universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„NMR spectroscopy of intrinsically disordered protein complexes“

verfasst von / submitted by

Andrea Gabriele Flamm, BSc MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2016 / Vienna 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Robert Konrat

Acknowledgments

The last three years have been an exciting and challenging time and I am deeply grateful for this experience.

First of all, I want to thank my supervisors Robert Konrat and Nicolas Coudevylle for their guidance and help during my PhD studies.

Working in the VBC 5 in the Department of Computational and Structural Biology has been fantastic from both the scientific and personal point of view. Thanks to all former and present people working here and of course especially everyone from the NMR group. Karin Ledolter helped me to conduct the work in the lab and Georg Kontaxis helped me with the spectrometers.

I would like to thank Miquel Pons and Anabel-Lise Le Roux as well as Wiktor Koźmiński and his group and all other people I worked with during the last years. Also, I spent two months in the group of Geoffrey Bodenhausen and Fabien Ferrage working with Pavel Kadeřávek where I learnt a lot about NMR.

Finally, I want to thank my friends and family for all their support over the last years.

This cumulative thesis is based on the following three manuscripts:

Chapter 1

Flamm, A. G.; Le Roux, A.-L.; Mateos, B.; Díaz-Lobo, M.; Storch, B.; Breuker, K.; Konrat, R.; Pons, M.; Coudevylle, N., N-Lauroylation during the Expression of Recombinant N-Myristoylated Proteins: Implications and Solutions. *ChemBioChem,* **2016**, 17, 82-89.

Chapter 2

Flamm, A.G.; Żerko, S.; Zawadzka-Kazimierczuk, A.; Koźmiński, W.; Konrat, R.; Coudevylle, N., ^{1}H, ^{15}N, ^{13}C resonance assignment of human GAP-43. *Biomol. NMR Assign***. 2016,** 10, 171-174.

Chapter 3

Flamm, A.G.; Konrat, R.; Coudevylle, N.; The membrane binding properties of two neuronal intrinsically disordered proteins: GAP-43 and BASP1. *Structure,* **2016,** manuscript number: STRUCTURE-D-16-00375, submitted (24. Oct. 2016).

Other Publication

Kurzbach, D.*; Vanas, A. *; **Flamm, A. G.***; Tarnoczi, N.; Kontaxis, G.; Maltar-Strmecki, N.; Widder, K.; Hinderberger, D.; Konrat, R., Detection of correlated conformational fluctuations in intrinsically disordered proteins through paramagnetic relaxation interference. *Phys. Chem. Chem. Phys.* **2016**, 18, 5753-5758. (* shared first authorship)

Kurzbach, D.; **Flamm, A. G.**; Sára, T., Network representation of protein interactions—Experimental results. *Prot. Sci.* **2016**, 25, 1628-1636.

Vuichoud, B.; Bornet, A.; de Nanteuil, F.; Milani, J.; Canet, E.; Ji, X.; Miéville, P.; Weber, E.; Kurzbach, D.; **Flamm A. G.**; Konrat, R.; Gossert, A. D.; Jannin, S.; Bodenhausen, G., Filterable Agents for Hyperpolarization of Water, Metabolites, and Proteins. *Chem. Eur. J.* **2016**, 22, 14696-14700.

Kurzbach, D.; Canet, E.; **Flamm, A.G.**; Jhajharia, A.; Weber, E.; Konrat, R.; Bodenhausen, G., Casting Light on Unfolding Processes of Intrinsically Disordered Proteins by Selective Exchange with Hyperpolarized Water. *Angew. Chem. Int. Ed.,* **2016**, Accepted (Oct 2016).

Abstract

Since the rather recent discovery of their crucial role in protein interaction networks intrinsically disordered proteins (IDPs) are subject to intense investigations. IDPs are devoid of a stable tertiary structure or pronounced secondary structure elements. This feature enables them to sample an enormous and heterogeneous conformational space. Continuous interconversion between states within this space is an intrinsic feature. Many intrinsically disordered proteins have the potential to interact simultaneously or sequentially with a range of different partner molecules. Still, the description of the complex formation with their binding partners remains elusive in many cases. This work focuses on two neuronal human intrinsically disordered proteins, the Growth Associated Protein 43 (GAP-43) and the Brain Acid Soluble Protein 1 (BASP1). Special attention is paid to their interactions with different binding partners as well as to the structural and functional implications that follow these interactions.

GAP-43 and BASP1 do not share significant sequence identity yet they fulfill similar biological functionalities. Both are associated with neuronal membranes and are found on the outside of synaptic vesicles. They are known to bind Calmodulin (by very distinct modes), are substrates of Protein Kinase C and co- localize with phosphatidylinositol-4,5-bisphosphate. Their membrane interaction properties are influenced by acylation of the proteins' N-termini. Therefore, these two proteins are models of choice to understand the complex formation of structurally different but functionally related intrinsically disordered proteins.

In order to better understand the complex formation of GAP-43 and BASP1, nuclear magnetic resonance (NMR) spectroscopy in combination with other biophysical methods are employed for this work. Some light is thus shed on the binding of intrinsically disordered proteins to membrane systems and on the accompanying structural changes.

Zusammenfassung

Aufgrund der entscheidenden Rolle von intrinsisch ungeordneten Proteinen (IDP) in Proteininteraktionsnetzwerken stellen diese Proteine die Grundlage eines intensiven Forschungsfeldes dar. IDP besitzen weder stabile Tertiärstruktur noch ausgeprägte Sekundärstrukturelemente. Dadurch ist es für sie möglich, einen großen, heterogenen Konformationsraum einzunehmen und in diesem eine kontinuierliche Umwandlung zwischen den darin enthaltenen Konformationen durchzuführen. Viele intrinsisch ungeordnete Proteine können sowohl simultan als auch sequentiell mit einer Reihe von verschiedenen Partnermolekülen interagieren. Dennoch bleibt die Beschreibung der Komplexbildung mit ihren Bindungspartnern oftmals wage. Diese Arbeit fokussiert sich auf zwei neuronale humane Proteine, Growth Associated Protein 43 (GAP-43) und Brain Acid Soluble Protein 1 (BASP1), und deren Interaktion mit verschiedenen Molekülen sowie die daraus resultierenden strukturellen und funktionellen Auswirkungen. Diese beiden Proteine teilen keine signifikante Sequenzhomologie, trotzdem erfüllen sie ähnliche biologische Funktionen. Man findet sie assoziiert mit neuronalen Membranen und auf der Außenseite von synaptischen Vesikeln. Beide binden Calmodulin (jedoch auf unterschiedliche Weise), sind Substrat der Protein Kinase C und co-lokalisieren mit Phosphatidylinositol-4,5-Bisphosphat. Außerdem werden ihre Membraninteraktionseigenschaften durch Acylierung am N-Terminus beeinflusst.

Deswegen sind diese Proteine ideale Modelle an Hand derer es zu verstehen gilt, wie die Komplexbildung von substantiell unterschiedlichen aber funktionell verwandten intrinsisch ungeordneten Proteinen zustande kommt und wie sich deren Konformationsensemble aufgrund dieser Komplexbildung ändert.

Um die Komplexbildung von GAP-43 und BASP1 besser zu verstehen, wurde in dieser Arbeit eine Kombination von Kernmagnetresonanzspektroskopie mit anderen biophysikalischen Techniken verwendet. Diese führte zu wichtigen Erkenntnissen bezüglich der Bindung von IDP an Membransystemen und die damit verbundenen strukturellen Veränderungen.

Table of Contents

1. **Introduction to intrinsically disordered proteins**

Over billions of years living matter on earth evolved creating the modern world we find ourselves in nowadays. “Zooming in” we find that life is based on different entities on different length scales. On the micrometer length scale cells constitute reproducible entities that build the basis for our life. On a nanometer length scale we find that cells are amongst other units built up on proteins. These themselves are biopolymers that are made up of amino acids as basic building blocks. Consequentially, we arrive at atoms that give rise to the chemical structure of the different amino acids.

The construction of living matter can thus be viewed as a bottom-to-top process. Over time, evolution resulted in more and more complex combinations of different biomolecules resulting finally in multicellular organisms and increasing specialization of different cell types within one organism. The more complex the cellular environment gets, the higher is the demand for cross-talk among the different cellular units, i.e., the cells require a vast network of proteins and molecules to function properly. For many years scientists used to associate this higher complexity with more complicated protein architectures: highly sophisticated three-dimensional structures that fulfill the requirements necessary for the regulation of complex cellular networks. However, evolution partly solved the problem of the growing demand of regulation in a cell via a somewhat counterintuitive approach. Comparing the proteomes of bacteria and archaea to those of eukaryotes, we find that the latter has a significantly higher percentage of so-called intrinsically disordered proteins (IDP) and intrinsically disordered regions (IDR).[1] Preferentially found in signaling pathways, transcription and regulating functions, this class of biomolecules offers unique advantages over globular proteins with a well-defined structure.[2]

The benefits of IDPs and IDRs can be attributed to a lack of stable secondary structure elements – and therefore a rigid three dimensional structure – under native conditions. They rather exist in a range of different conformations that dynamically commute into each other, yet retaining biological functionality.[3] As they are more common in complex organisms, apparently these proteins offer some advantages that are needed for setting up complex cellular networks. IDPs frequently function as “molecular hubs”, meaning that they are embedded in an interaction network in which they can both interact with or be targeted by multiple molecules simultaneously or sequentially.[4] It is widely recognized that this versatility is based on the lack of 3D structure augmenting their exposed molecular surface and accessibility. This extended conformation thus provides a higher interaction surface per residue and less steric hindrance during the interaction.[2]

Since IDPs are devoid of a stable hydrophobic core, the whole protein is more accessible for binding partner(s) but also for post-translational modifications (PTMs). Some PTMs seem to occur preferentially in IDRs, e.g., phosphorylation.[5] PTMs enable quick adaptations to the continuously changing environmental conditions, which is crucial for efficient signaling in the cell. IDPs are also tightly regulated [6] by rapid degradation and cellular turnover rates because of their promiscuous binding abilities. This feature is simultaneously a major drawback for humankind because once deregulated, they are likely to be associated with many pathologies such as neurodegenerative diseases and cancer.[7] Prominent examples include BRCA1 in breast cancer, alpha-synuclein and Tau in Alzheimer's disease or p53 in even a plethora of cancer types.

The field of IDP research is rather young and became a popular topic in the beginning of the 2000's. At this time the first disorder prediction techniques for IDPs were developed.[8, 9] The prediction of disordered regions is based on the amino acid bias that is found between globular, well- folded proteins and IDR.[10] On the primary sequence level, IDPs differ from folded proteins in their amino acid composition and complexity, net charge and hydropathy. (Special attention has to be paid to disordered integral membrane proteins whose amino acid composition differs from that of soluble IDRs.[11]) Often rich in polar, charged amino acids, proline and low in hydrophobic and aromatic residues, IDPs accumulate disorder promoting residues in their sequence and are depleted of amino acids found in rigid structures. This difference enables the prediction of disorder not only on a local scale but even among whole genomes. This led to the important conclusion that the occurrence of protein disorder is significantly higher in eukaryotes than in both bacteria and archaea.[1] Interestingly the highest percentage of IDR among genomes was found for small viruses.[12] 37-50% of all amino acids of the human proteome belong to disordered regions according to various disorder predictors.[13]

IDPs possess an inherent flexibility that goes along with their structural heterogeneity which has many consequences for their function. The corresponding motions in their native state, which entail very complicated transitions between various structural and dynamic sub-states, are often described via an associated energy landscape. It describes the energy barriers between different sub-states in the ensemble of possible structures of a protein (see Figure 1).[14] There are two extremes of this picture: first a completely folded protein that preferentially exists in a state of a single global minimum. Second, an IDP that does not feature a global energetic minimum, but rather a rugged energy surface with energy barriers that separate different minima of similar depth. While in the former case the global energy minimum determines a more or less constant protein conformation, the latter case allows for a continuous transition between a range of conformers as hopping from minimum to

minimum of the energy surface entails only small activation barriers. Therefore, from a structural point of view IDPs can be described as an ensemble of conformers. The associated energy landscape of an IDP can be altered by, e.g., binding of another molecule, PTM or changes in external parameters (see Figure 1) thereby manipulating the proteins structural properties.

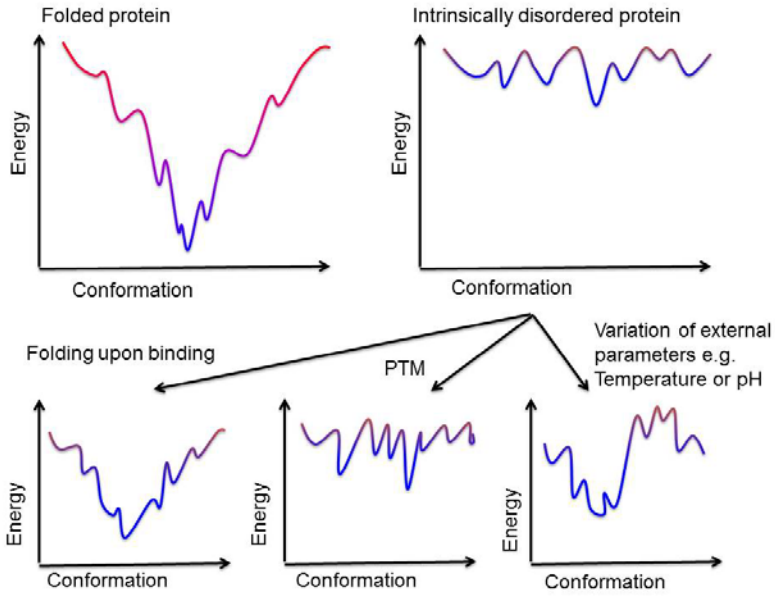


Figure 1: Depiction of the energy landscape of a folded protein and an IDP. Various factors can influence the associated energy surface of an IDP: formation of a complex that is accompanied by folding, PTM or a change in the local environment.

1.1. Biophysical characterization of intrinsically disordered proteins

For a long time, the classical structure- function principle dominated structural biology, thus, many of the established methods are tailored for structure determination of proteins that exist in a single conformation or at least display a considerable amount of secondary structure. In contrast, one unique structure cannot properly describe an IDP or IDR since dynamic interconversion within the conformational ensemble is an integral part of their native state. Therefore, a classic structural characterization is impossible. Yet, what can be achieved by experimental approaches is a better understanding of this dynamic behavior by describing parameters like the overall average shape, residual secondary structure propensities, long-range transient contacts and differences in flexibility and mobility along the primary sequence.

Most structures of proteins that are known today have been determined by X-ray crystallography.[15] This led to the impression of a requirement of a static structure for a proper function. The so-called lock-key principle [16] of protein function is a prominent variant of this deduction. However, there are numerous examples in the protein database (PDB) [17] of structures including IDR. Thus, over the years doubts prevailed concerning the necessity of a rigid structure for protein function. In crystal structures, IDR are commonly lacking electronic density (or have a high B-factor [18]) due to their inherent flexibility.

Solely because IDPs cannot be portrayed by one single three dimensional structure does not mean that their structural behavior is completely irregular. Many IDPs display propensities of secondary structure content of, e.g., alpha-helices or beta sheets that they transiently populate and that are present to a certain percentage in the ensemble of structures.[19] They can be restricted to certain areas along the protein and might also play an essential role in the interaction with binding partners that can stabilize preformed structural elements.[20]

It is important to note that an IDP does not display the same structural ensemble as a denatured protein. Studies of the hydrodynamic radius clearly show that there is a difference between them.[21] IDPs are not just extended but contain cooperatively folded parts that can be probed by classical unfolding studies. One example of such a study is the IDP Osteopontin that was shown to sample compact conformations directly related to its binding to heparin.[22] The cooperatively folded compact core of Osteopontin can be probed by classical titration experiments with urea and sodium chloride and illustrates the non-random behavior of the transiently sampled tertiary contacts of IDPs.[23]

1.1.1. Overview over the most important experimental techniques

To get an overview of the overall disorder and the occurrence (propensity) of secondary structures, Circular Dichroism (CD) [24] can be used. CD is based on the unequal absorption

of left- and right- handed circular polarized light by chiral molecules. It makes use of the fact that different elements of secondary structure give rise to characteristic absorption of this kind of light. Although CD can report on overall disorder properties, residual secondary structures cannot be assigned to a particular part of the protein and are potentially hard to detect when the structure is transient. Despite being a low-resolution technique, it can yield valuable information concerning characterization of IDPs under different conditions, which, e.g., can induce changes in secondary structure content.[25] Similar information can also be obtained with Optical Rotatory Dispersion (ORD), Fourier-Transform Infrared Spectroscopy (FTIR) and Raman Spectroscopy.

Long range contacts between different parts of the protein and low- resolution structural constraints can be probed by Small angle X-ray scattering (SAXS) [26, 27] that can determine the Radius of gyration and in some cases the relative topology of domains. SAXS can also provide the oligomeric state and overall shape of a bio- macromolecule as the hydrodynamic radius of a folded protein differs significantly from a disordered protein of similar size.[21]

Electron paramagnetic resonance (EPR) can yield detailed information about pairwise distance distribution functions (PDDF) of doubly site directed spin-labeled proteins by double electron electron resonance (DEER) spectroscopy creating a detailed picture of intra- or intermolecular interactions.[23] Likewise, fluorescence based measurements can be used for identifying long-range contacts.[28] If present, intrinsic fluorescence (e.g., through tryptophans) can be studied although IDPs are often depleted in bulky amino acids. Yet, labeling with donor and acceptor fluorophores allows for, e.g., fluorescence resonance transfer (single molecule FRET) experiments.[29] Mass spectrometry is a valuable technique for structural and/or binding studies of IDPs, e.g., based on hydrogen/deuterium exchange [30] or on crosslinking techniques, which highlight potential binding partners and the associated epitopes.[31]

A further, simple way to characterize unfolded regions is limited proteolysis. Here, one observes the extent of proteolytic digestion by proteases like trypsin, which correlates with amino acid exposure and flexibility.[32]

Furthermore, the overall shape and weight of a protein can be determined by gel filtration together with static and dynamic light scattering. A noteworthy, special feature of IDPs is their abnormal migration in sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis. Due to the large number of polar amino acid in IDPs, they bind less SDS molecules per protein than ordered proteins leading to an apparent higher molecular weight.[33]

1.1.2. Contribution of nuclear magnetic resonance to the characterization of intrinsically disordered proteins

A key player in the study of IDPs and IDRs is nuclear magnetic resonance (NMR) spectroscopy. This technique has evolved into a versatile toolbox that allows the study of various aspects of protein structure and function. The continuous technical improvement in liquid state NMR renders it most promising in obtaining various information about IDPs at atomic resolution.

