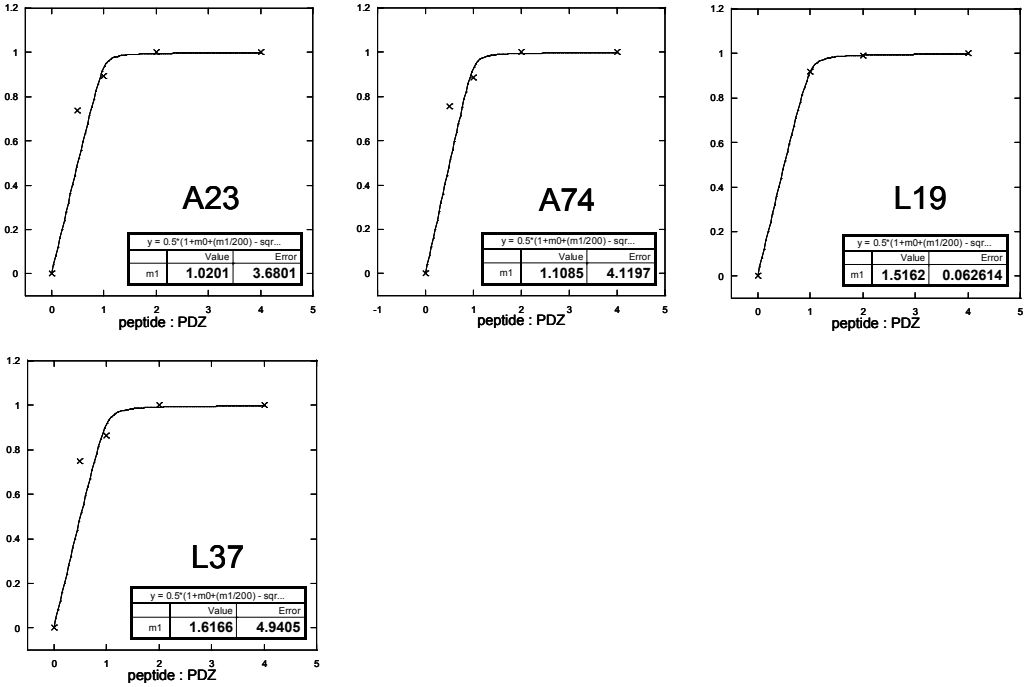


Figure S14. Determination of dissociation constants from NMR (continued)

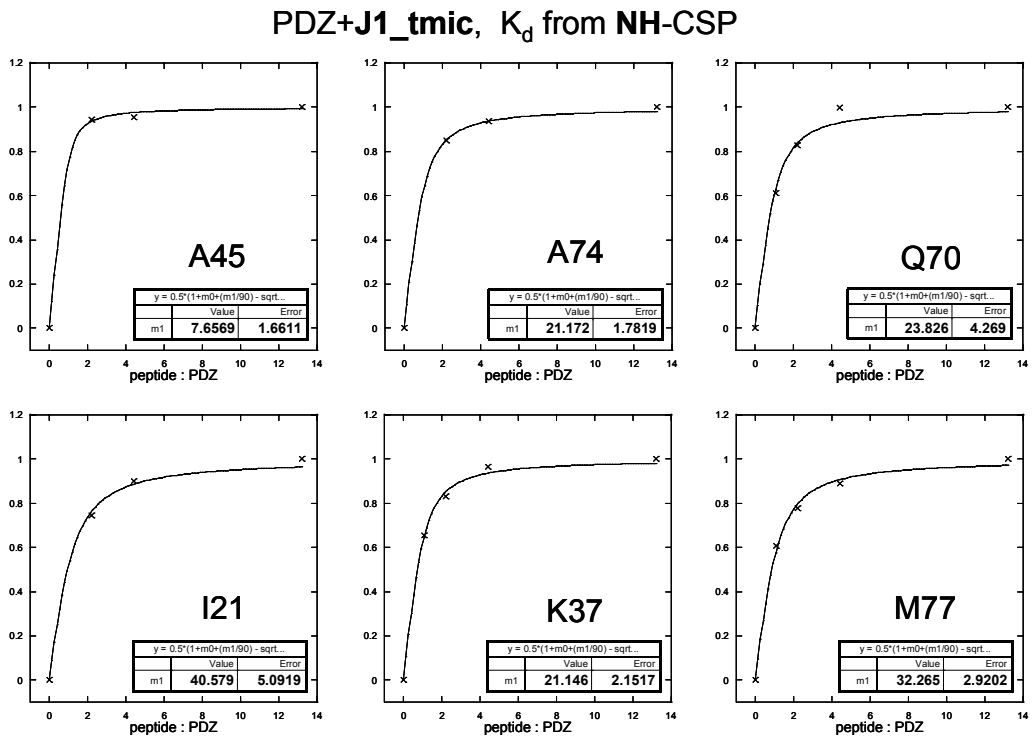


Figure S14. Determination of dissociation constants from NMR. Dissociation constants (K_d) were calculated plotting normalized CSP values against the molar ratios X (peptide:PDZ) and fitting by nonlinear regression analysis according to the formula CSP/CSPmax = 0.5*{1+X+K_d/[P] - [(1 + X + K_d/[P])2 - 4X]1/2} where [P] is a total concentrations of AF6_PDZ. In each plot, the inset reports the K_d (μM) and the fitting error.