Chemical Shift Distribution

The NMR resonance frequency of an amino acid in an IDP – as quantified by the chemical shift (CS) – carries an immense information content that can be exploited for the description of IDP properties. Typically, one monitors amide protons and nitrogens in the backbone of a protein. The local electronic environment of their nuclei influences the CS. Therefore, variations in tertiary structure contacts and secondary structure population change the electronic environment of the nucleus, and therefore its CS. A prominent example of the utility of this phenomenon is the following: Compared to globular proteins, the IDPs' proton chemical shift distribution is very narrow due to the averaging of the local contributions and the rather similar environment of the different nuclei due to the lack of stable secondary structure and tertiary structure contributions (see Figure 2). Thus, from a simple analysis of the width of the CS distribution one can often evaluate whether a protein is tightly folded or not.

NMR Resonance Assignment

To perform efficient NMR spectroscopy on proteins one needs to know which resonance line corresponds to which nucleus. A signal assignment, a table with chemical shifts associated to a certain nucleus is, hence, necessary. For folded proteins one typically performs triple resonance (3D) experiments, in which nuclear magnetization is transferred from amide protons to their bound nitrogens and subsequently to surrounding carbon atoms. By matching the resonance frequencies of carbons that connect neighboring amides one can eventually work its way through the primary sequence of a protein and associate each nucleus of each amino acid with its CS values.[34]

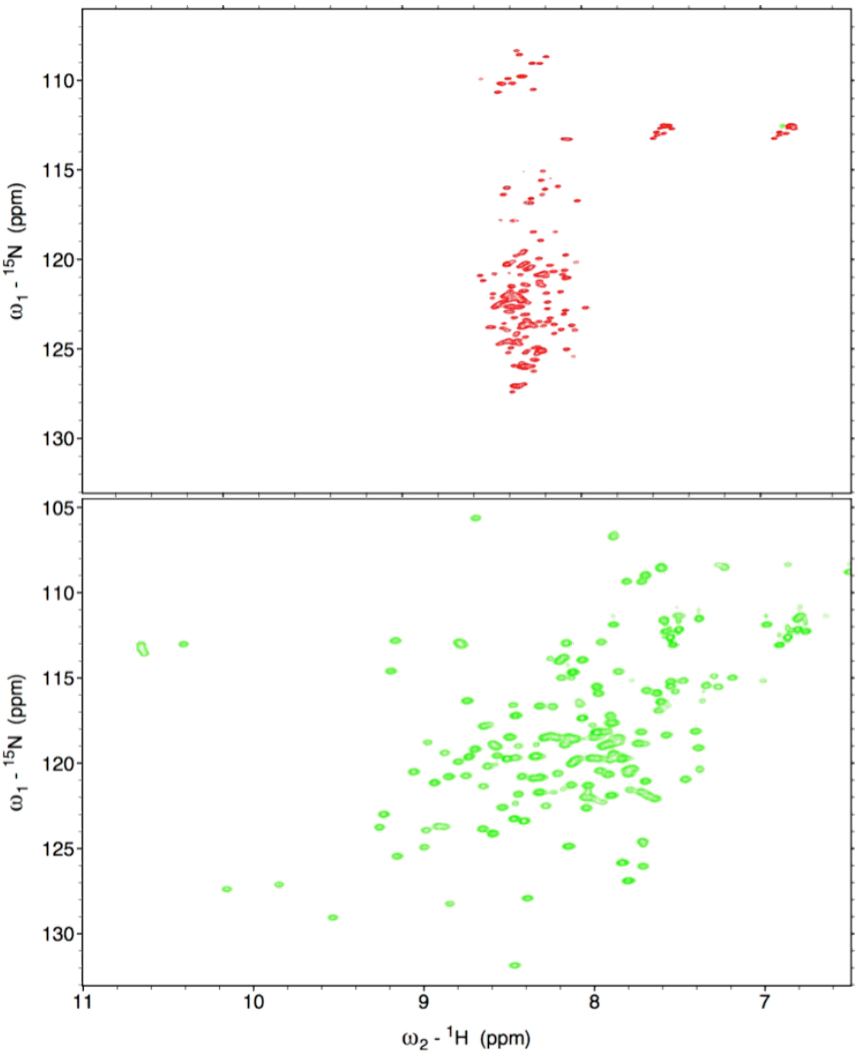


Figure 2: Comparing the proton chemical shift distribution in ^{1}H-^{15}N HSQC (Heteronuclear Single Quantum Coherence) of an IDP (top, BASP1) and a folded protein (bottom, CaM). Protein expression, data acquisition and analysis were done as described in [43]. Note that CaM (~17 kDa) has ~65% of the total number of amino acids, i.e. cross peaks in the HSQC, than BASP1 (~25 kDa) and yet shows a tremendously broader CS distribution.

Unfortunately, in the case of IDPs, the narrow CS distributions (especially $^{1}H_{N}$, see Figure 2 top) lead to relatively crowded spectra. These can potentially impede the assignment procedure as the peaks cannot be separated in conventional 3D spectra. Recent developments in higher dimensionality (number of dimensions >3) assignment methods in combination with non-uniform sampling (NUS) helped to overcome this limitation and allowed the successful NMR assignment of a range of IDPs.[35-38] While the addition of dimensions can help to resolve crowded spectral areas,[39] NUS greatly reduces the acquisition time.[40] Helpful sequence optimization applicable for IDPs further reduce the time needed for data acquisition as has been shown by the employment of BEST (band selective excitation short transient) [41] strategies as for example in the SOFAST- HMQC (band selective optimized flip angle short transient heteronuclear multiple quantum coherence) experiment that additionally uses Ernst angle excitation.[42]

Structural Information from NMR

Once the assignment is completed, residual secondary structure can be calculated from ^{13}C alpha, beta and carbonyl chemical shifts. One makes use of the fact that the ^{13}C CS is dependent on the local backbone geometry of the protein and can report on the population of secondary structure elements such as alpha helix or beta sheets. The notion of propensity of a secondary structure element takes into account the fact that many (ID) proteins sample a secondary structure only transiently. The propensity, hence, indicates the probability to find this element in a particular stretch of the protein. To do a ^{13}C secondary CS analysis of IDPs one needs a reference CS table for the different amino acids. There are many reference chemical shift sets reported in the literature that can be used to compare the chemical shifts of amino acids in a particular protein (found via the abovementioned resonance assignment) to shifts that were obtain from “random coil-type” proteins or from short peptides.[44] Note that in this respect pH and temperature [45] as well as neighboring amino acids have to be taken into account as they influence the random coil CS.[46] Thus, by evaluating the ^{13}C CS for every amino acid in the backbone of a protein the distribution of secondary structure elements in the protein can be determined.

Other NMR parameters that can offer information about IDP residual structure are coupling constants. Couplings arise from the interaction of neighboring nuclei via chemical bonds (scalar coupling, J-coupling) or through space (dipolar coupling).

Scalar coupling leads to splitting of NMR lines. It is quantified in Hertz corresponding to the frequency difference between the split NMR lines. These couplings form not only the basis for the coherence transfer in multidimensional NMR experiments, but can yield additional information. For example, the three- bond coupling is sensitive towards the dihedral angles in a peptide bond. In the case of IDPs the fast (on the NMR time scale) structural fluctuations result in an averaged observable J-value, which reports on partial populations of secondary structure elements in certain regions of a protein.[47] An important concept for the interpretation of this kind of data is the Karplus relation that connects the J-coupling with dihedral angels between coupled atoms.

Dipolar couplings are inversely proportional (r^{-6}) to the distance between two interacting spins.[48] They are also dependent on the angle between the connecting vector of the two coupling partners and the static magnetic field and are consequently averaged to zero in solution (fast rotation of the connecting vector). Under these conditions the couplings do not give rise to a splitting. However, the angular dependence can be exploited by anisotropic alignment of the sample resulting in so-called residual dipolar couplings (RDCs). RDCs rely on the weak preferential orientation in an anisotropic medium or natural magnetic anisotropy

of a molecule and enables the otherwise isotropically averaged spin-spin dipolar interaction in solution to be measured as line splitting.[48] RDC report on structural parameters (relative vector information) and can be used for the characterization of biomolecules.[49]

Dynamic Information from NMR and Nuclear Spin Relaxation

Beside the structural characterization, NMR can also provide information about dynamic parameters of IDPs. Since an IDP typically constitutes an inhomogeneous entity in terms of steric constraints along its protein backbone, it displays varying motions and dynamics in dependence of the position in the primary sequence. Additionally, these motions take place on different timescales. Many factors influence the mobility of a given segment of an IDP, such as long-range contacts, residual structure or a more hydrophobic amino acid composition that can lead to more compact regions. Thus, these factors may lead to a deviation in measured dynamics from values found for more disordered/extended regions in the protein. Frequently, such deviations indicate functional importance of the identified patch in the protein.[50] Important reporters of a protein's dynamics include the ^{15}N longitudinal and transverse relaxation times (T_1 and T_2 respectively) and heteronuclear ^{1}H-^{15}N heteronuclear NOE (η, HetNOE).

Relaxation is concerned with the return of the system from an exited to its ground state after excitation by radio frequency pulses. It is caused by fluctuating magnetic fields that result from the motion of the molecule in solution.[48] This process is often described by a mono-exponential decay of the magnetization with a characteristic relaxation time.[48] The return to the thermal equilibrium by loss of energy is described in the longitudinal relaxation time T_1 (or rate $R_1 = 1/T_1$). The transverse relaxation time T_2 (or rate $R_2 = 1/T_2$) is mainly due to the dephasing of coherences.[48] T_2 defines the linewidth of the detected NMR signals. In both cases, longitudinal and transverse relaxation, the relaxation rates are dependent on the rotational diffusion of the molecule.[48] The rotational movement is quantified by the rotational correlation, τ_c, time that characterizes the time needed for a spherical molecule to rotate through one radian.[48] In the case of IDPs T_2 is relatively long as the movements of IDPs are quite fast compared to globular proteins.[51] The corresponding NMR lines become, thus, narrow. This feature is also exploited for the design of higher- dimensional NMR experiments with the detection of an increasing number of indirect dimensions to compensate for the signal overlap in two- and three dimensional spectra.

An analysis of R_1 and R_2 (and also of heteronuclear NOEs) as a function of the primary sequence of an IDP can help to identify regions of slow or fast dynamics, i.e., of compacted or extended structural elements, respectively. Changes in R_1, R_2 and HetNOE are also informative when looking at the formation of complexes involving an IDP. Depending on the

presence or absence of a ligand molecule, the dynamics of an IDP can change. Thus, the differential relaxation parameters will reveal regions of adapted dynamics such as binding sites or compensatory/allosteric sites. This has been shown for the IDP Osteopontin and its binding partner Heparin.[22, 52] The complex formation leads to a rigidification in the binding site, but to more flexibility in a compensatory site that was initially in contact with the binding site as an intramolecular contact.[22]

Structural Constraints from NMR

The most macroscopic, yet, still molecular measure of a protein structure is probably the protein's hydrodynamic radius or its radius of gyration. NMR can be employed to determine these values by the measurement of the diffusion constant and the hydrodynamic radius by Pulsed Field Gradient-NMR.[21] Empirically, distinct differences between R_h and R_g of IDPs and folded proteins were found.[21] A comparison to fully folded or denatured proteins with globular proteins indicates a relation between R_h and R_g, the compactness and the degree of disorder in a protein. IDPs typically display values between those of folded and denatured proteins.[21]

On the intramolecular length scale, the NMR parameter called the Nuclear Overhauser effect (NOE) is employed to determine long-range contacts.[48] These distance constraints are often used in NMR structure calculations. However, this is not applicable in the case of IDPs because of the typically very short-lived nature of long-range contacts within IDPs and the spectral crowding that hinders unambiguous NOE assignment. A promising alternative to NOE-based structure determination is paramagnetic relaxation enhancement (PRE). PRE developed into an important tool for characterizing structural ensembles of intrinsically disordered proteins by NMR. Based on the introduction of a paramagnetic label (carrying an unpaired electron) – for example a nitroxide attached to a cysteine or a paramagnetic metal that is chelated by the protein at a known position along the primary sequence– the relaxation rates of neighboring (NMR active) nuclei are enhanced by dipolar interaction with the unpaired electron.[53] These changes in relaxation rates depend on the distance (up to 3.5 nm; dependent on the paramagnetic entity) [53] between the electron and the nucleus. The attachment of the paramagnetic label leads to signal loss in the direct vicinity of the attachment site (by increased relaxation rates) but also at distant sites that are in (transient) contact with the labeled part of the protein (see Figure 3). With NMR and its atomic resolution, one can measure this effect for each amino acid. Thereby one may detect non-

random long-range contacts and compaction along the protein backbone to get a better picture of intra- and intermolecular interactions.[23]

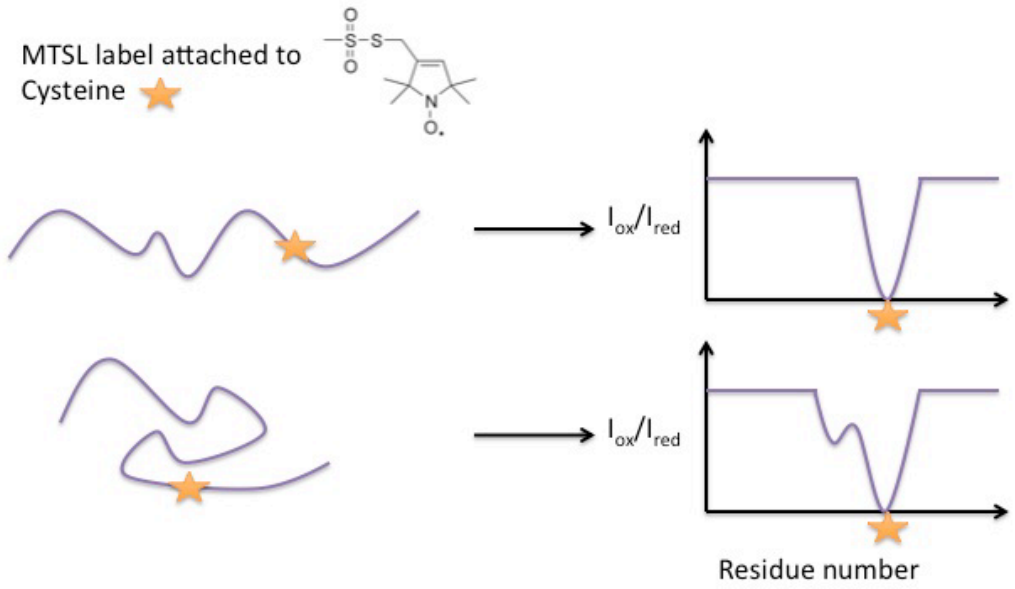


Figure 3: Graphical overview of PRE. The unpaired electron of an MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate)- label is attached at a known site along the primary sequence. This will lead to additional relaxation and consequently decreased signal intensities in its vicinity that can be plotted as a funtion of the primary sequence to identify long- range contacts and compaction. I_{ox}= Intensity of the cross peaks in the HSQC in the oxidized state of MTSL; I_{red}= Intensity of the cross peaks in the HSQC in the reduced state of MTSL after the addition of ascorbic acid.

Although NMR offers many unique advantages, there are numerous limitations for performing experiments on IDRs/IDPs. The major limitation is the protein size. In general, the bigger the protein the more unfavorable the relaxation times although this is not as crucial for IDPs as for global proteins. Yet, for IDPs spectra become increasingly crowded due to the similar chemical environment of the nuclear spins. The latter problem can be partially overcome by selective labeling procedures.[54] However, problems due to overlapping resonances in sequences with tandem repeats remain as these will inevitably lead to similar chemical shifts and signal overlap.

It is furthermore often impossible to measure ^{1}H-^{15}N correlation spectra (^{1}H-^{15}N HSQC/HMQC) at physiological conditions (pH, temperature) as the exchange of solvent exposed amide protons broadens the resonances. This problem can be circumvented by ^{13}C-direct detection- based experiments.[55] ^{13}C direct detection benefits from the larger chemical shift dispersion and additionally prolines give rise to detectable signals, a feature that is missing in conventional ^{1}H-^{15}N correlation spectra.[55]

As NMR provides many different experiments one should also pay attention to the type of desired information. In many cases the choice of an observable like R_2 or HetNOE will entail a length and time scale-window which naturally displays the investigated system through a

filter. Data obtained from NMR are therefore necessarily selective for a certain type of information. Thus, one should always consider combining other biophysical methods with NMR in order to get a better overall picture of the IDP under investigation.

1.2. Changing the structural and dynamical ensemble

As mentioned above NMR is well-suited for the observation of *changes* in observable parameters. A change in the structural ensemble, i.e., a shift in conformational space will almost always entail changes in one of the many NMR observables. In the following the most common processes that lead to structural and dynamical changes will be introduced.

1.2.1. Co-, and post translational modification and the influence of external factors

Despite the disorder, the interactions of IDPs are multifaceted as one IDP can frequently interact with a range of proteins. They need to be tightly controlled otherwise uncontrolled interaction can lead to serious cellular dysfunction.[7] The need to strongly regulate and tune an IDPs function is partially met by post-translational modifications that are used by cells to regulate IDPs. PTMs meet the necessity to quickly adapt in rapidly changing cellular situations.[56] PTMs can be classified by the moiety attached to a protein and the modified residue. Prominent PTMs include reversible cycles of phosphorylation and dephosphorylation, acylation (acetylation, myristoylation, palymitoylation), glysocylation, oxidation, alkylation, sumoylation or ubiquitylation.[57] The vast range of possible PTMs and the various combinations of them can lead to a specific context dependent behavior of the modified IDPs. PTMs can influence important factors including the turnover rate of the protein, the cellular location, the binding properties, the self- association or the secretion.[57] The applied modification can alter a proteins electrostatic or hydrophobic properties or act via steric changes and therefore influence a range of parameters including the stabilization but also a destabilization of local or global secondary structure elements, the intramolecular contacts or interaction properties.[58]

One example for a PTM is phosphorylation which is the reversible replacement of an OH group by a dianionic PO_4^{2-} group through kinases on different residues (serine, threonine and tyrosine in mammalian proteomes). The reverse step is carried out by phosphatases and the interplay between cycles of phosphorylation, interacting proteins and dephosphorylation is carefully balanced in the cell. Bioinformatics studies showed, that among already identified phosphorylation sites, the majority is found to be disordered.[5] Many regulatory pathways include important interaction motifs that are based on phosphorylation. One example for an IDP that is targeted by phosphorylation is 4E-BP2. Upon phosphorylation, the protein undergoes phosphorylation induced folding of a four stranded beta domain and thereby inducing novel intramolecular contacts that drastically decreased the binding to its natural ligand eIF4E.[59]

1.2.2. Influence of the Cellular Environment

Not only PTMs are important for an IDPs structure-function relationship but also the local biochemical environment can alter the interaction properties of the IDP. Since the ensemble of structures of a particular IDP is highly dependent on the cellular environment, one could argue whether non-native *in vitro* conditions induce the disorder, also raising the question what the situation *in vivo* would be. It is now well established, that IDPs can remain disordered also in the context of a cellular environment. The associated experimental approach is in-cell NMR where the protein can be studied in its “native” environment within cells.[60] It is based on measuring the isotopically labeled over-expressed protein directly within cells in an NMR spectrometer. The protein can be either directly measured in the over-expressing cell or introduced into a range of cell types by different approaches including cytoplasmic microinjection, usage of cell penetrating peptides, cell permeabilizing toxins or electroporation.[61, 62] Thereby one can conduct binding studies or look at the impact of PTM in real time in a native-like environment.[63]

External factors like the pH, salt concentration, molecular crowding and temperature can trigger substantial changes in the structural ensemble in some cases but in other cases the protein behaves comparably in the *in vivo* and the *in vitro* condition or the changes are rather subtle.[64] This is an important point for the validity of *in vitro* experiments in NMR solution state and it is important to keep in mind the potential influence of external factors and the structure- function relation of an IDP when one is analyzing the *in vitro* NMR experiments.