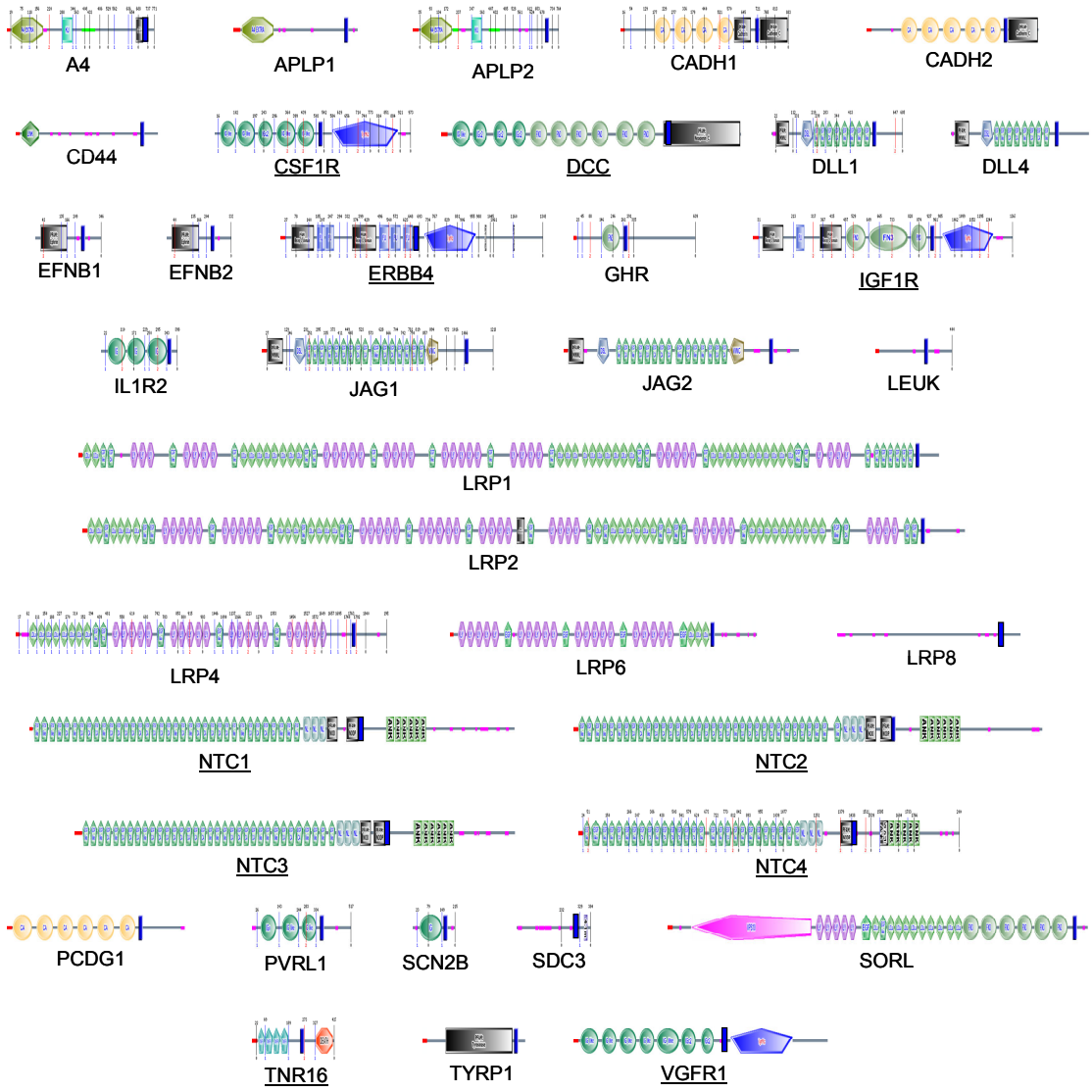


Figure S15. Intrinsic disorder. Domain architecture, as calculated by SMART, of human RIP substrates analyzed in this work. A blue vertical bar represents the transmembrane segment. Proteins containing a cytoplasmic region that is predicted to be at least partially globular are underlined, and correspond to the black filled circles in **Figure 3.2**.