1.2.3. Principles of intrinsically disordered protein complexes

For a long time, the key-lock principle[16] dominated our view of protein interactions. But a defined, rigid structure is not a prerequisite for the interaction of two proteins.[3] IDPs can form complexes in many different ways that significantly differ in their understanding of the functional role of disorder. The two extremes in the large spectrum of IDP-partner interaction are the “folding upon binding” and the “fuzzy complex formation”.

“Folding upon binding” has emerged from the observation that some IDPs partially fold into well-defined structures upon interaction with their targets.[65] The interaction sites in specific patches of an IDP, ranging from short regions up to large parts, are frequently already partially folded in the proteins unbound state (often characterized via residual structure propensities). The binding to the partner molecule triggers the final folding of these parts. One well-studied example is the helix formation of phosphorylated kinase inducible domain (pKID) of CREB that binds to KIX.[66] In both, the phosphorylated and the non-phosphorylated KID, the region that undergoes folding upon binding shows a small (~10%) helical content in the unbound form.[67] This observation makes the secondary structure propensity valuable information indicating potential interaction sites. Once folded, the former IDR can be studied by conventional methods including techniques like X-ray crystallography.[68] “Folding upon binding” became very popular and frequently leads to the false assumption that the formation of a rigid structure is a prerequisite for IDP function.

Next to the aforementioned “folding upon binding”, the other extreme approach to explain IDP- partner interaction that can be found in the literature is the “fuzzy complex”. There are examples of IDP complexes where no folding occurs. These interactions are summarized under the term “fuzzy complex”.[69] The IDP remains unfolded in the bound state. In some cases, this enables the IDP to wrap around the partner molecule to get into contact with multiple sites simultaneously. One example is the complex between cyclin dependent kinase inhibitor Sic1 with Cdc4 receptors, which is modulated by phosphorylation of Sic1 and subsequent binding with retention of the mostly disordered state of Sic1.[70] This exemplifies that disorder can be retained in the complex and represents an important functional feature of IDPs.

IDPs can contain short linear interaction motifs that enable a specific interaction with the partner molecule as has been shown by bioinformatic analysis.[71] In contrast, globular proteins frequently feature spatially organized interaction surfaces.

1.3. Intrinsically disordered proteins under investigation

In this section the target proteins of this thesis will be introduced together with their binding partners and most important biological functions.

1.3.1. The human Brain Acid Soluble Protein 1

The Brain Acid Soluble Protein 1 (BASP1, alternative names: neuronal axonal membrane protein NAP-22 and cortical associated protein 23 CAP-23) has been first described as an acidic calmodulin (CaM) binding protein isolated from the growth cone enriched fraction of newborn rat brain and associated with chicken brain development.[33, 72] The rather unusual amino acid composition and the aberrant behavior on SDS- acrylamide gels together with a documented [33] heat stability and denaturation resistance already pointed towards a disordered protein. Indeed, the NMR assignment of human BASP1 does indicate that this protein belongs to the class of IDPs.[35] This protein has further been described as a substrate of Protein Kinase C and a binding partner of Ca^{2+}- loaded CaM with both of these interactions being mutual exclusive.[73] The phosphorylation at Serine 6 of BASP1 by PKC inhibits the binding to CaM and vice versa. Both processes are influenced by the myristoylation of BASP1s N-terminus that occurs co-translational although some BASP1 was reported to be unmodified in the cellular environment.[74] The phosphorylation by PKC yet occurs for both myristoylated and unmyristoylated BASP1 but faster for the acylated protein.[75] The crystal structure of Ca^{2+}- loaded CaM with a N-terminal peptide of BASP1 shows that the myristoyl moiety inserts into a hydrophobic pocket of CaM.[76] Not only is the interaction with proteins influenced by the myristoylation but also the membrane association of BASP1. BASP1 colocalizes with phosphatidylinositol-4,5-bisphosphate (PIP2) in neuronal cell lines.[77] In cells, BASP1 overexpression can lead to neurite outgrowth in neuronlike cell lines that is dependent on both the Serine 6 phosphorylation and the myristoylation.[78]

BASP1 is essential in the development of mice as can be seen from the high postnatal lethality of BASP1 deficient mice that show a lack of nerve sprouting at the neuromuscular junction.[79] Overexpression of BASP1 in transgenic mice induces nerve sprouting at the neuromuscular junction.[80] Concerning the cellular location, BASP1 is found in punctual patterns in neuronal cell, where it is associated with the growth cone,[33, 81] nerve terminals and synaptic vesicle membranes.[82, 83] Additionally, BASP1 can be found in the nucleus of non-neuronal cells.[72, 84, 85] Interestingly, BASP1 has been attributed as a potential tumor suppressor. It was shown that BASP1 protein and mRNA levels are lowered in fibroblasts that are transformed with the v-myc oncogene while the simultaneous expression of both BASP1 and v-myc does not render the transformed phenotype.[86] Additionally, ectopic expression of BASP1 seems to prevent v-myc gene activation and cell transformation.[86]

1.3.2. The human Growth Association Protein 43

Similar to BASP1, NMR investigations on the growth associated protein 43 (GAP-43; alternatively known as neuromodulin, B-50, F1, F-57) show that it is an intrinsically disordered protein.[36] The two proteins share a number of common characteristics yet still remain significantly different. Both are classified as intrinsically disordered proteins and therefore have a typical amino acid composition found for this class of proteins. However, they do not share a significant amount of sequence homology. Both feature a high amount of charged residues and a low amount of bulky, hydrophobic amino acids. Both cluster positive charges on their N-terminus (see Chapter 3). In contrast to BASP1, GAP-43 contains an additional structural element, an IQ- motif.[87] This motif is involved in CaM binding and in the case of GAP-43, this interaction is taking place in both the Ca^{2+}-depleted (apo) and Ca^{2+}-loaded (holo) state of CaM.[68] Opposed to BASP1 where the interaction takes place exclusively with holo- CaM [76] some results found in the literature suggest that GAP-43 binds apo- CaM with a higher affinity.[68] The crystal structure of the GAP-43 IQ- domain with CaM [68] shows an alpha helical adoption of the IQ domain. Another feature that these two IDPs share is PKC interaction.[88] GAP-43 gets reversibly phosphorylated on Serine 41, which is located within the IQ domain.[88] As for BASP1 this phosphorylation abolishes the interaction with CaM [89] and the association to the membrane.[90] Another similarity is the N-terminal acylation. In the case of GAP-43, Cystein 3 and Cystein 4 can undergo palmitoylation, a feature that influences the membrane binding ability of this protein.[91] The palmitoylation is highly heterogeneous, with the majority of the protein being non-acylated and further small shares of the protein being singly or doubly acylated.[92]

GAP-43 plays an important role in neuronal development, as can be seen by the lethality of its knock out in mice that results in malfunction in neuronal pathfinding.[93] An interesting finding is that the normally highly lethal deletion of BASP1 in mice is rendered less fatal by replacing the BASP1 gene with GAP-43 under the promotor of BASP1.[79] This clearly shows an overlap of the two proteins functionalities within the cell. Although both proteins are abundant during neuronal development, they are found to occur at different time points and in different cell types, indicating that their function does not overlap completely.[78] GAP-43 also possesses the ability to bind actin fibrils.[94]

1.4. Membranes in the cell and experimental *in vitro* membrane models for studying protein- lipid interactions

A major part of this work is dedicated to protein-membrane interaction; therefore a brief outline of some biochemical principles of membranes will follow in this section. Biological membranes are in general composed of different lipids and proteins and are crucial for the cells' structure and function.[57] They form a lipid bilayer (~5 nm thickness) that acts as a selectively permeable entity.[57] Typically, membranes are associated with various proteins that can be an integral or peripheral part of the lipid bilayer. The lipids that constitute the lipid bilayer are amphipathic meaning that they have a hydrophobic tail and a hydrophilic, polar head group. The polar head groups point towards the outer water facing side of the bilayer and the hydrophobic tails are facing inwards. The most common lipids are phospholipids that are composed of two hydrocarbon tails with variable length and different polar head groups. Their fatty acid tail can be saturated or contain unsaturated double bonds that affect the packing and thereby the fluidity of the membrane. Prominent phospholipids in mammalian biological membranes include phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, sphingomyelin and inositol phospholipids.[57] The major lipid component that is found in higher eukaryotes next to the phospholipids is cholesterol.[57] Cholesterol orients in the lipid bilayer in a way that the hydroxyl group comes close to the polar head groups of the phospholipids. The lipid composition of a membrane is found to be different in diverse organisms, organs (cell types) and organelles.[57] Importantly, both lipids and proteins possess a certain degree of rotational freedom and the ability for lateral diffusion (with respect to the plane spanned by the membrane). More rarely events of motion within a lipid bilayer include the flip flop which is the transition of a lipid from one monolayer to the other (cytosolic side to extracellular side). This process is catalyzed by specialized enzymes (phospholipid translocators).[57] A major modification of animal cell membranes is the glycosylation of both lipids and proteins.[57] The sugars are attached in the Golgi apparatus and endoplasmic reticulum and are functionally important in, for example, cell recognition by lectins.[57]

A phospholipid bilayer is asymmetric, i.e., the composition of lipids, glycosylation and associated proteins on the inner and outer leaflet can differ on the two sides of the membrane.[57] That asymmetry is functionally important because it enables different molecules to target the head groups that are exposed to the solvent specifically in the intra- or extracellular leaflet. Especially glycosylation is found to be extremely asymmetrically distributed among the two monolayers that compose the lipid bilayer.[57] Glycolipids are found predominantly on the non-cytosolic side. The charged lipids including

phosphatidylserine and inositol phospholipids as well as the neutral phosphoethanolamine are preferentially found on the cytosolic side.[57]

1.4.1. Membrane binding motifs

There are a range of different binding domains in folded proteins that can target a membrane.[95] Prominent examples include the membrane binding motifs first described in PKC, namely the C1 and the C2 domain. While the C1 domain – a zinc finger of approximately 50 residues – specifically binds phorbol esters and diacylglycerol (DAG), the C2 domain is a beta sandwich that in some cases binds to Ca^{2+}.[95] Another example is the Pleckstrin homology domain (~100 residues). It features several basic loops which connect beta strands and a pocket to bind phosphoinositides.[95] Other domains include the FYVE zinc finger, the PX domain, the ENTH and the BAR domains.[95] Next to the well- structured domains, there are reports on membrane associated proteins that bind membranes without a folded domain. One example is MARCKS, which is a membrane anchored IDP that is known to bind phosphoinositides by electrostatic interactions of basic parts of the protein.[96]

1.4.2. Membrane mimics for studying protein- lipid interactions

Mimicking a membrane for *in vitro* experiments is a non- trivial task due to the many variables that need to be considered. In aqueous solution, lipid bilayers or micelles form spontaneously since the polar head groups of the lipids can engage in favorable interaction with water while the hydrophobic tails aggregate to minimize the contact surface with water.[57] It is important to have prior knowledge of the lipid composition as well as the approximate molar percentage of these lipids to ensure the right conditions. Of course buffer salt concentrations and pH are important like in every *in vitro* experiment (see Section Changing the structural and dynamical ensemble) Another factor that needs to be considered is the curvature of the membrane model because some proteins are sensitive to this parameter.[97] For the experimental setup one further needs to consider the phase transition temperature that indicates the transition from the liquid state to a gel or crystalline state (phase) of the membrane. This temperature is influenced by the fluidity of a membrane that is dependent on the lipid composition. Short lipids with more double bonds are less prone to interact with each other and therefore increase the fluidity at lower temperatures.

In general, most biomembrane mimics can be used in combination with NMR. However, a frequent problem is the large size that upon binding will alter the relaxation times of the binding residues of the protein and therefore will drastically reduce their signal intensity (this is shown in Chapter 3 in more detail). Yet, there are different types of mimics that allow the observation of an IDP in the membrane bound state and to look directly at the changes that occur upon binding.

Liposomes

Liposomes are a good model for biological membranes since their composition can be adjusted according to the required distribution of lipids. They structurally resemble spherical lipid bilayers (see Figure 4). Liposomes can be regarded as hollow spheres that usually include an aqueous solution that is separated from the outside of the sphere.[57] They can, for example, be used for co-sedimentation assays for characterization of protein lipid interaction. In brief, they are produced by evaporating an apolar solvent that initially dissolved the monomeric lipids. To prevent oxidation weighting the lipid powder is performed under a stream of nitrogen or argon. Subsequently, the lipids are hydrated with buffer that will yield multilamellar vesicles of roughly 1 μm diameter.[98] In order to control the size of the vesicles, an extruder with a membrane of defined pore size can be employed. This will lead to large unilamellar vesicles with diameters between 100 and 500 nm. To achieve smaller sizes sonication can be used to produce small unilamellar vesicles (< 50 nm diameter).[98] Co-sedimentation assays are based on the fact that the liposomes will form pellets upon centrifugation. Therefore, liposome-bound proteins will be drawn into the pellets, too. This enables the quantification of bound proteins when comparing the amount in the pellet and the supernatant.

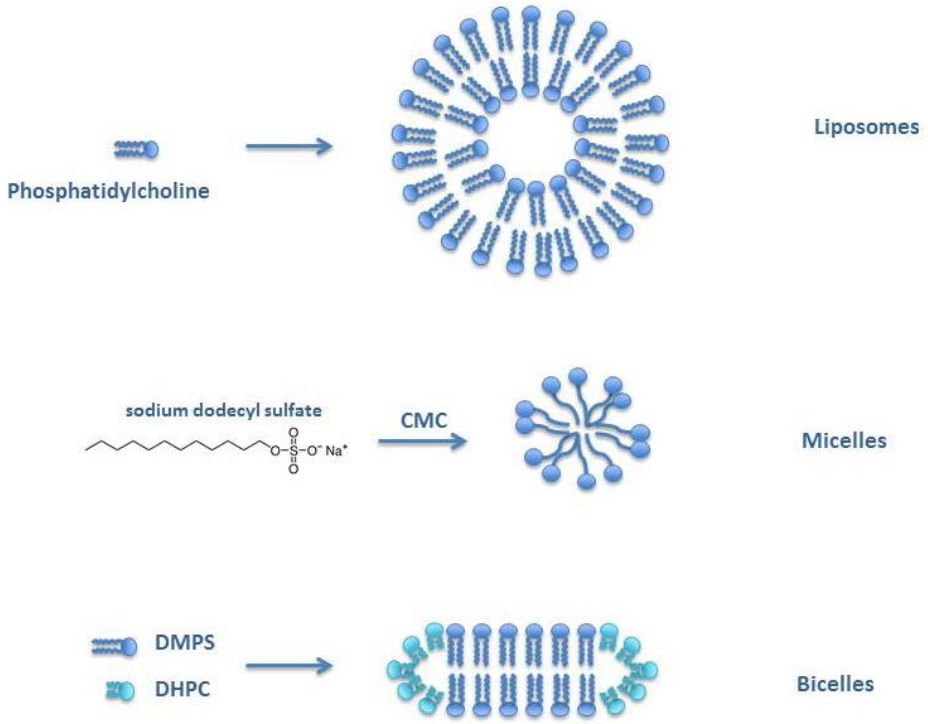


Figure 4: Graphical representation of different membrane mimics that can be used for binding studies.

Micelles

Micelles usually form spontaneously from cone- shaped amphiphiles when mixed with aqueous buffers once the critical micelle concentration (CMC) is reached.[57] The size of a micelle is dependent on the total concentration of the amphiphile, the temperature and the molecule used.[98] Their size is typically 2-20 nm. Micelles can not only be formed in polar solvents, but also in non-polar ones, yielding reverse micelles where the hydrophilic head groups are found on the inside of the micelle and the hydrophobic tails are on the outside.[99] Popular micelles forming agents are sodium dodecyl sulfate (SDS) and dodecylphosphocholine (DPC).[98]

Bicelles

Bicelles are formed by mixing two lipids with different chain size together in order to form a disc-shaped bilayer of one lipid while the other forms a micelle-like assembly to protect the disc from the solvent. A major advantage is their small size, as compared to liposomes, as they are only approximately double the size of a micelle.[98] They represent an authentic lipid bilayer that is preferred over micelles for protein binding studies. Bicelles have been used in this work for a detailed analysis of the structural rearrangements taking place once GAP-43 and BASP1 are bound (see Chapter 3).

Other membrane mimics

An alternative approach is, e.g., the supported lipid bilayer where one side faces a solid surface and the other side faces the buffer solution.[98]

Nanodiscs are composed of a bilayer coated by an amphipathic protein.[98] They constitute suitable membrane models for studying integrated membrane proteins.[98]

1.5. Description of the thesis work

The main topic of this thesis is the formation of intrinsically disordered protein complexes. Several different approaches were combined. The most important biophysical methods used for the data acquisition in the following manuscripts beside nuclear magnetic resonance are surface plasmon resonance (SPR) and liposome co-sedimentation assays. The NMR methods used include assignment procedures and secondary chemical shift analysis, studies of protein dynamics, PRE and chemical shift changes upon titration with a ligand.