	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1_tmic**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**	PDZ + **J1C24RQ**	PDZ + **J1C6pY**
res	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N	^{1}H / ^{15}N
K2	8.47/123.80	8.47/123.68	8.46/123.71	8.51/123.75	8.46/123.59	8.45/123.74	8.47/123.53	8.49/123.67
E3	8.48/123.92	8.48/123.88	8.48/123.91	8.52/123.88	8.48/123.91	8.48/123.91	8.48/123.68	8.52/123.85
E5	8.46/122.83	8.46/122.81	8.46/122.81	8.49/122.76	8.46/122.81	8.46/122.80	8.46/122.75	8.49/122.85
I6	8.25/125.29	8.25/125.29	8.25/125.31	8.29/125.22	8.25/125.31	8.24/125.27	8.28/125.24	8.29/125.32
I7	9.13/123.11	9.13/122.67	9.13/122.67	9.16/122.66	9.13/122.66	9.12/122.72	9.15/122.65	9.16/122.70
T8	8.44/118.46	8.41/117.96	8.41/117.96	8.44/117.97	8.41/117.97	8.40/117.96	8.42/117.87	8.44/117.91
V9	9.27/126.93	9.23/126.92	9.23/126.92	9.26/126.95	9.23/126.88	9.23/126.89	9.26/126.87	9.26/126.96
T10	7.97/123.52	7.91/124.32	7.91/124.32	7.95/124.42	7.90/124.37	7.90/124.28	7.92/124.24	7.93/124.39
L11	9.12/125.98	9.07/125.78	9.07/125.78	9.10/125.82	9.06/125.78	9.05/125.77	9.08/125.82	9.08/125.82
K12	8.60/124.06	8.62/123.79	8.62/123.80	8.65/123.83	8.62/123.75	8.59/123.87	8.64/123.78	8.65/123.81
K13	8.52/123.24	8.68/123.42	8.67/123.42	8.71/123.51	8.68/123.44	8.66/123.44	8.70/123.37	8.71/123.42
Q14	8.00/121.05	7.89/120.92	7.89/120.92	7.93/120.93	7.88/120.94	7.90/120.85	7.90/120.90	7.91/120.88
N15	8.75/121.33	8.56/121.87	8.56/121.88	8.60/121.92	8.55/121.88	8.58/121.84	8.58/122.13	8.58/121.91
G16	8.11/106.84	8.08/107.02	8.08/107.03	8.11/107.02	8.08/107.03	8.08/107.03	8.09/107.02	8.11/107.12
M17	9.03/118.87							
G18	8.66/106.15	9.47/108.77	0.00/0.0000	9.50/108.83	9.48/108.91	0.00/0.0000	9.50/108.85	9.47/108.77
L19	7.80/119.00	8.29/121.43	8.28/121.41	8.33/121.33	8.31/121.56	8.27/121.28	8.32/121.50	8.32/121.54
S20	8.52/117.64	8.57/116.40	8.57/116.45	8.61/116.47	8.57/116.41	8.55/116.39	8.61/116.40	8.60/116.24
I21	8.67/117.50	8.83/119.28	8.84/119.24	8.88/119.50	8.83/119.55	8.85/119.19	8.85/119.67	8.86/119.33
V22	9.00/117.09	8.93/116.16	8.92/116.09	8.94/116.05	8.92/116.09	8.89/115.97	8.92/115.70	8.92/115.83
A23	8.29/127.07	8.35/125.08	8.49/125.18	8.52/125.24	8.55/125.16	8.43/125.77	8.52/124.91	8.48/125.42
A24	9.40/125.43	9.41/124.93	9.40/124.53	9.43/124.45	9.38/124.22	9.40/124.79	9.41/124.36	9.45/124.96
K25	8.56/121.05	8.56/120.99	8.57/120.95	8.60/120.97	8.58/120.97	8.57/120.93	8.59/120.97	8.59/120.97
G26	8.46/112.30	8.47/112.34	8.47/112.34	8.51/112.33	8.46/112.19	8.45/112.18	8.48/112.25	8.49/112.32
A27	8.43/123.93	8.43/123.92	8.43/123.92	8.47/124.01	8.43/123.91	8.42/123.90	8.45/123.96	8.47/123.95
G28	8.75/110.25	8.76/110.24	8.76/110.24	8.80/110.26	8.76/110.16	8.75/110.20	8.77/110.22	8.80/110.28

	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1_tmic**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**	PDZ + **J1C24RQ**	PDZ + **J1C6pY**
Q29	7.90/118.32	7.90/118.21	7.90/118.25	7.93/118.29	7.90/118.28	7.90/118.28	7.92/118.20	7.92/118.23
D30	8.49/119.14	8.49/119.17	8.49/119.15	8.52/119.22	8.48/119.22	8.48/119.17	8.51/119.07	8.52/119.14
K31	7.69/118.46	7.68/118.44	7.68/118.44	7.72/118.43	7.68/118.44	7.67/118.36	7.69/118.36	7.70/118.43
L32	8.36/122.01	8.35/121.86	8.35/121.76	8.39/121.87	8.35/121.72	8.33/121.69	8.35/121.71	8.36/121.71
G33	8.70/110.39	8.69/110.30	8.70/110.32	8.74/110.38	8.70/110.31	8.67/110.16	8.73/110.30	8.70/110.14
I34	8.76/119.55	8.76/119.55	8.75/119.56	8.79/119.65	8.74/119.53	8.74/119.41	8.77/119.62	8.78/119.43
Y35	9.25/124.20	9.28/124.13	9.30/124.22	9.32/124.22	9.30/124.22	9.29/124.22	9.31/124.22	9.36/124.30
V36	8.96/119.82	8.83/119.49	8.83/119.50	8.85/119.56	8.81/119.47	8.82/119.46	8.84/119.49	8.86/119.54
K37	9.60/134.18	9.80/134.38	9.78/134.23	9.80/134.17	9.79/134.22	9.75/134.26	9.82/134.25	9.85/134.60
S38	7.46/109.71	7.68/109.53	7.67/109.53	7.71/109.54	7.68/109.53	7.66/109.52	7.71/109.55	7.74/109.58
V39	8.47/121.12	8.60/122.35	8.59/122.31	8.63/122.35	8.59/122.34	8.56/122.05	8.59/122.18	8.62/122.30
V40	7.88/127.21	7.84/127.59	7.84/127.55	7.88/127.58	7.84/127.66	7.82/127.51	7.86/127.57	7.86/127.60
K41	9.08/133.22	9.06/133.37	9.06/133.34	9.10/133.42	9.06/133.44	9.06/133.33	9.07/133.33	9.09/133.42
G42	10.7/117.37	10.7/117.33	10.7/117.28	10.7/117.31	10.7/117.19	10.7/117.19	10.7/117.26	10.7/117.53
G43	7.79/106.84	7.76/106.78	7.76/106.78	7.80/106.83	7.76/106.87	7.75/106.80	7.78/106.78	7.79/106.85
A44	8.66/119.00	8.83/119.67	8.83/119.66	8.87/119.74	8.84/119.69	8.82/119.59	8.92/119.46	8.86/119.74
A45	7.83/118.59	8.39/119.21	8.36/119.20	8.41/119.29	8.39/119.22	8.32/119.08	8.41/119.22	8.40/119.15
D46	8.98/123.52	8.98/124.66	8.99/124.59	9.03/124.71	8.99/124.69	9.00/124.54	9.01/124.63	9.03/124.71
V47	8.30/116.95	8.24/117.35	8.24/117.34	8.28/117.45	8.24/117.34	8.25/117.37	8.25/117.30	8.27/117.33
D48	7.65/118.87	7.68/118.75	7.68/118.75	7.71/118.83	7.68/118.75	7.69/118.75	7.69/118.81	7.70/118.77