Chapter 1 is concerned with the production of isotopically labeled uniformly myristoylated human BASP1 for NMR studies. We observed the occurrence of a modification with both myristic and lauric acid and could show by SPR that these modifications have different liposome binding properties. Furthermore, we showed by NMR that the myristoylation leads to oligomerization of BASP1. Chapter 2 describes the resonance assignment of human GAP-43. A full backbone and side chain assignment could be achieved for a ^{15}N- and ^{13}C-enriched construct by means of higher dimensionality methods as explained in section ‘NMR Resonance Assignment’. Chapter 3 makes use of this assignment for NMR studies of the membrane binding properties of both IDPs. We observe structural rearrangements upon binding of these molecules to different membrane mimics as monitored by NMR.

The scope of the presented thesis covers, hence, the entire process form substrate production over optimization of experimental conditions to the finding and interpretation of biological meaningful and useful information that might help to understand the interaction properties of IDPs.

2. Publications

2.1. Chapter 1

Flamm, A. G.; Le Roux, A.-L.; Mateos, B.; Díaz-Lobo, M.; Storch, B.; Breuker, K.; Konrat, R.; Pons, M.; Coudevylle, N., N-Lauroylation during the Expression of Recombinant N-Myristoylated Proteins: Implications and Solutions. *ChemBioChem,* **2016**,17, 82-89.

ChemPubSoc Europe
DOI: 10.1002/cbic.201500454
CHEMBIOCHEM Full Papers

VIP *Very Important Paper*

N-Lauroylation during the Expression of Recombinant N-Myristoylated Proteins: Implications and Solutions

Andrea Gabriele Flamm,[a] Anabel-Lise Le Roux,[b, c] Borja Mateos,[b] Mireia Díaz-Lobo,[c] Barbara Storch,[d] Kathrin Breuker,[d] Robert Konrat,[a] Miquel Pons,[b] and Nicolas Coudevylle*[a]

Incorporation of myristic acid onto the N terminus of a protein is a crucial modification that promotes membrane binding and correct localization of important components of signaling pathways. Recombinant expression of N-myristoylated proteins in *Escherichia coli* can be achieved by co-expressing yeast N-myristoyltransferase and supplementing the growth medium with myristic acid. However, undesired incorporation of the 12-carbon fatty acid lauric acid can also occur (leading to heterogeneous samples), especially when the available carbon sources are scarce, as it is the case in minimal medium for the expression of isotopically enriched samples. By applying this method to the brain acid soluble protein 1 and the 1–185 N-terminal region of c-Src, we show the significant, and protein-specific, differences in the membrane binding properties of lauroylated and myristoylated forms. We also present a robust strategy for obtaining lauryl-free samples of myristoylated proteins in both rich and minimal media.

Introduction

Over the last twenty years, important progress in molecular biology and recombinant technologies has led to new applications and areas of research in the field of protein NMR. The ability to produce, by recombinant technologies, pure, highly concentrated, and isotopically labeled samples has allowed both solution- and solid-state protein NMR to gain insight into the structural dynamics of many biomolecular systems with an atomic resolution. On the other hand, the very use of recombinant technology has led the field largely to neglect the importance of post-translational modifications (PTMs), an indispensable step in the maturation of proteins towards their functional forms that regulate the activity, localization, stability, and physicochemical properties of proteins. PTMs have become especially relevant in the study of intrinsically disordered proteins (IDPs), which are attracting considerable interest among the biomolecular NMR community. Indeed, IDPs are more susceptible to undergoing PTMs than folded proteins.[1] Because of the inherent flexibility of these proteins, PTMs will have an important impact on their activity, cellular localization, and interaction properties, by modulating their structural dynamics.[2–8] For example, N-terminal acetylation of α-synuclein[9] increases the helical propensity of the N-terminal segment[7] and enhances the affinity of α-synuclein for calmodulin by a factor of 10.[10] Thus, neglecting the impact of PTMs on the general properties of proteins will necessarily lead to an inaccurate description of their biochemical properties and ultimately of their physiological functions.

N-Myristoylation is the covalent attachment, catalyzed by the enzyme N-myristoyl transferase (NMT), of a 14-carbon saturated fatty acid to the N-terminal glycine residue through an amide bond; in eukaryotic systems it usually occurs co-translationally. N-Myristoylation is a very common PTM. By increasing the hydrophobicity of the modified protein, myristoylation is generally involved in membrane binding, targeting, and subcellular trafficking.[11] As such, many myristoylated proteins are involved in important physiological processes such as signaling pathways, oncogenesis, or viral replication.[12] Myristoylation can also be used to tune the activity of a protein through the effect known as the myristoyl-electrostatic switch. This effect consists of a conformational change of the protein (usually triggered by the binding of a ligand) that will expose the myristoyl group, previously sequestered in a hydrophobic

[a] *A. G. Flamm,[+] Dr. R. Konrat, Dr. N. Coudevylle*
Department of Computational and Structural Biology
F. Max Perutz Laboratories, University of Vienna
Campus Vienna Biocenter 5, 1030 Vienna (Austria)
E-mail: nicolas.coudevylle@univie.ac.at

[b] *A.-L. Le Roux,[+] B. Mateos, Dr. M. Pons*
Biomolecular NMR Laboratory, Department of Organic Chemistry
University of Barcelona
Baldiri Reixac 10–12, 08028 Barcelona (Spain)

[c] *A.-L. Le Roux,[+] M. Díaz-Lobo*
Institute for Research in Biomedicine (IRB Barcelona)
Baldiri Reixac 10–12, 08028 Barcelona (Spain)

[d] *B. Storch, Dr. K. Breuker*
Institute of Organic Chemistry
Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck
CCB, Innrain 80/82, 6020 Innsbruck (Austria)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201500454.

© 2015 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

pocket.[13] This mechanism is at the core of the function of proteins such as the ADP ribosylation factor or recoverin. Finally, myristoylation appears to play an important role in apoptosis, because many proteolitic products of caspase 3 are myristoylated and subsequently up- or down-regulate apoptosis.[14]

Myristoylation does not naturally occur in *Escherichia coli*, the most commonly used organism for protein expression. Therefore, to obtain large amounts of isotopically labeled myristoylated proteins suitable for biomolecular NMR studies, two options are available. The first one is in vitro myristoylation, which requires, additionally to the protein of interest, purified NMT, myristoyl-CoA (the co-substrate of NMT), and an additional purification step to separate the myristoylated protein from the by-products. An alternative approach is based on the co-expression of the NMT and the substrate protein by use of a bicistronic vector, in order to enable in-cell modification.[15] However, Liu et al. reported that in minimal medium this procedure leads to mixtures of myristoylated and lauroylated forms of the recombinant protein.[16] However, the factors leading to lauroylation remain unclear, and the degree of lauroylation seems to be highly variable and inconsistent. Indeed, other studies based on this myristoylation strategy (both in rich and in minimal media) did not report the formation of a lauroylated form.[15,17,18]

Here we confirm the observation of Liu et al. and demonstrate that myristoylated and lauroylated forms of the same protein exhibit significantly different lipid-binding properties, emphasizing the need to ensure the homogeneity of the attached fatty acid chains. Consequently, we also report an optimized strategy for obtaining pure myristoylated proteins. This strategy has been applied to the production of two myristoylated proteins: brain acid soluble protein 1 (BASP1, also known as NAP-22 and CAP-23) and the first two domains of c-Src: namely the Unique and SH3 domains (USH3). BASP1 is a 25 kDa intrinsically disordered protein highly abundant in the brain during development, involved in growth cone guidance and actin cytoskeleton organization[19] and interacting with holo-CaM specifically in its myristoylated form.[20] c-Src is the leading member of the Src family of non-receptor tyrosine kinases (SFKs), which are involved in many signaling pathways. Deregulation of these kinases, and in particular of c-Src itself, affects cell migration, proliferation, and survival, all of which contributes to its oncogenic potential. The N-terminal (SH4) region of c-Src is co-translationally myristoylated at the N-terminal glycine unit. SH4 is a basic peptide situated at the beginning of the Unique Domain of Src, an intrinsically disordered domain not conserved among the Src family. Myristoylated SH4 is responsible for c-Src membrane anchoring, through concurrent hydrophobic and electrostatic interactions.

Results and Discussion

Expression of recombinant N-myristoylated protein in minimal medium leads to a mixture of N-myristoylated and N-lauroylated protein

We expressed recombinantly myristoylated BASP1 (MyrBASP1) by using the bicistronic vector designed by Glück et al.[15] In order to measure the extent of myristoylation by mass spectrometry, we first performed the expression in minimal medium supplemented with unlabeled (^{14}N) ammonium chloride. The LC-MS analysis revealed the presence of two species differing in mass by 27 Da (Figure 1), thus suggesting that

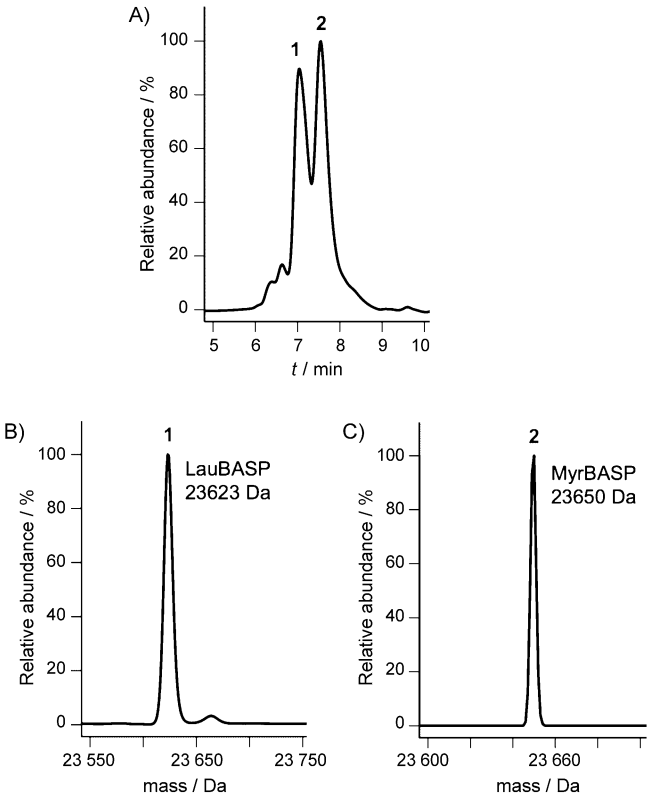


Figure 1. Expression of BASP1 in minimal medium leads to two different species. A) RP-HPLC analysis of purified BASP1 co-expressed with NMT in minimal medium. B) and C) MS analysis of products corresponding to peaks 1 and 2, respectively.

both N-myristoylated and N-lauroylated forms of BASP1 (differing in mass by 28 Da) had been produced, as reported previously in the case of Arf1.[16] NMR analysis of the putative (later confirmed) myristoylated BASP reveals that the N-terminal acylation leads to a disappearance of the $^{1}H,^{15}N$ HSQC crosspeaks corresponding to the N terminus (Figures S1 and S2 in the Supporting Information), probably due to micelle formation.

Further MS analysis of the two species by means of electron capture dissociation (ECD) and collisionally activated dissociation (CAD) confirmed that the difference in mass is 28.0302 ± 0.0035 Da (CH_2CH_2, calcd mass 28.0313 Da) and is due to myris-

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

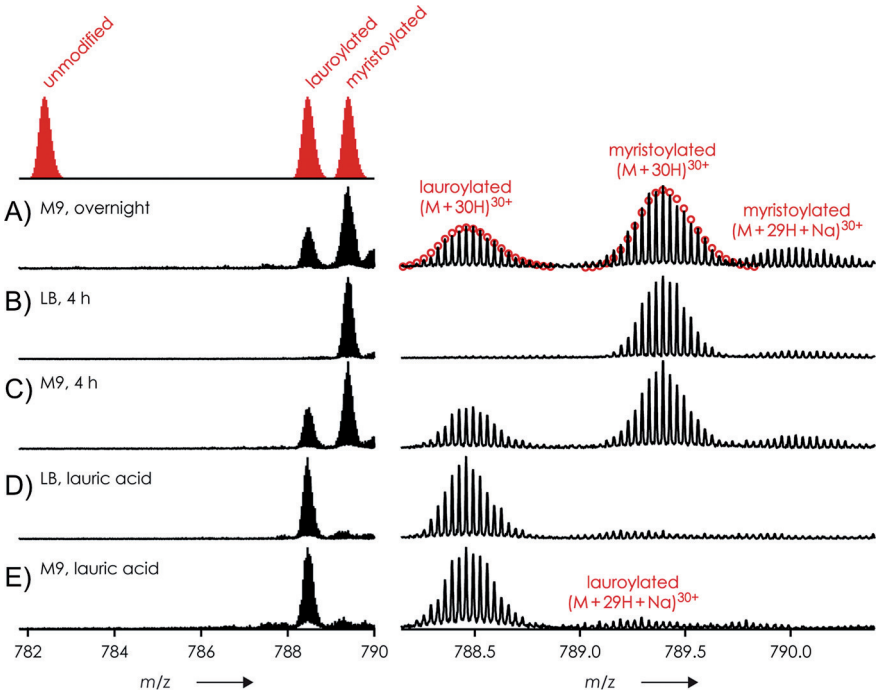


Figure 2. Proportion of N-myristoylated and N-lauroylated forms of BASP under different expression conditions. ESI spectra of human BASP co-expressed with hNMT in A) M9 overnight, supplemented with myristic acid, B) LB for 4 h, supplemented with myristic acid, C) M9 for 4 h, supplemented with myristic acid, D) and E) LB and M9, respectively, supplemented with lauric acid; red circles in (A) show calculated isotopic profiles for $[M+30H]^{30+}$ ions of lauroylated and myristoylated protein.

toylation and lauroylation at the N terminus of BASP1 (Supporting Information, data 1).

In order to explore the origin of the lauroylated form of BASP1 (LauBASP1), we measured the proportions of N-myristoylated and N-lauroylated forms of BASP1 in samples expressed under different conditions (Figure 2). When expressed in rich medium (lysogeny broth, LB), BASP1 is only found in its N-myristoylated form (Figure 2B). When expressed in minimal medium BASP1 is found in both its N-myristoylated and N-lauroylated forms, and the proportion of the N-lauroylated form increases with the length of the expression time (Figure 2A and C). This suggests that under scarce conditions (minimal medium and/or long expression times), myristoyl-CoA is converted into lauroyl-CoA through β-oxidation and then used by the NMT (which is known to show weak discrimination between myristoyl-CoA and shorter acyl-CoA) to acylate BASP1. Expressing BASP1 in the presence of lauric acid only leads to the lauroylated form of BASP1 both in rich and in minimal medium (Figure 2D and E). The same observations were made for USH3 (with a different expression strain of *E. coli*), regardless of whether it is expressed in rich or minimal medium (see below), thus suggesting that this problem is not specific to BASP1 but is instead a general problem that occurs especially in minimal medium but can also be observed in rich medium.

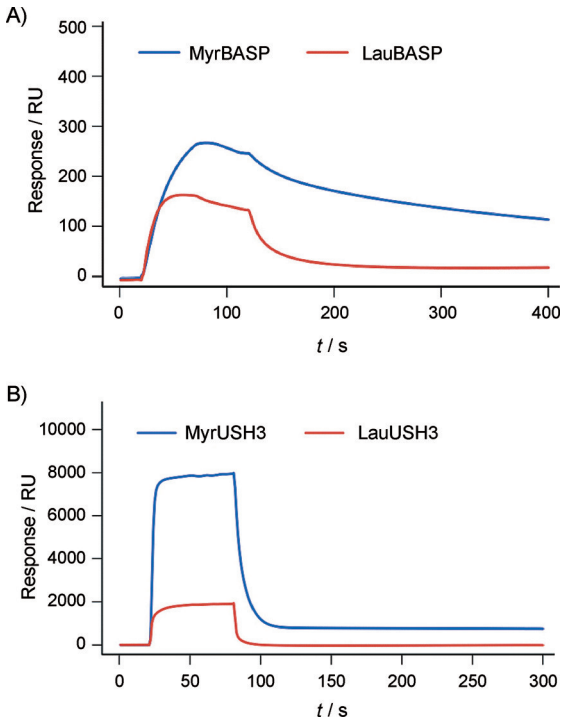


Figure 3. SPR analysis of the binding properties of different acylated forms of BASP and USH3 towards liposomes. A) Real-time SPR response showing binding of MyrBASP and LauBASP to surface-immobilized DMPC/PS (3:1) liposomes. The protein concentration was 0.2 μM. B) Real-time SPR response showing binding of MyrUSH3 and LauUSH3 to surface-immobilized DMPC/DMPG (2:1) liposomes; the protein concentration was 20 μM.

N-Myristoylated and N-lauroylated forms of the same protein have different biochemical properties

To test whether the biochemical properties of BASP1 and USH3 are affected by the length of the N-acyl chain, we per-

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

formed surface plasmon resonance (SPR) experiments with immobilized liposomes. Because this method does not require isotopic labeling, we obtained fully myristoylated or lauroylated samples by expression in rich medium supplemented with the corresponding fatty acid (lauric or myristic acid). In the case of BASP1 (Figure 3A), the binding to 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/L-α-phosphatidylserine (DMPC/PS) liposomes of LauBASP1 (K_D = 86 nM) shows an affinity reduced by a factor of 10 relative to MyrBASP1 (K_D = 7.9 nM). Additionally, the length of the acyl chain also seems to affect the dissociation behavior because MyrBASP1 dissociates from the liposome more slowly than LauBASP1 (Figure 3A).

In the case of myristoylated USH3 (MyrUSH3), the difference is even more pronounced. Lauroylated USH3 (LauUSH3) binding to DMPC/DMPG (2:1; DMPG = 1,2-dimyristoyl-*sn*-glycero-3-phospho-*rac*-(1-glycerol)) liposomes is four times weaker than for MyrUSH3 (Figure 3B). Dissociation of the lauroylated form from the liposomes is also faster than for the myristoylated form.

Our SPR results clearly show that the length of the N-acyl chain influences biochemical properties. Therefore the two species should be separated in order to obtain biochemically pure samples.

Production of pure N-myristoylated USH3

Our SPR results show that the lengths of the N-acyl chains influence the properties of N-acylated proteins. Therefore, it is necessary to obtain a pure N-myristoylated sample for the pursuit of subsequent biochemical and NMR studies. As described above, the lauroylated and myristoylated forms of BASP1 can be easily separated by RP-HPLC. However, the use of organic solvent is not generally applicable to proteins containing folded domains, as in the case of USH3.