G49	7.40/103.01	7.40/102.81	7.40/102.81	7.44/102.85	7.40/102.81	7.40/102.81	7.44/102.83	7.43/102.83
R50	7.96/118.46	7.93/118.73	7.93/118.75	7.96/118.78	7.93/118.75	7.93/118.75	7.96/118.71	7.95/118.75
L51	8.52/119.96	8.73/120.77	8.72/120.72	8.77/120.87	8.73/120.78	8.70/120.62	8.75/120.80	8.76/120.77
A52	8.94/124.34	8.98/123.91	8.97/123.91	9.01/123.97	8.97/123.91	8.96/124.03	8.99/123.91	9.00/123.90
A53	8.48/121.47	8.43/121.22	8.43/121.25	8.46/121.27	8.43/121.25	8.42/121.24	8.44/121.18	8.46/121.23
G54	9.12/111.89	9.24/112.03	9.26/112.04	9.34/112.11	9.28/112.03	9.27/112.04	9.31/112.05	9.29/112.05
D55	8.07/121.88	8.03/121.91	8.03/121.91	8.07/121.95	8.03/121.88	8.03/121.88	8.06/121.94	8.06/121.99

	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1_tmic**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**	PDZ + **J1C24RQ**	PDZ + **J1C6pY**
Q56	8.63/121.46	8.59/121.03	8.59/120.99	8.62/121.09	8.58/121.09	8.59/121.04	8.60/121.06	8.63/121.03
L57	8.94/126.39	8.93/126.86	8.92/126.85	8.96/126.88	8.92/126.87	8.93/126.75	8.92/126.74	8.98/126.91
L58	9.14/124.34	9.10/124.53	9.10/124.53	9.13/124.53	9.10/124.53	9.10/124.53	9.12/124.46	9.12/124.64
S59	7.67/111.62	7.68/111.56	7.68/111.56	7.72/111.64	7.68/111.56	7.68/111.56	7.70/111.62	7.71/111.55
V60	8.63/117.50	8.71/117.34	8.71/117.34	8.75/117.42	8.71/117.34	8.71/117.34	8.73/117.37	8.75/117.33
D61	10.0/130.76	9.96/130.62	9.96/130.62	9.99/130.65	9.96/130.62	9.96/130.62	9.99/130.62	9.99/130.66
G62	8.83/102.60	8.79/102.50	8.79/102.50	8.83/102.51	8.79/102.50	8.78/102.45	8.82/102.47	8.82/102.52
R63	8.18/122.29	8.26/122.34	8.26/122.34	8.30/122.38	8.26/122.34	8.25/122.34	8.28/122.31	8.31/122.42
S64	8.58/118.18	8.58/118.14	8.58/118.13	8.63/118.20	8.58/118.13	8.58/118.14	8.60/118.05	8.62/118.17
L65	8.04/124.75	8.05/125.01	8.05/125.01	8.11/125.01	8.04/125.00	8.03/124.99	8.08/125.04	8.06/125.13
V66	7.97/122.83	7.97/122.80	7.98/122.82	8.02/122.90	7.99/122.81	7.97/122.81	8.01/122.79	7.98/122.73
G67	8.66/116.27	8.74/116.10	8.74/116.10	8.78/116.14	8.75/116.09	8.72/116.10	8.78/116.08	8.75/116.17
L68	7.90/121.47	7.96/121.41	7.97/121.40	8.01/121.45	7.98/121.41	7.99/121.42	7.99/121.37	8.02/121.50
S69	7.77/112.71	7.86/112.96	7.85/112.98	7.90/113.04	7.85/113.12	7.81/112.81	7.86/112.99	7.85/112.89
Q70	8.97/120.37	8.90/119.70	8.95/119.90	8.99/119.96	8.96/120.00	8.86/119.70	9.01/119.97	8.93/119.52
E71	8.86/117.09	8.77/117.56	8.76/117.53	8.77/117.56	8.74/117.50	8.69/117.50	8.80/117.72	8.69/118.03