After protein expression, the soluble fraction of USH3 elutes from the size-exclusion column in the form of three different peaks (Figure 4A). SDS-PAGE (Figure 4B) reveals that peak 1 contain both USH3 and NMT, whereas peak 2 only contain USH3 (peak 3 contains a degraded form of the protein). Further MS analysis (Figure 4C and D) showed that in peak 1 (which contains both USH3 and NMT) USH3 is exclusively myristoylated whereas in peak 2 USH3 (Figure 4E) is only found in its unmodified and lauroylated forms. Hence, it appears that MyrUSH3 co-purifies with NMT whereas LauUSH3 remains free of NMT due to the lower affinity of the enzyme for the shorter acyl chain forms.

This co-purification might be exploitable, because the purification of the complex could lead to a fully myristoylated sample. However, we were not able to dissociate the complex fully, either by chromatographic methods (such as ion exchange or hydrophobic separation) or by the use of myristoylated coenzyme A as a competitor as suggested elsewhere.[21] Consequently, we applied and refined an alternative method originally proposed by Ha et al.[22] This exploits the strong affinity of the myristoylated protein towards membrane: a significant amount of myristoylated protein remains in the cell pellet

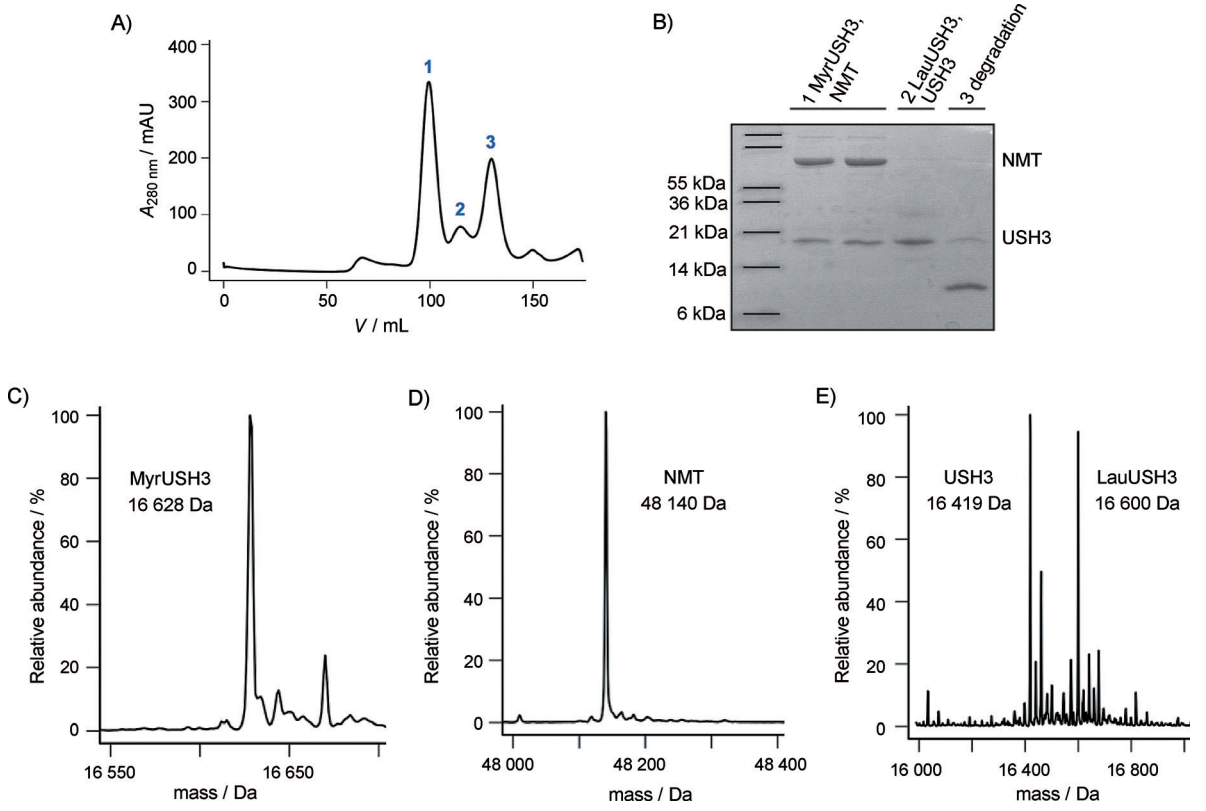


Figure 4. A) Size-exclusion chromatography profiles of acylated USH3 purified from the supernatant. B) SDS-PAGE analysis of pooled fractions of the peaks observed in (A). C) and D) MS analysis of peaks 1. E) MS analysis of the peaks 2.

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

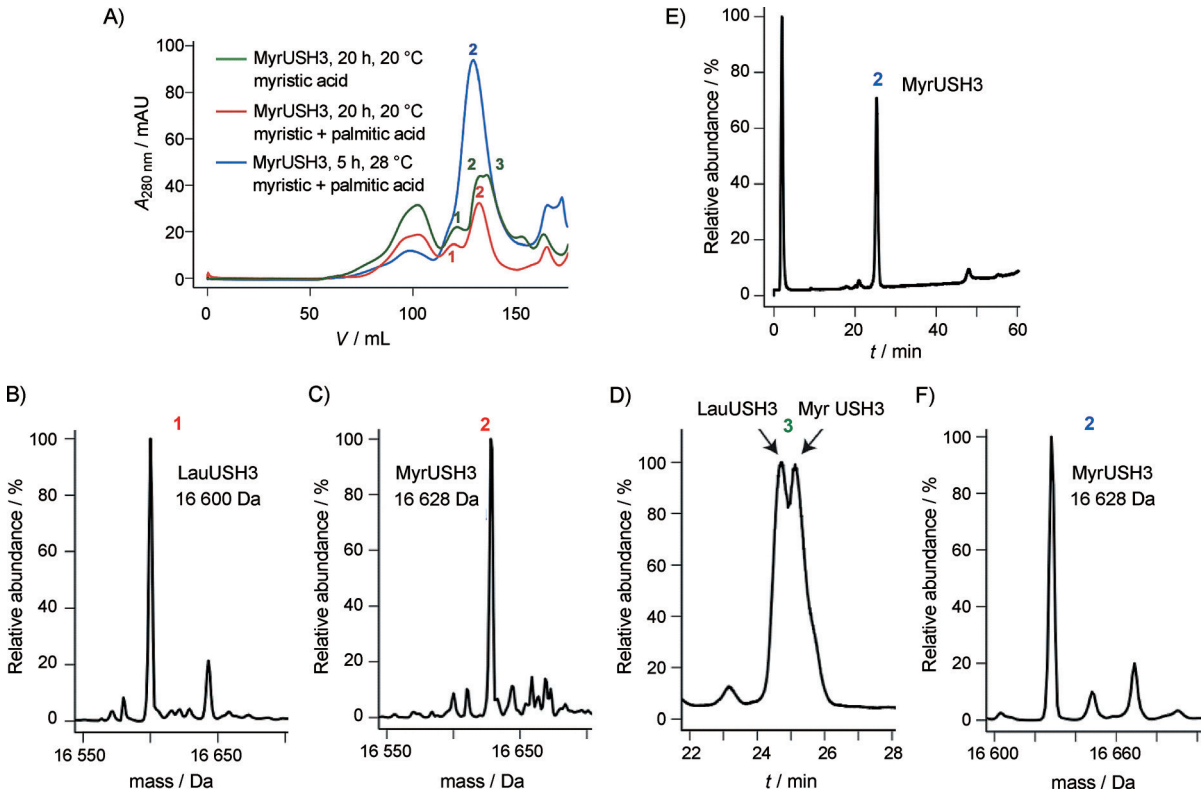


Figure 5. A) Size-exclusion chromatography profiles of acylated USH3 expressed under different conditions. Red curve: in Rosetta cells, in the presence of both myristic and palmitic acid (20 h, 20 °C), two elution peaks are observed. Green curve: in T7 cells, in the presence only of myristic acid (20 h, 20 °C), three elution peaks are observed. Blue curve: in T7 cells, in the presence of both myristic and palmitic acid (5 h, 28 °C), one elution peak is observed. B) and C) Mass spectrometric analysis after separate pooling of the fractions corresponding to peaks 1 and 2, respectively. D) RP-HPLC analysis of peak 3 of the red curve in (A) after pooling of its fractions. E) and F) RP-HPLC and MS analysis of peak 2 of the blue curve after pooling of its fractions.

upon centrifugation, whereas most of the NMT remains in the supernatant. This procedure relies on the resuspension and resolubilization of the membrane-bound protein, with use of Triton, before further purification. This procedure leads to a mixture of lauroylated and myristoylated USH3 that can be separated by size-exclusion chromatography. As can been seen from the size-exclusion profile (Figure 5A, red curve), the first peak corresponds to LauUSH3 (Figure 5B) whereas the second corresponds to MyrUSH3 (Figure 5C). Intriguingly, when large amounts of LauUSH3 are produced (Figure 5A, green curve), a third peak corresponding to a mixture of Myr and LauUSH3 is observed (Figure 5D). This last peak merges with the MyrUSH3 peak, compromising its purity. When the expression conditions are optimized to minimize the production of LauUSH3 (see below for a detailed explanation), pure MyrUSH3 can easily be obtained (Figure 5A, blue curve). Thus, in order to be able to prepare pure myristoylated protein efficiently, the expression protocols have to be optimized to minimize the presence of lauroylated forms before the purification. In addition, size separation of the two acylated forms is a property of our particular system and might not occur for other proteins, thus emphasizing the necessity to use conditions that ensure the presence of pure myristoylated proteins.

We hypothesized that lauroylCoA originates from β-oxidation of added myristic acid. This hypothesis would explain the observation that lauroylation increases under conditions of limited availability of carbon sources (minimal medium) or after extended expression (presumably associated with nutrient depletion). Thus, one of the strategies that we tested was the addition of palmitic acid. NMT does not incorporate palmitoyl groups, but β-oxidation of palmitic acid would contribute to replenish the myristoylCoA pool.

Table 1 summarizes the effects of different expression factors on the relative amounts of MyrUSH3 and LauUSH3 obtained

Table 1. Effects of expression conditions in LB medium on the relative amounts of myristoylated and lauroylated USH3.

Cell type	Acyl chain addition	*t* [h] after induction	Lauroylated USH3 [%]	Myristoylated USH3 [%]
Rosetta	myristic acid	5[a]	20	80
T7	myristic acid	5[a]	20	80
Rosetta	myristic + palmitic acid	5[a]	5	95
T7	myristic + palmitic acid	5[a]	2	98
Rosetta	myristic acid	20[b]	80	20
T7	myristic acid	20[b]	35[c]	65[c]
Rosetta	myristic + palmitic acid	20[b]	18	82
T7	myristic + palmitic acid	20[b]	3	97

The relative amounts of myristoylated and lauroylated USH3 were determined by integrating the elution peaks from size-exclusion chromatography performed on the mixtures obtained after the Triton wash purification procedure. [a] At 28 °C. [b] At 20 °C. [c] Myristoylated and lauroylated species are poorly separated.

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

after the Triton wash procedure, nickel column purification, and size-exclusion chromatography. The induction time appears to be the most important factor, because longer expression times lead to greater amounts of lauroylated form. Nevertheless, some factors such as the expression strain or the addition of palmitic acid to the growth medium also have a significant impact on the myristoylation level. More specifically, expression in T7 cells yields lower lauroylation levels than expression in Rosetta cells. In rich medium, USH3 is found as a mixture of myristoylated and lauroylated forms even after short induction times (whereas under the same conditions BASP1 was fully myristoylated). However, adding palmitic acid to the growth medium leads to almost pure myristoylated USH3. To sum up, the set of best conditions consists of a short induction time in the presence of both palmitic and myristic acid in the expression medium. Use of these conditions together with the Triton wash purification procedure leads to pure and homogeneous natural-abundance MyrUSH3 samples (Figure 5 A, blue curve, and E/F). However, expression in minimal media, required in order to produce isotopically labeled proteins, did not provide pure MyrUSH3, presumably due to limited availability of carbon sources. This limitation was overcome by using the Marley method, which consists of generating cell mass in unlabeled rich media and subsequent transfer into labeled media just before induction.[23] In that case we were able to obtain sufficient amounts of labeled proteins (95%) to enable NMR measurement. Finally, this protocol was also successfully used for the expression and purification of MyrBASP1, thus demonstrating the robustness of the method.

Conclusion

Myristoylation is an important post-translation modification generally involved in subcellular trafficking and membrane association. Correct structural characterization of myristoylated proteins requires the production of their isotopically labeled forms. Co-expression of yeast NMT with the protein of interest in *E. coli* provides an efficient method but is complicated by the simultaneous formation of lauroylated and myristoylated proteins. Although they are chemically similar, SPR experiments showed that the variation in chain length greatly alters the lipid-binding properties. Therefore, a suitable method to obtain purely myristoylated proteins in *E. coli* needs to be developed. We suggest here that the addition of the shorter acyl chain appears to be due to the conjunction of two factors: 1) the availability of a shorter acyl-CoA generated through β-oxidation of the myristic acid that needs to be supplemented to the growth medium, and 2) the poor discrimination of the human NMT for the shorter acyl-CoA.[24] On the basis of this hypothesis, we have shown that addition of palmitic acid, which is not incorporated into the protein by NMT but can resupply the MyrCoA pool, contributes, under some conditions, to minimize the formation of the lauroylated form.

The capacity to isolate lauroylated and myristoylated proteins from mixtures of the two is dependent upon the physiochemical properties of the proteins of interest. In the case of a fully disordered protein (BASP1), the two forms can be separated by RP-HPLC thanks to their different hydrophobicities. In the case of a partially or fully folded protein (USH3), we had to develop a more sophisticated procedure based on 1) optimization of the expression conditions to minimize the formation of LauUSH3, 2) use of the Triton wash purification protocol to obtain NMT-free myristoylated protein with a low proportion of the lauroylated form, and 3) final purification by affinity and size-exclusion chromatography. Interestingly, although both proteins have acyl moieties attached to a disordered segment, their biochemical properties seem to be affected in different ways by the length of the acyl chain. It is known that myristoylation by itself is not enough for stable membrane attachment. A second binding event is usually required.[11] Examples include electrostatic interactions, a second aliphatic anchor, or additional contacts with amphipathic regions of the peptide backbone. Clearly, the myristoyl chain plays a stronger role in lipid binding by USH3 than by BASP1. This might be due to the proximity of the charged residues and the acyl chain in USH3, coupling the strength of the electrostatic interaction with the insertion of the acyl chain in the lipid bilayer, or might be the result of an indirect effect caused by modulation of the SH4 and SH3 interactions by the presence of distinct fatty acid chains attached to the SH4 domain.[25]

Thus, if pure lauroylated and myristoylated forms of the same protein can be obtained independently or separated from mixtures, the comparison of their lipid-binding properties provides additional insight into the processes involved in fatty-acid-mediated lipid interactions.

To summarize, we provide evidence as to why the production of myristoylated proteins in *E. coli* by co-expressing NMT results in a heterogeneously N-acylated sample. We have demonstrated that the different N-acylated forms have different biochemical properties and we have developed an expression/purification protocol to generate a homogeneously N-myristoylated sample. We expect these observations to be especially relevant for the conduct of NMR studies in which the homogeneity of the sample is essential.

Experimental Section

Protein expression and purification of myristoylated BASP1: Expression was achieved with *E. coli* strain T7 Express (New England BioLabs) and the bicistronic vector pETDuet-1Δ6His_hNMT_hBASP1_6His as already described.[15] Expression of BASP1 for the analysis of the protein by mass spectrometry was done either in rich medium (LB medium) or in M9 minimal medium with unlabeled (^{14}N) ammonium chloride (1 g L^{-1}) and (^{12}C) glucose (4 g L^{-1}). In order to produce myristoylated protein, myristic acid was added to the growth medium at a final concentration of 50 μM 10 min before induction. Fresh myristic acid stock solution was prepared as described elsewhere.[15] The cells were grown at 37 °C to an OD_{600} of 0.8 and induced by adding isopropyl β-D-thiogalactopyranoside (IPTG, 0.8 mM). The expression temperature was 28 °C for 4 h or overnight. The cells were pelleted by centrifugation at 2862 *g* at 4 °C for 20 min. The pellet was resuspended in lysis buffer (PBS) with protease inhibitors (cOmplete, Mini Protease Inhibitor Tablets, EDTA-free, Roche) and sonicated on ice with a Branson W-450 D sonifier with a microtip (3 min, 50% amplitude) before centrifugation at 36223 *g* for 20 min at 4 °C. The super-

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

natant was applied to a Ni^{2+}-loaded HisTrap HP column (5 mL, GE Healthcare) that was pre-equilibrated with PBS. The column was washed with PBS and high-salt PBS [NaCl (1.5 M), imidazole (20 mM)] and five column volumes of PBS before elution with a linear gradient of PBS/PBS with imidazole (0.5 M).

FT-ICR mass spectrometry: Protein samples were desalted with Vivaspin 500 centrifugal concentrators (Sartorius, Germany, PES membrane, MWCO 5000) as described previously[26] Briefly, a protein solution (≈10 μM, 500 μL) was concentrated to 100 μL, and an aqueous ammonium acetate solution (100 mM, 400 μL) was added. The process was repeated five times, followed by six cycles of concentration and dilution with H_2O. For ESI (flow rate 1.5 μL min^{-1}), desalted protein was diluted to ≈2 μM in H_2O/CH_3OH 1:1 with acetic acid (1% vol) as additive. H_2O was purified to 18 MΩ·cm at room temperature with a Milli-Q system (Millipore, Austria); CH_3OH (Acros, Austria) was HPLC-grade. Experiments were performed with a 7 Tesla Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer equipped with an electrospray ionization source, a collision cell for CAD, and a hollow dispenser cathode for ECD.

Myristoylated USH3 expression: USH3 expression was performed in *E. coli* Rosetta (DE3)pLysS cells (Novagen) or T7 Express (New England BioLabs) cells and using the bicistronic vector pETDuet-1Δ6His_hNMT_USH3_6His. The cells were grown at 37 °C to an OD_{600} of 0.8, and 10 min before induction with IPTG (final concentration 1 mM, Melford), freshly prepared myristic and/or palmitic acid (Sigma) were added to the cell culture, at a final concentration of 200 μM. For the expression of isotopically labeled samples the protocol developed by Marley et al. was used:[23] cells were first grown in LB to an OD_{600} of 0.4 and were then centrifuged at 1000 *g* and 4 °C for 20 min. The pellet was resuspended in half the volume of minimal medium containing ammonium chloride and glucose. After 20 min at 37 °C, myristic and palmitic acid were added as described above. The expression temperature was either 28 °C (for 5 h) or 20 °C (for 20 h). The cells were pelleted by centrifugation at 3993 *g* at 4 °C for 30 min. The pellet was resuspended in lysis buffer [Tris·HCl (20 mM), NaCl (300 mM), imidazole (5 mM), pH 8], to which protease inhibitors (protein inhibitor cocktail and phenylmethanesulfonyl fluoride (PMSF); 1 mM), both from Sigma) were added.

MyrUSH3 purification: The resuspended pellet was sonicated on ice before centrifugation at 25 000 rpm for 45 min at 4 °C. The protein appeared to be distributed between the supernatant and the pellet, so two different purification methods were used. For the soluble fraction, the supernatant was applied to a Ni-NTA column (Qiagen) followed by size-exclusion chromatography in a Superdex 75 column in sodium phosphate buffer (Na_3PO_4) [NaP (50 mM), NaCl (150 mM), EDTA (0.2 mM), pH 7.5]. For the insoluble fraction, the pellet was resuspended in lysis buffer containing Triton X-100 (1%). The resuspended pellet was centrifuged again for 30 min at 75 600 *g*, and the procedure was repeated twice or three times. The supernatant from the Triton washes was purified by immobilized metal affinity chromatography as described above. If lauroylated species were present, they eluted from the size-exclusion chromatographic column at an apparent higher molecular weight than the myristoylated ones. This method enabled the two different acylated species to be separated and their respective amounts quantified. A comparison of various expression and purification protocols is presented in the Results section.

LC-MS: The purities and identities of the products were established by UPLC coupled to MS [Acquity chromatograph with a BioSuite pPhenyl column (1000RPC 2.0×75 mm)] coupled to a LCT-Premier spectrometer (Waters corporation).

Liposome preparation: DMPC, DMPG, and PS were purchased from Avanti Polar Lipids, Inc. The lipids were dissolved in chloroform or, in the case of DMPG, in chloroform/methanol/H_2O (65:35:8). Liposomes were prepared by mixing the appropriate amount of lipids in the solvent. The solvent was removed in a rotary evaporator, followed by rehydration and vortexing at 40 °C with the buffer used for SPR analysis, with a final lipid concentration of 1 mM. The different liposomes were prepared with a DMPC/DMPG ratio of 2:1 or a DMPC/PS ratio of 4:1. Large unilamellar vesicles were mechanically extruded at 40 °C by use of a Thermobarrel extruder (10 mL Thermobarrel extruder; Lipex Northerns Lipids Inc. Burnaby, Canada) with at least ten cycles of extrusion and use of a polycarbonate filter (100 nm). To verify the appropriate size of the liposomes, the mean diameter was checked by dynamic light scattering (Zetasizer Nanoseries S, Malvern Instruments).

Surface plasmon resonance (SPR): SPR experiments with BASP1 were carried out with a Biacore 2000 instrument (Biacore, GE Healthcare) and SPR sensor chip (L1, Biacore, GE Healthcare). Liposomes were injected for 500 s at a flow rate of 5 μL min^{-1}. The reference channel was coated with BSA by use of a 200 s injection of BSA (1 mg mL^{-1}, Sigma, fatty acid free) at a flow rate of 10 μL min^{-1}. Protein binding experiments were performed at 50 μL min^{-1}. The interaction of MyrBASP1 or LauBASP1 with liposomes was followed by observing the SPR response when a solution of protein was injected for 100 s (association phase), followed by a 300 s washing period (dissociation phase). MyrBASP1 binding to DMPC/PS liposomes was monitored by injections at concentrations ranging from 10 nM to 2.5 μM. LauBASP1 binding to DMPC/PS liposomes was monitored by injections at concentrations ranging from 20 nM to 7.5 μM. All experiments were performed in the running buffer, which consisted of Na_3PO_4 (20 mM), NaCl (50 mM), pH 7.4. The surface was regenerated with a series of CHAPS (20 mM)/HCl (10 mM)/CHAPS (20 mM) pulses, each pulse for 30 s at 100 μL min^{-1}. Each experiment was started with freshly captured liposomes. Liposome coating was reproducible, with a variation smaller than 4% between subsequent coatings, ensuring very reproducible protein binding curves.

SPR experiments with MyrUSH3 were performed in a very similar fashion, with slight modifications. The SPR chip (a 2D carboxymethyldextran surface) was purchased from Xantec and modified by covalent attachment of phytosphingosine (TebuBio) to allow the capture of DMPC:DMPG (2:1) liposomes. Liposomes were injected at 10 μL min^{-1} for 200 s before the protein binding experiment: MyrUSH3 or LauUSH3 was injected at 50 μL min^{-1} for 60 s, and dissociation was allowed for 300 s. The running buffer was composed of Na_3PO_4 (50 mM), NaCl (150 mM), EDTA (0.2 mM, pH 7.5). Liposome coating was reproducible, with a variation of about 1% between the subsequent coatings, ensuring very reproducible protein binding curves. Data analysis was performed with the Biacore Biaevalution software.

Acknowledgements

The bicistronic vector encoding NMT was a kind gift from Dr. Dieter Willbold (Forschungszentrum Juelich Institute of Complex Systems). This work was supported by grants from the Austrian Science Foundation (FWF P26317 and Y372, to K.B.), as well as

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

by funds from MINECO (BIO2013–45793-R) and la Fundació Marató TV3. A.L.L.R. holds an IRB/La Caixa fellowship. We thank Markus Hartl and the Mass Spectrometry Facility (Campus Science support Facilities) of the F. Max Perutz Laboratories in Vienna, as well as Christian Becker and Aleksandr Kravchuk from the Institute of Biological Chemistry in Vienna. We gratefully acknowledge the technical help of Marta Taules (CCiTUB) in the SPR experiments and the support of Maria Antonia Busquets (Physical Chemistry Department, Faculty of Pharmacy, UB) in the preparation of liposomes, as well as the Mass Spectrometry Core Facility, in particular Marta Vilaseca (Institute for Research in Biomedicine). We also thank Mariano Maffei for his input.

Keywords: intrinsically disordered proteins · lauroylation · myristoylation · surface plasmon resonance · transferases

[1] G. A. Khoury, R. C. Baliban, C. A. Floudas, *Sci. Rep.* **2011**, *1*, 90.
[2] N. Errington, A. J. Doig, *Biochemistry* **2005**, *44*, 7553 – 7558.
[3] B. Meyer, H. Möller, *Top. Curr. Chem.* **2006**, *267*, 187 – 251.
[4] J. Liu, J. R. Faeder, C. J. Camacho, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19819 – 19823.
[5] Y. L. Deribe, T. Pawson, I. Dikic, *Nat. Struct. Mol. Biol.* **2010**, *17*, 666 – 672.
[6] A. H. Mao, S. L. Crick, A. Vitalis, C. L. Chicoine, R. V. Pappu, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8183 – 8188.
[7] A. S. Maltsev, J. Ying, A. Bax, *Biochemistry* **2012**, *51*, 5004 – 5013.
[8] F. X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, *Chem. Rev.* **2014**, *114*, 6661 – 6714.
[9] T. Bartels, J. G. Choi, D. J. Selkoe, *Nature* **2011**, *477*, 107 – 110.
[10] J. M. Gruschus, T. L. Yap, S. Pistolesi, A. S. Maltsev, J. C. Lee, *Biochemistry* **2013**, *52*, 3436 – 3445.
[11] M. D. Resh, *Biochim. Biophys. Acta Mol. Cell Res.* **1999**, *1451*, 1 – 16.
[12] M. D. Resh, *Nat. Chem. Biol.* **2006**, *2*, 584 – 590.
[13] S. McLaughlin, A. Aderem, *Trends Biochem. Sci.* **1995**, *20*, 272 – 276.
[14] D. D. Martin, E. Beauchamp, L. G. Berthiaume, *Biochimie* **2011**, *93*, 18 – 31.
[15] J. M. Glück, S. Hoffmann, B. W. Koenig, D. Willbold, *PLoS One* **2010**, *5*, e10081.
[16] Y. Liu, R. A. Kahn, J. H. Prestegard, *Structure* **2009**, *17*, 79 – 87.
[17] S. Breuer, H. Gerlach, B. Kolaric, C. Urbanke, N. Opitz, M. Geyer, *Biochemistry* **2006**, *45*, 2339 – 2349.
[18] S. Lim, I. Peshenko, A. Dizhoor, J. B. Ames, *Biochemistry* **2009**, *48*, 850 – 862.
[19] D. Frey, T. Laux, L. Xu, C. Schneider, P. Caroni, *J. Cell Biol.* **2000**, *149*, 1443 – 1454.
[20] M. Matsubara, T. Nakatsu, H. Kato, H. Taniguchi, *EMBO J.* **2004**, *23*, 712 – 718.
[21] C. R. Morgan, B. V. Miglionico, J. R. Engen, *Biochemistry* **2011**, *50*, 3394 – 3403.
[22] V. L. Ha, G. M. Thomas, S. Stauffer, P. A. Randazzo, *Methods Enzymol.* **2005**, *404*, 164 – 174.
[23] J. Marley, M. Lu, C. Bracken, *J. Biomol. NMR* **2001**, *20*, 71 – 75.
[24] T. A. Neubert, R. S. Johnson, J. B. Hurley, K. A. Walsh, *J. Biol. Chem.* **1992**, *267*, 18274 – 18277.
[25] M. Maffei, M. Arbesu, A. L. Le Roux, I. Amata, S. Roche, M. Pons, *Structure* **2015**, *23*, 893 – 902.
[26] M. Hartl, A. M. Mitterstiller, T. Valovka, K. Breuker, B. Hobmayer, K. Bister, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4051 – 4056.

Manuscript received: September 5, 2015
Accepted article published: November 2, 2015
Final article published: December 3, 2015

 © 2016 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

CHEMBIOCHEM

Supporting Information

N-Lauroylation during the Expression of Recombinant N-Myristoylated Proteins: Implications and Solutions

Andrea Gabriele Flamm,[a] Anabel-Lise Le Roux,[b, c] Borja Mateos,[b] Mireia Díaz-Lobo,[c] Barbara Storch,[d] Kathrin Breuker,[d] Robert Konrat,[a] Miquel Pons,[b] and Nicolas Coudevylle*[a]

cbic_201500454_sm_miscellaneous_information.pdf

NMR Analysis of Myristoylated BASP:

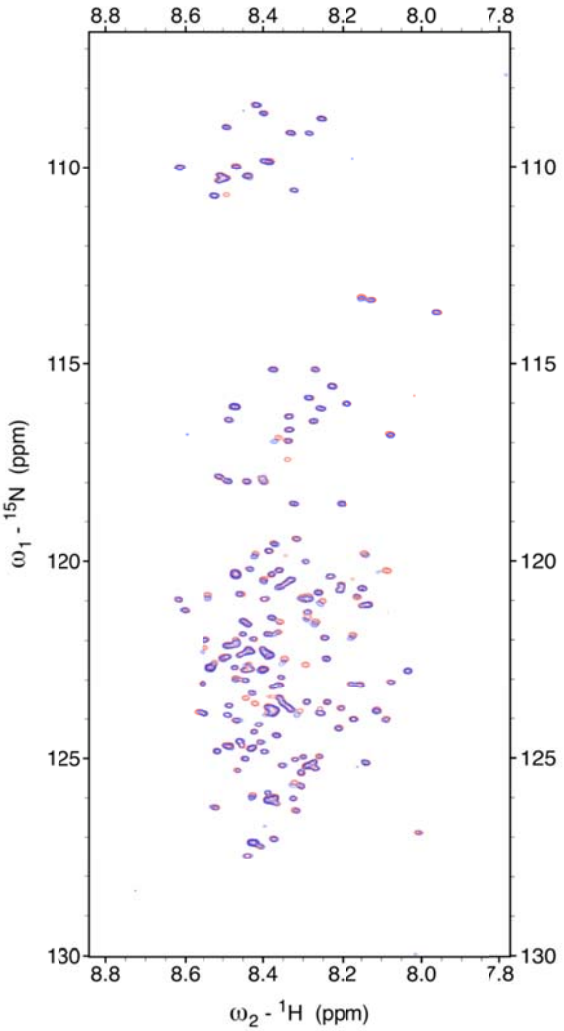


Figure S1. Overlay of the ^{1}H-^{15}N HSQC spectra of unmodified (red resonances) and N-myristoylated (blue resonances) BASP1. The spectra were acquired at 25 °C on Varian Inova spectrometers operating at 800 MHz.

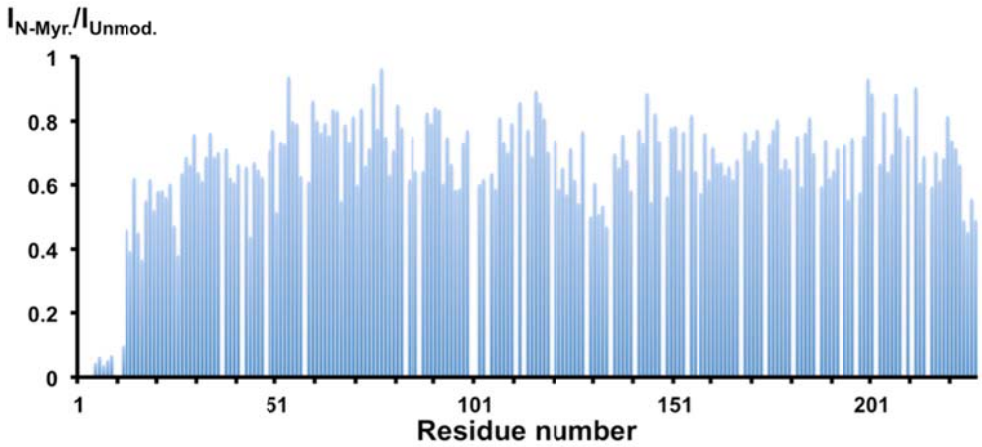


Figure S2. Normalised intensity ratio between the ^{1}H-^{15}N HSQC spectra of N-myristoylated and unmodified BASP1.

Collisionally Activated Dissociation and Electron Capture Dissociation Mass spectrometry

Electron capture dissociation (ECD, electron energy <1 eV) of the $(M + 30H)^{30+}$ ions of human BASP1 from overnight expression in *E. coli* M9 medium containing 50 µM myristic acid at *m/z* ~788.5 and ~789.4 (Figure S3) gave $z^{\bullet}$ fragment ions (Scheme S1) consistent with the predicted sequence of unmodified BASP1, which were used for internal calibration of the ECD spectrum (standard deviation 0.4 ppm). The measured protein mass values (most abundant isotope) derived from signals of the undissociated protein were 23651.59 and 23623.55 Da, which agree to within 0.40 and 0.03 ppm with the predicted mass of singly myristoylated and lauroylated BASP1, respectively.

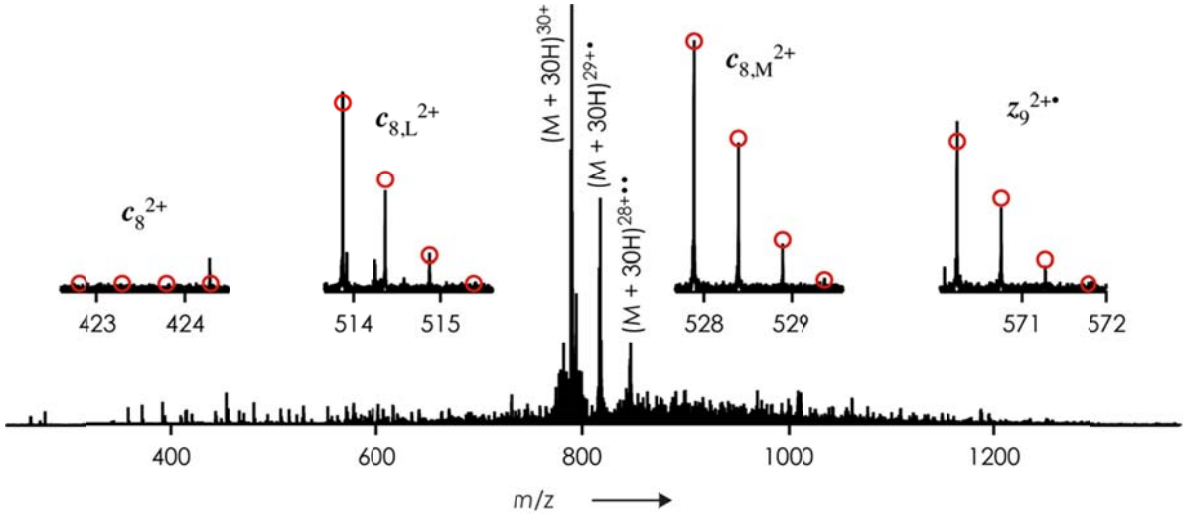
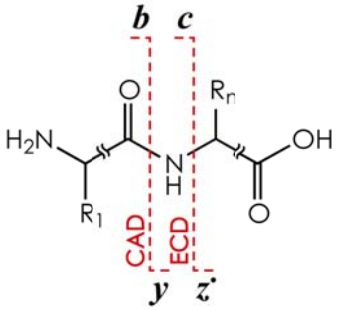


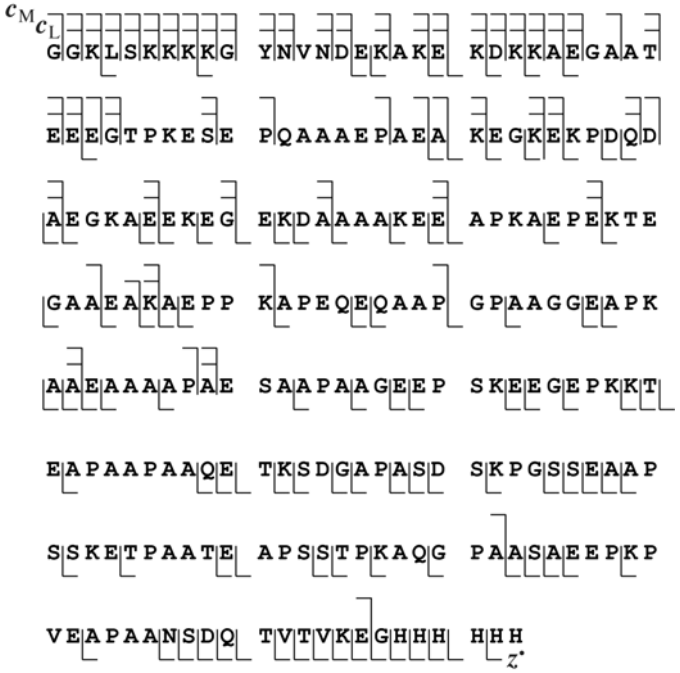
Figure S3. ECD spectrum of $(M + 30H)^{30+}$ ions at *m/z* ~788.5 and ~789.4 from ESI of human BASP1 from overnight expression in *E. coli* M9 medium containing 50 µM myristic acid; insets show measured (lines) and calculated (circles) isotopic profiles of c_8, $z_9^{\bullet}$, lauroylated c_8 (c_{8L}), and myristoylated c_8 ($c_{8,M}$) fragment ions.