R72	7.72/120.64	7.63/120.62	7.62/120.61	7.66/120.60	7.61/120.47	7.64/120.47	7.59/120.50	7.66/120.39
A73	8.54/122.56	8.59/122.01	8.56/122.04	8.60/122.00	8.56/122.03	8.59/122.01	8.57/122.02	8.68/122.01
A74	8.58/119.28	8.27/118.95	8.28/119.05	8.30/119.13	8.26/119.06	8.40/119.18	8.29/119.07	8.43/119.25
E75	7.62/119.28	7.58/119.39	7.58/119.38	7.64/119.50	7.58/119.38	7.60/119.37	7.58/119.33	7.65/119.54
L76	7.69/119.69	7.74/118.90	7.71/118.83	7.77/118.90	7.71/118.75	7.67/118.74	7.72/118.78	7.76/118.94
M77	8.02/114.90	8.14/114.09	8.13/114.14	8.17/114.16	8.13/114.06	8.15/114.34	8.15/114.07	8.20/114.17
T78	8.02/114.63	8.08/114.13	8.08/114.07	8.11/114.06	8.09/114.06	8.10/113.91	8.12/113.96	8.14/114.30
R79	7.41/123.52	7.42/123.91	7.42/123.91	7.45/123.95	7.42/123.91	7.40/123.78	7.45/123.88	7.44/123.99
T80	7.46/110.66	7.39/110.48	7.40/110.48	7.44/110.58	7.39/110.47	7.39/110.47	7.41/110.47	7.41/110.45
S81	9.06/118.18	8.86/117.54	8.86/117.63	8.89/117.59	8.86/117.50	8.85/117.50	8.89/117.52	8.86/117.41
S82	8.19/113.67	8.16/113.30	8.16/113.30	8.21/113.44	8.17/113.28	8.16/113.40	8.19/113.28	8.18/113.30

	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1_tmic**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**	PDZ + **J1C24RQ**	PDZ + **J1C6pY**
V83	7.56/119.55	7.50/119.51	7.50/119.52	7.53/119.53	7.50/119.53	7.50/119.47	7.52/119.50	7.52/119.48
V84	9.01/127.07	8.97/126.51	8.97/126.56	9.01/126.55	8.97/126.56	8.96/126.45	8.99/126.49	8.99/126.39
T85	8.89/122.83	8.83/121.90	8.83/121.96	8.86/122.02	8.83/121.88	8.82/121.95	8.86/121.91	8.85/121.85
L86	9.61/130.22	9.61/130.21	9.61/130.25	9.65/130.31	9.61/130.31	9.61/130.31	9.62/130.22	9.63/130.27
E87	7.50/121.60	7.50/121.38	7.50/121.41	7.54/121.36	7.51/121.41	7.49/121.34	7.53/121.42	7.53/121.45
V88	9.29/125.84	9.26/125.80	9.26/125.79	9.30/125.81	9.26/125.79	9.26/125.79	9.29/125.84	9.29/125.86
A89	9.19/128.30	9.20/128.27	9.20/128.28	9.23/128.30	9.20/128.28	9.19/128.28	9.23/128.24	9.23/128.31
K90	8.93/125.57	8.93/125.76	8.93/125.72	8.96/125.76	8.93/125.78	8.92/125.75	8.95/125.71	8.97/125.76
Q91	8.78/122.56	8.79/122.58	8.79/122.63	8.83/122.64	8.79/122.66	8.78/122.65	8.81/122.58	8.82/122.59
G92	8.56/111.10	8.55/111.07	8.55/111.09	8.59/111.09	8.55/111.09	8.55/111.09	8.57/110.96	8.59/111.10
A93	7.97/129.40	7.98/129.37	7.98/129.37	8.02/129.43	7.98/129.37	7.97/129.38	8.00/129.37	8.01/129.44

Table S1. 1H, 15N-Chemical shifts. The table shows chemical shifts of the backbone amides of AF6_PDZ alone, and titrated with different Jagged-1 peptides.