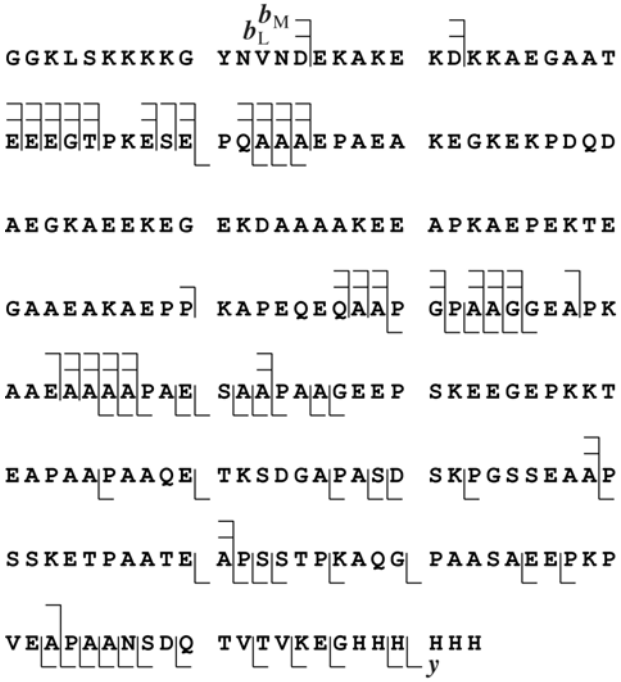
Scheme S1. Schematic illustration of protein backbone dissociation into complementary b/y or $c/z^{\bullet}$ fragment ion pairs by CAD or ECD, respectively.

Only unmodified $z^{\bullet}$ fragment ions (Figure S3) were observed in the ECD spectrum of human BASP1 from overnight expression in *E. coli* M9 medium containing 50 µM myristic acid (Scheme S2),

indicating myristoylation and lauroylation at the N terminus. Consistent with this, only modified ***c*** fragments were observed, both in myristoylated ($\boldsymbol{c}_M$) and lauroylated ($\boldsymbol{c}_L$) form (Scheme 2). Likewise, collisionally activated dissociation (CAD) of the $(M + 30H)^{30+}$ ions of human BASP from overnight expression in *E. coli* M9 (Figure 1b) gave only unmodified ***y*** ions, and only myristoylated ($\boldsymbol{b}_M$) and lauroylated ($\boldsymbol{b}_L$) ***b*** fragments (Scheme S3).



Scheme S2. Sequence map illustrating ***c*** and ***z*** $^{\bullet}$ fragment ion coverage from ECD of $(M + 30H)^{30+}$ ions at *m/z* ~788.5 and ~789.4 from ESI of human BASP1 from overnight expression in *E. coli* M9 medium containing 50 μM myristic acid; $\boldsymbol{c}_M$ and $\boldsymbol{c}_L$ indicate myristoylated and lauroylated ***c*** fragments, respectively.



Scheme S3. Sequence map illustrating ***b*** and ***y*** fragment ion coverage from CAD of $(M + 30H)^{30+}$ ions at *m/z* ~788.5 and ~789.4 from ESI of human BASP1 from overnight expression in *E. coli* M9 medium containing 50 μM myristic acid; $\boldsymbol{b}_M$ and $\boldsymbol{b}_L$ indicate myristoylated and lauroylated ***b*** fragments, respectively.

Separate ECD and CAD spectra of the $(M + 30H)^{30+}$ ions of human BASP1 from expression in *E. coli* in LB and M9 medium containing 50 μM lauric acid (*m/z* ~788.5, Figure 2e, f) showed that lauroylation also occurs at the N terminus of the protein.

2.2. Chapter 2

Flamm, A.G.; Żerko, S.; Zawadzka-Kazimierczuk, A.; Koźmiński, W.; Konrat, R.; Coudevylle, N., ^{1}H,^{15}N,^{13}C resonance assignment of human GAP-43. *Biomol. NMR Assign*., **2016,** 10, 171-174.

Biomol NMR Assign (2016) 10:171–174
DOI 10.1007/s12104-015-9660-9

ARTICLE

^{1}H, ^{15}N, ^{13}C resonance assignment of human GAP-43

Andrea Gabriele Flamm[1] · Szymon Żerko[2] · Anna Zawadzka-Kazimierczuk[2] · Wiktor Koźmiński[2] · Robert Konrat[1] · Nicolas Coudevylle[1]

Received: 17 August 2015 / Accepted: 28 December 2015 / Published online: 9 January 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract GAP-43 is a 25 kDa neuronal intrinsically disordered protein, highly abundant in the neuronal growth cone during development and regeneration. The exact molecular function(s) of GAP-43 remains unclear but it appears to be involved in growth cone guidance and actin cytoskeleton organization. Therefore, GAP-43 seems to play an important role in neurotransmitter vesicle fusion and recycling, long-term potentiation, spatial memory formation and learning. Here we report the nearly complete assignment of recombinant human GAP-43.

Keywords GAP-43 · Neuromodulin · Intrinsically disordered protein · Growth cone

Biological context

GAP-43 is a 25 kDa neuronal intrinsically disordered protein (IDP) highly abundant in the neuronal growth cone during development and regeneration. The exact molecular function(s) of GAP-43 remains unclear but it appears to be involved in growth cone guidance and actin cytoskeleton organization (Frey et al. 2000). Therefore, GAP-43 seems to play an important role in neurotransmitter vesicle fusion and recycling, long-term potentiation, spatial memory formation and learning (Denny 2006). GAP-43 experiences S-palmitoylation on positions Cys3 and Cys4 (Liu et al. 1993). Once acylated, it is bound to the inner leaflet of the plasma membrane in the growth cone of axons where it sequesters phosphatidylinositol-4,5-bisphosphate (PIP2) into lipid rafts via electrostatic interactions involving polybasic stretches (Laux et al. 2000). GAP-43 appears to be regulated by calcium- and PIP2-signaling cascades as it can bind calmodulin (CaM) and is the substrate of protein kinase C (PKC) (Apel et al. 1990; Maekawa et al. 1994). GAP-43 interacts with both apo and holo-CaM via its IQ domain (from His32 to Leu51), which adopts a helical conformation upon binding (Kumar et al. 2013). Phosphorylation by PKC occurs within the IQ domains at Ser41. Phosphorylation prevents the association with CaM and influences membrane binding (Maekawa et al. 1994; Tejero-Diez et al. 2000). Additionally, the phosphorylated form of GAP-43 seems to stabilize actin filaments via a direct interaction (He et al. 1997). Interestingly, GAP-43 is functionally related to the neuronal IDP BASP1, which is also under investigation in our group.

In order to gain insight into the intriguing membrane and ligand binding properties of GAP-43 and how post-translational modifications affect its structural dynamics and therefore its binding properties, we first achieved the near complete chemical shift assignment of human GAP-43.

Electronic supplementary material The online version of this article (doi:10.1007/s12104-015-9660-9) contains supplementary material, which is available to authorized users.

✉ Nicolas Coudevylle
nicolas.coudevylle@univie.ac.at

[1] Department of Computational and Structural Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria

[2] Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland

Methods and results

Protein expression and purification

The coding region for human GAP-43 was obtained by assembly PCR (Stemmer et al. 1995), this procedure allows to directly optimize the coding sequence for the chosen expression system and to avoid undesired restriction sites inside the coding sequence. 32 oligonucleotides, covering the entire coding sequence of GAP-43, were designed using the online software DNAworks (Hoover and Lubkowski 2002). The PCR gene assembly was realized according to a published protocol (Stemmer et al. 1995). The final gene amplification was performed from 1 µl of the gene assembly mixture by introducing 5′ NdeI and 3′ NotI restriction sites. Subsequently, the fragment was inserted in-frame into the NdeI and NotI sites of the bacterial expression vector pET-41b, yielding pET-41b-hGAP-43, encoding human GAP-43 fused to a C-terminal His6-tag. The two cysteines on position 3 and 4 were replaced by glycines by site-directed mutagenesis. $^{15}N/^{13}C$ labeled GAP-43 was expressed in the *Escherichia coli* strain T7 express (New England BioLab) in minimal medium containing ^{15}N-labeled ammonium chloride and ^{13}C-glucose as sole nitrogen and carbon sources, respectively. GAP-43 expression was induced at an A_{600nm} of 0.6 by addition of 0.8 mM IPTG. The cells were collected after 16 h of expression at 30 °C by centrifugation at 6000 rpm for 10 min and resuspended in 30 ml of ice-cold lysis buffer (20 mM NaP_i pH 7.4, 50 mM NaCl) per liter of the original bacterial culture. Bacteria were lysed by passing through a French press, and the cell lysate was cleared by centrifugation at 18,000 rpm for 20 min. The supernatant containing the soluble protein fraction was loaded onto a Ni^{2+} loaded HiTrap 5 ml affinity column (GE Healthcare), washed with 2 column volumes of high salt buffer (20 mM NaP_i pH 7.4, 1.5 M NaCl, 10 mM imidazole) and eluted with high imidazole buffer (20 mM NaP_i pH 7.4, 50 mM NaCl, 0.5 M imidazole) using a linear gradient of 15 column volumes. The GAP-43 containing fractions were collected and the buffer was exchanged by dialysis in the measurement buffer, 20 mM NaPi pH 7.4, 50 mM NaCl and subsequently concentrated to 0.5 mM in an Amicon Ultra-15 centrifugal filter device 10 K NMWL (Amicon).

NMR experiments

The backbone ^{1}H, ^{13}C and ^{15}N resonances were assigned using a set of high dimensionality experiments exploiting sparse random sampling of indirectly detected time domains, in order to increase resolution. A 3D HNCO and 3D CACON experiments were used as a base spectra for sparse multidimensional Fourier transform (SMFT) processing of higher dimensionality experiments (Kazimierczuk et al. 2009). Backbone assignment was achieved using 4D HN(CA)NH (Zawadzka-Kazimierczuk et al. 2010) 5D HN(CA)CONH (Kazimierczuk et al. 2010), (HACA)CON(CA)CONH (Zawadzka-Kazimierczuk et al. 2012b), (H)NCO(NCA)CONH (Zawadzka-Kazimierczuk et al. 2012b) experiments. Side-chain assignments were obtained using 5D HabCabCONH, 5D HNCOCACB (Zawadzka-Kazimierczuk et al. 2012b) and HC(CC-tocsy)CONH experiments (Kazimierczuk et al. 2009; Hiller et al. 2008). All those5D experiments were acquired at 298 K on an Agilent Direct Drive 700 MHz spectrometer using the standard 5 mm $^{1}H–^{13}C–^{15}N$ triple-resonance probe head. Additionally, to obtain proline residues assignment two directly ^{13}C detected experiments: 3D CACON and 5D (HACA)CONCACON (Bermel et al. 2013) were performed using Agilent Direct Drive 2 800 MHz spectrometer equipped with cryogenically cooled 5 mm $^{1}H–^{13}C–^{15}N$ triple-resonance probe head.

All NMR data sets were processed by multidimensional Fourier transformation using the home written software package (Kazimierczuk et al. 2006, 2009; Stanek and Kozminski 2010; Stanek et al. 2012) (http://nmr.cent3.uw.edu.pl/software). Sampling artefacts from 3D HNCO and 4D HN(CA)NH were removed using cleaner3d (Stanek and Kozminski 2010) and cleaner4d (Stanek et al. 2012) programs, respectively. The resonance assignment was performed using the TSAR program (Zawadzka-Kazimierczuk et al. 2012a). The input data for TSAR was prepared using Sparky software.

Extent of assignment and data deposition

As can be clearly seen from the very narrow peak dispersion in the ^{1}H dimension of the $^{1}H–^{15}N$ HSQC spectrum (Fig. 1), GAP-43 is an intrinsically disordered protein. Extensive signal overlap in conventional 2D and 3D spectra could be overcome by using the aforementioned 5D experiments. Interestingly, for the full-length protein, the IQ domain (from His32 to Leu51) could not be assigned, probably due to conformational exchange originating from the pronounced helical propensity of this segment. However, the $^{1}H–^{15}N$ HSQC spectrum of the IQ domain could be assigned in the context of a shorter construct consisting of the first 59 residues of GAP-43 (GAP-43-NTD) and the assignment transposed to the spectrum of the full-length protein (see supplementary material for full details). However, the side chain assignment of the IQ domain for the full-length protein could not be obtained. Consequently, 98 % of backbone ^{15}N, 98 % of ^{1}HN resonance

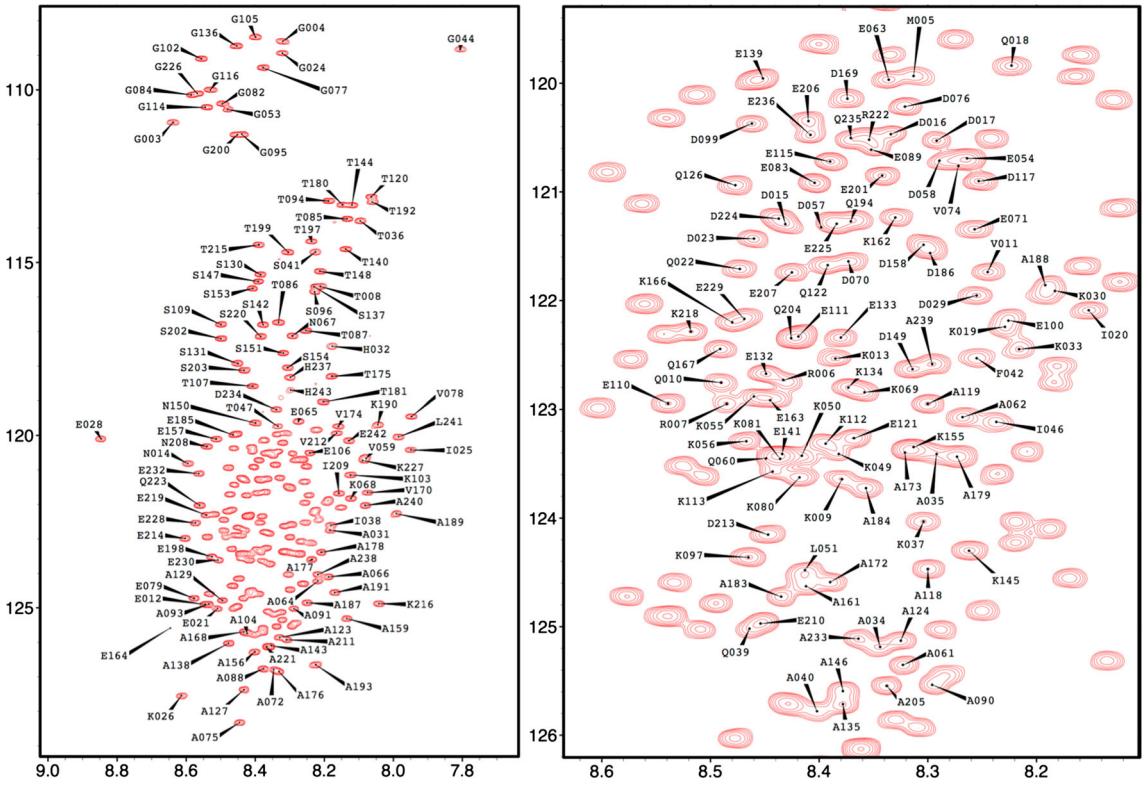


Fig. 1 *Left* ^{1}H–^{15}N HSQC spectrum of human GAP-43 at pH 6 and 298 K. *Right* magnification of the central region of the spectrum

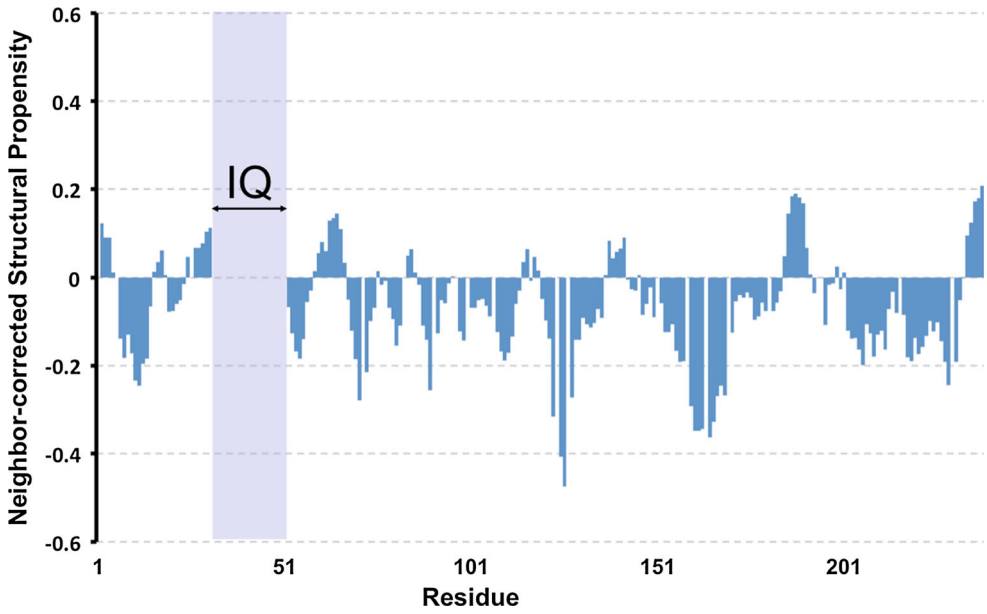


Fig. 2 Neighbor-corrected structural propensity index (Tamiola and Mulder 2012) of GAP-43 at pH 6 and 298 K. The IQ domain (from His32 to Leu51), for which only the backbone ^{15}N and ^{1}HN resonance assignment is available, is highlighted in *blue*

have been assigned and, excluding the IQ domain, 89 % of ^{13}Cα, 85 % of ^{1}Hα, 88 % of ^{13}Cβ, 77 % of ^{1}Hβ and 97 % of ^{13}C′ resonances have been assigned. Additionally, HC(CC-tocsy)CONH spectra allowed the assignment of several side-chain atoms. The neighbor-corrected structural propensity index (Tamiola and Mulder 2012) clearly shows that GAP-43 is devoid of any stable secondary structure element (Fig. 2) but seems to preferentially adopt extended structures (negative index values).

The ^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under the BMRB accession number 19246.

Acknowledgments All NMR experiments were carried out in the Structural Research Laboratory at the Faculty of Chemistry, University of Warsaw. This work was supported by the Grant P24761 from the Austrian Science Foundation FWF. S.Ż and W.K. thank the Foundation for Polish Science for support with the TEAM Programme. Co-financed by the EU European Regional Development.