residue (CH3)	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**
I6(1)	0.70/17.53	0.70/17.49	0.70/17.47	0.70/17.47	0.69/17.48
I6(2)	0.73/12.75	0.72/12.71	0.72/12.72	0.72/12.71	0.71/12.71
I7(1)	0.71/14.40	0.70/14.43	0.70/14.43	0.69/14.45	0.69/14.42
I7(2)	0.83/17.53	0.83/17.47	0.82/17.46	0.82/17.46	0.82/17.46
T8	1.00/21.71	1.01/21.71	1.01/21.71	1.01/21.71	1.00/21.72
V9(1)	0.83/21.04	0.84/21.20	0.84/21.18	0.83/21.21	0.83/21.17
V9(2)	0.73/21.59	0.72/21.59	0.72/21.58	0.72/21.59	0.72/21.61
T10	1.00/21.61	0.99/21.64	0.99/21.62	0.99/21.64	0.99/21.61
L11(1)	0.83/24.62	0.81/24.41	0.81/24.39	0.81/24.39	0.81/24.39
L11(2)	0.69/25.49	0.67/25.55	0.67/25.54	0.67/25.55	0.67/25.51
L19(1)	0.67/25.77	0.64/25.88	0.64/25.87	0.64/25.87	0.63/25.87
L19(2)	0.70/26.59	0.80/25.50	0.80/25.52	0.80/25.46	0.78/25.65
I21(1)	0.76/14.57	0.72/14.55	0.72/14.56	0.72/14.55	0.73/14.59
I21(2)	0.63/18.63	0.67/18.83	0.67/18.82	0.68/18.82	0.66/18.78
V22(1)	0.97/20.32	0.90/19.38	0.88/19.31	0.88/19.23	0.90/19.61
V22(2)	0.86/20.37	0.94/21.77	0.93/21.88	0.93/21.99	0.90/21.59
A23	1.38/20.72	1.34/21.08	1.35/21.10	1.35/21.12	1.34/20.89
A24	1.21/22.59	1.18/22.54	1.17/22.55	1.17/22.55	1.18/22.59
A27	1.33/18.22	1.33/18.21	1.33/18.23	1.33/18.24	1.32/18.22
L32(1)	0.74/25.02	0.72/24.99	0.72/25.03	0.72/25.03	0.70/25.01
L32(2)	0.71/23.65	0.69/23.56	0.70/23.64	0.70/23.66	0.68/23.50
I34(1)	0.74/19.13	0.76/18.63	0.76/18.64	0.76/18.61	0.75/18.72
I34(2)	0.59/9.09	0.56/9.11	0.57/9.10	0.57/9.09	0.55/9.05
V36(1)	0.89/22.50	0.90/22.51	0.89/22.52	0.89/22.52	0.89/22.50
V36(2)	0.64/22.36	0.64/22.74	0.63/22.72	0.63/22.75	0.63/22.68
V39(1)	0.78/21.74	0.79/21.68	0.79/21.68	0.79/21.68	0.78/21.69
V39(2)	0.65/20.49	0.64/20.45	0.64/20.45	0.64/20.45	0.64/20.46
V40(1)	0.82/21.01	0.78/20.35	0.78/20.38	0.78/20.36	0.78/20.43
V40(2)	0.89/20.50	0.88/20.37	0.88/20.39	0.88/20.38	0.87/20.41
A44	1.48/18.78	1.51/18.46	1.51/18.46	1.51/18.45	1.50/18.47
A45	1.38/18.74	1.31/18.84	1.31/18.85	1.31/18.85	1.31/18.83
V47(1)	0.92/21.06	0.92/21.05	0.92/21.06	0.92/21.05	0.92/21.06
V47(2)	0.95/22.30	0.96/22.38	0.96/22.38	0.96/22.38	0.95/22.36

Table S2. ^{1}H, ^{13}C-Chemical shifts.

residue (CH3)	PDZ **free**	PDZ + **J1C6**	PDZ + **J1C24**	PDZ + **J1C24pS**	PDZ + **J1C24pSpY**
L51(1)	0.71/26.15	0.73/26.09	0.73/26.07	0.73/26.06	0.72/26.08
L51(2)	0.82/25.77	0.85/25.96	0.85/25.96	0.85/25.96	0.85/25.92
A52	1.28/23.00	1.29/23.23	1.29/23.24	1.29/23.24	1.28/23.21
A53	1.24/17.86	1.26/18.09	1.26/18.09	1.26/18.12	1.25/18.09
L57(1)	0.64/25.90	0.64/26.17	0.64/26.17	0.64/26.17	0.63/26.09
L57(2)	0.63/24.04	0.63/23.91	0.63/23.93	0.62/23.92	0.62/23.91
L58(1)	0.75/26.15	0.75/26.18	0.74/26.18	0.74/26.18	0.74/26.14
L58(2)	0.60/21.80	0.59/21.76	0.59/21.75	0.59/21.75	0.59/21.75
V60(1)	0.83/19.64	0.85/19.35	0.85/19.34	0.85/19.34	0.85/19.37
V60(2)	0.79/22.08	0.81/22.32	0.81/22.31	0.81/22.32	0.80/22.30
L65(1)	0.55/26.44	0.56/26.55	0.56/26.55	0.56/26.57	0.55/26.55
L65(2)	0.77/24.30	0.77/24.21	0.77/24.22	0.77/24.22	0.77/24.25
V66(1)	0.92/21.10	0.91/21.20	0.91/21.20	0.91/21.22	0.91/21.22
V66(2)	1.04/21.78	1.03/21.77	1.03/21.74	1.03/21.74	1.02/21.77
L68(1)	0.79/25.78	0.78/26.03	0.77/26.03	0.77/26.04	0.77/25.99
L68(2)	0.76/23.83	0.75/23.71	0.74/23.72	0.74/23.74	0.74/23.76
A73	1.32/19.18	1.28/19.07	1.29/19.07	1.28/19.06	1.29/19.12
A74	1.48/17.76	1.27/17.71	1.22/17.69	1.19/17.68	1.32/17.67
L76(1)	0.77/23.37	0.74/23.69	0.74/23.67	0.74/23.71	0.74/23.62
L76(2)	0.73/25.81	0.73/25.68	0.73/25.68	0.73/25.68	0.73/25.69
T78	1.32/21.69	1.38/21.46	1.38/21.49	1.38/21.49	1.38/21.53
T79	1.29/22.51	1.26/22.44	1.26/22.45	1.26/22.44	1.26/22.42
V83(1)	0.80/20.70	0.78/20.54	0.78/20.55	0.78/20.55	0.78/20.58
V83(2)	0.72/21.55	0.70/21.55	0.70/21.55	0.70/21.55	0.69/21.56
V84(1)	0.95/21.50	0.96/21.55	0.96/21.55	0.96/21.55	0.95/21.55
V84(2)	0.98/21.99	0.97/21.73	0.97/21.73	0.97/21.73	0.97/21.74
L86(1)	0.70/24.40	0.70/24.18	0.70/24.18	0.70/24.18	0.70/24.25
L86(2)	0.80/26.31	0.84/26.48	0.84/26.48	0.84/26.48	0.82/26.44
V88(1)	0.76/21.19	0.75/21.17	0.75/21.17	0.75/21.17	0.74/21.16
V88(2)	0.72/21.31	0.71/21.27	0.71/21.27	0.71/21.27	0.71/21.27
A89	1.30/19.67	1.30/19.68	1.30/19.68	1.29/19.67	1.29/19.66
A93	1.30/20.24	1.29/20.26	1.29/20.26	1.29/20.26	1.29/20.22

Table S2. ^{1}H, ^{13}C-Chemical shifts. The table shows chemical shifts of assigned methyl groups of AF6_PDZ alone, and titrated with different J1C24 peptides. For residues with two methyl groups (valines, leucines and isoleucine) chemical shifts of both groups are shown.