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

References

Apel ED, Byford MF, Au D, Walsh KA, Storm DR (1990) Identification of the protein kinase C phosphorylation site in neuromodulin. Biochemistry 29(9):2330–2335

Bermel W, Felli IC, Gonelli L, Koźmiński W, Piai A, Pierratelli R, Zawadzka-Kazimierczuk A (2013) High-dimensionality ^{13}C direct-detected NMR experiments for the automatic assignment of intrinsically disordered proteins. J Biomol NMR 57(4):353–361

Denny JB (2006) Molecular mechanisms, biological actions, and neuropharmacology of the growth-associated protein GAP-43. Curr Neuropharmacol 4(4):293–304

Frey D, Laux T, Xu L, Schneider C, Caroni P (2000) Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity. J Cell Biol 149(7):1443–1454

He Q, Dent EW, Meiri KF (1997) Modulation of actin filament behavior by GAP-43 (neuromodulin) is dependent on the phosphorylation status of serine 41, the protein kinase C site. J Neurosci Off J Soc Neurosci 17(10):3515–3524

Hiller S, Joss R, Wider G (2008) Automated NMR assignment of protein side chain resonances using automated projection spectroscopy (APSY) experiments. JACS 130(36):12073–12079

Hoover DM, Lubkowski J (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res 30(10):e43

Kazimierczuk K, Zawadzka A, Kozminski W, Zhukov I (2006) Random sampling of evolution time space and Fourier transform processing. J Biomol NMR 36(3):157–168

Kazimierczuk K, Zawadzka A, Koźmiński W (2009) Narrow peaks and high dimensionalities: exploiting the advantages of random sampling. J Magn Reson 197(2):219–228

Kazimierczuk K, Zawadzka-Kazimierczuk A, Koźmiński W (2010) Non-uniform frequency domain for optimal exploitation of non-uniform sampling. J Magn Reson 205(2):286–292

Kumar V, Chichili VP, Zhong L, Tang X, Velazquez-Campoy A, Sheu FS, Seetharaman J, Gerges NZ, Sivaraman J (2013) Structural basis for the interaction of unstructured neuron specific substrates neuromodulin and neurogranin with Calmodulin. Sci Rep 3:1392

Laux T, Fukami K, Thelen M, Golub T, Frey D, Caroni P (2000) GAP43, MARCKS, and CAP23 modulate PI(4,5)P(2) at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. J Cell Biol 149(7):1455–1472

Liu Y, Fisher DA, Storm DR (1993) Analysis of the palmitoylation and membrane targeting domain of neuromodulin (GAP-43) by site-specific mutagenesis. Biochemistry 32(40):10714–10719

Maekawa S, Murofushi H, Nakamura S (1994) Inhibitory effect of calmodulin on phosphorylation of NAP-22 with protein kinase C. J Biol Chem 269(30):19462–19465

Stanek J, Kozminski W (2010) Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets. J Biomol NMR 47(1):65–77

Stanek J, Augustyniak R, Kozminski W (2012) Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm. J Magn Reson 214(1):91–102

Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 164(1):49–53

Tamiola K, Mulder FA (2012) Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. Biochem Soc Trans 40(5):1014–1020

Tejero-Diez P, Rodriguez-Sanchez P, Martin-Cofreces NB, Diez-Guerra FJ (2000) bFGF stimulates GAP-43 phosphorylation at ser41 and modifies its intracellular localization in cultured hippocampal neurons. Mol Cell Neurosci 16(6):766–780

Zawadzka-Kazimierczuk A, Kazimierczuk K, Kozminski W (2010) A set of 4D NMR experiments of enhanced resolution for easy resonance assignment in proteins. J Magn Reson 202(1):109–116

Zawadzka-Kazimierczuk A, Koźmiński W, Billeter M (2012a) TSAR: a program for automatic resonance assignment using 2D cross-sections of high dimensionality, high-resolution spectra. J Biomol NMR 54(1):81–95

Zawadzka-Kazimierczuk A, Koźmiński W, Sanderova H, Krasny L (2012b) High dimensional and high resolution pulse sequences for backbone resonance assignment of intrinsically disordered proteins. J Biomol NMR 52(4):329–337

Supplementary data to:

^{1}H, ^{15}N, ^{13}C resonance assignment of human GAP-43

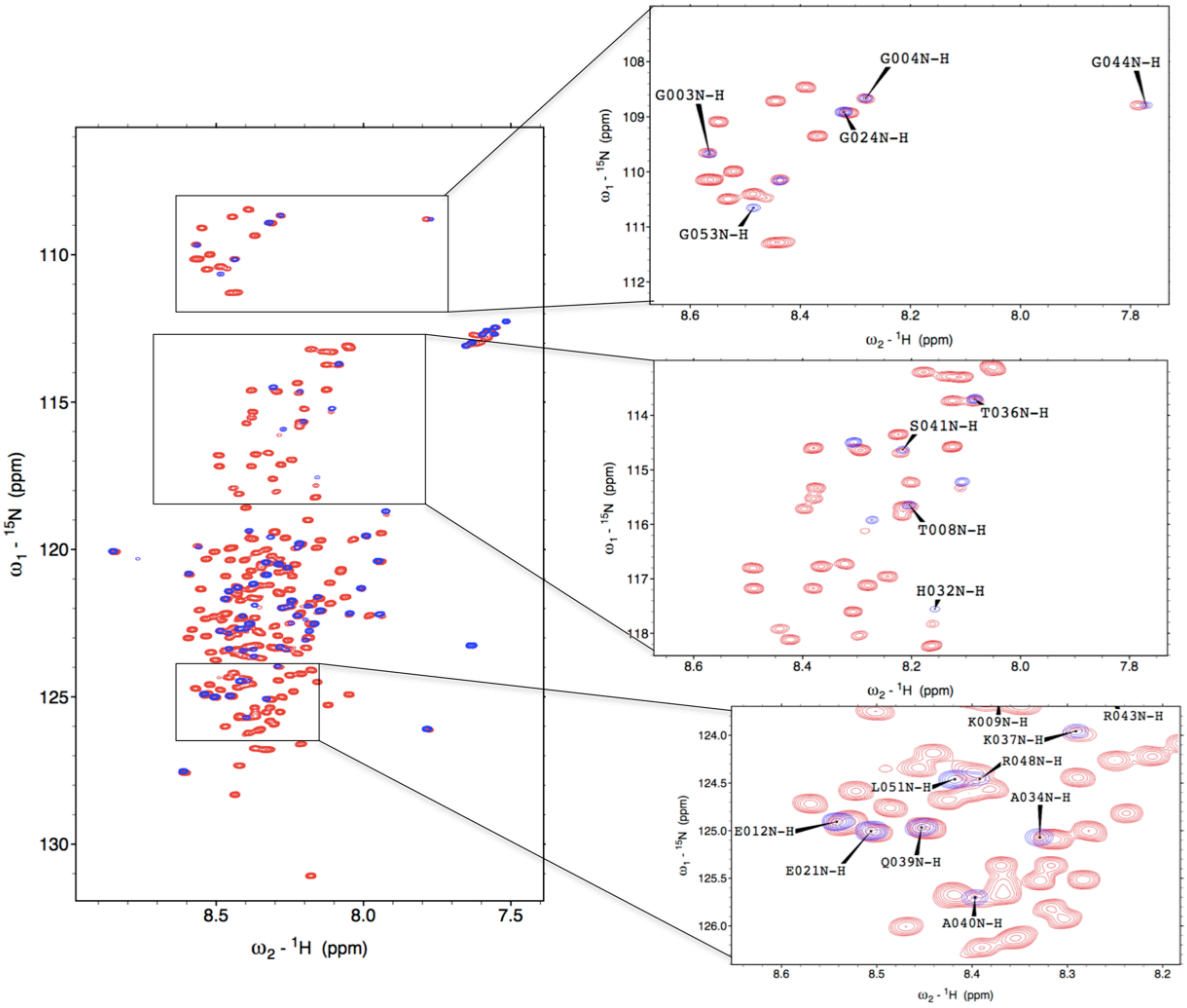
Andrea Gabriele Flamm[1], Szymon Żerko[2], Anna Zawadzka-Kazimierczuk[2], Wiktor Koźmiński[2], Robert Konrat[1] and Nicolas Coudevylle[1*]

Cloning and purification GAP-43-NTD

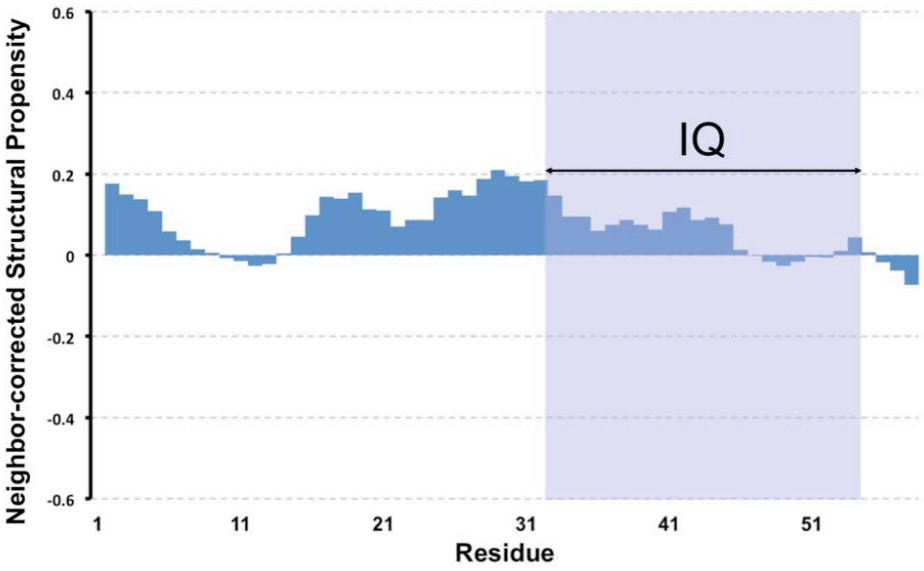
The GAP-43 N-terminal domain (amino acids 1-59 of the full length protein) was expressed using the vector pETM11 with a N-terminal His6-tag. Expression was done in *E. col*i strain T7 in minimal medium with ^{15}N Ammonium chloride and $^{13}C_6$ D- glucose as the sole nitrogen and carbon source. The expression is induced at an OD_{600} of 0.8 by adding 0.8 mM IPTG. Expression was carried out overnight (~12h) at 28°C. The cell pellet (after centrifugation at 5000 rpm for 15 min) was resuspended using PBS. Breaking the cells is done by sonication (3 min of 50% amplitude) and the supernatant after centrifugation at 18000 rpm for 20 min was pressed through a 0.45 μm filter before loading it on a Ni2+ -loaded HiTrap 5mL affinity column (GE healthcare). After a washing step with PBS containing 1.5M NaCl, the protein was eluted using 100% HI-PBS (~1-2 column volumes) The tag was cleaved with TEV-protease and the sample dialyzed against the measurement buffer

Assignment of GAP-43-NTD

Experiments with ^{15}N-^{13}C labeled GAP-43-NTD were performed at 298K at a 600Mhz Bruker spectrometer. The NTD construct overlaps nicely with the assignment of the full length protein (Supp. Figure 1), therefore the remaining amino acids (32-53) including the IQ-domain that were not assigned in the full length construct could be identified and assigned using (H)N(CA)NNH-, HNCO- and HNCACB- type of experiments. The neighbor-corrected structural propensity index for GAP-43-NTD clearly shows that in this construct, the IQ domain is devoid of any stable or even transient secondary structure element (supp. Figure 2).



Supp. Figure 1. ^{1}H-^{15}N HSQC spectra of full length GAP-43 (red resonances) and GAP-43-NTD (blue resonances).



Supp. Figure 2. Neighbor-corrected structural propensity index of GAP-43-NTD at pH 6 and 298 K. The IQ domain (from His32 to Leu51) is highlighted in blue.

2.3. Chapter 3

Flamm, A.G.; Konrat, R.; Coudevylle, N.; The membrane binding properties of two neuronal intrinsically disordered proteins: GAP-43 and BASP1. *Structure,* **2016***,* manuscript number: STRUCTURE-D-16-00375, submitted (24. Oct. 2016).

The membrane binding properties of two neuronal intrinsically disordered proteins: BASP1 and GAP-43

Andrea G. Flamm[1], Robert Konrat[1] and Nicolas Coudevylle[12*]

[1]Department of Computational and Structural Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria

[2]Lead Contact

*Correspondence: nicolas.coudevylle@univie.ac.at

SUMMARY

Peripheral membrane proteins usually bind to membranes using dedicated folded protein domains, such as C1, C2, or PH domains. However, it is still unclear how intrinsically disordered peripheral membrane proteins target specific parts of the membrane (such as phosphatidylinositol-4,5-bisphosphate) without a folded domain showing shape complementary with the target. Here we propose the use of small bicelles as membrane systems to study peripheral membrane proteins, as their small size and versatility make them ideally suitable for liquid-state NMR. Using small bicelles, NMR and liposome sedimentation assays we show that, for BASP1 and GAP-43, the density of positive charges along the protein backbone drives the interaction with negatively charged membranes. The specificity of these intrinsically disordered proteins for their target is mainly defined by electrostatic attractions rather than conformational complementarity. We also show that, in contrast to prototypical membrane binding proteins, Ca^{2+} prevents the binding of BASP1 and GAP-43 to membranes.

Keywords: Intrinsically disordered proteins, peripheral membrane proteins, small bicelle, liquid-state NMR, calcium binding protein

INTRODUCTION

It is well understood how peripheral membrane proteins interact with biological membranes. Hydrophobic anchor structures (hydrophobic helix, post translationally attached acyl group or GPI anchor) drive the association of the protein with the plasma membrane through unspecific hydrophobic interactions. More specific interactions (e.g. in order to target a defined part of the membrane) are driven by dedicated protein domains (such as PH, C2 or C1 domains) that recognize specific parts of the cellular membrane by harbouring shape and charge complementarity with the target lipid head-groups (Montaville et al., 2008). These domains also allow some level of regulation by effectors and second messengers. For example, calcium (Ca^{2+}) can facilitate the interaction between a protein and the membrane by stabilizing ionic bridges between acidic side chains and negative head-groups and/or through allosteric modulation of the protein's affinity for its target head-group (Coudevylle et al., 2008; Montaville et al., 2002). These exquisite membrane-binding properties are due to the stable three-dimensional organization of the protein domain that brings residues together in space with various properties of charge and shape.

However, it is now well recognized that some physiologically important peripheral membrane proteins are intrinsically disordered proteins (IDPs). Two examples are GAP-43 (Growth associated protein, also known as neuromodulin, B-50, P-57, F1 and pp46) and BASP1 (Brain-acid soluble protein 1, also known as NAP-22 and CAP-23) that are both 25 kDa IDPs highly abundant in the brain during development. Their exact functions remain unclear but they both appear to be involved in growth cone guidance and actin cytoskeleton organization (Frey et al., 2000). They seem to have similar functions *in vivo*, as knockout mice of BASP1 fail in stimulus-induced nerve sprouting at the neuromuscular junction but transgenic overexpression of BASP1 or GAP-43 re-induced nerve sprouting. Knock-in mice expressing GAP-43 under BASP1 promoter rescued the phenotype, but were sterile (Frey et al., 2000).

Both proteins are acylated; BASP1 undergoes N-myristoylation (Zakharov et al., 2003) whereas GAP-43 experiences S-palmitoylation on positions Cys3 and Cys4 (Liu et al., 1993). Once acylated, both proteins are bound to the inner leaflet of the plasma membrane in the growth cone of axons where they co-localize with phosphatidylinositol-4,5-bisphosphate (PIP2) *via* electrostatic interactions involving polybasic stretches of the proteins (Laux et al., 2000). Both proteins seem to form oligomers in presence of anionic phospholipids, which has been proposed to be playing a role in PIP2 signalling pathway (Forsova and Zakharov, 2016; Zakharov and Mosevitsky, 2010). They both appear to be regulated by calcium- and PIP2-

signaling cascades as they can bind calmodulin (CaM) and are substrates of PKC (Apel et al., 1990; Maekawa et al., 1994). GAP-43 interacts with both apo and holo-CaM *via* its IQ domain (from His32 to Leu51), which adopts a helical conformation upon binding (Kumar et al., 2013). BASP1 only interacts with holo-CaM in its myristoylated form by insertion of the myristoyl moiety into a hydrophobic pocket of calmodulin and the interaction of a patch of basic residues on the N-terminus of BASP1 (Matsubara et al., 2004). Phosphorylation by PKC occurs within the effector domains of GAP-43 at Ser41, and at Ser6 for BASP1 and prevents the association with CaM and influences membrane binding (Maekawa et al., 1994; Tejero-Diez et al., 2000).

GAP-43 and BASP1 appear functionally related although they share no significant sequence identity. Nevertheless, their amino acid composition is similar, both having a high percentage of acidic amino acids, lysine and alanine resulting in an acidic pI and high hydrophilicity. More importantly, they both are intrinsically disordered proteins (Flamm et al., 2016b; Geist et al., 2013) and it remains unclear how they can recognize and interact with specific parts of the membrane (e.g. PIP2) without a stable fold. Both GAP-43 and BASP1 were successfully expressed, purified and assigned in the laboratory (Flamm et al., 2016b; Geist et al., 2013). We also mastered the recombinant expression of myristoylated BASP1 (Flamm et al., 2016a). Therefore we decided to delineate the membrane binding properties of both proteins, especially the factors that allow them to specifically target lipid head groups, using liquid state NMR in conjunction with other biophysical methods.

RESULTS

The unacylated forms of GAP-43 and BASP1 bind to neuronal membranes and are released in the presence of calcium.

In the cell, GAP-43 and BASP1 are both found acylated. BASP1 is N-myristoylated (Frey et al., 2000), whereas GAP-43 is palmitoylated on positions 3 and 4 (Zakharov et al., 2003). These long-chain acylations drive the association of the protein with the inner leaflet of the plasma membrane through unspecific hydrophobic interactions. However, the polypeptidic part of the conjugated protein is thought to target specific areas of the membrane in the cell (e.g. co- localization with PIP2), mainly through electrostatic interactions between basic residues of the protein and negative head-groups from the membrane (Laux et al., 2000). In order to delineate the structural factors involved in the recognition of specific patches of the plasma membrane by the polypeptide, we check whether recombinant (un-acylated) BASP1 and GAP-43 bind to biological membranes *in vitro*.

We first used a sedimentation assay to monitor the binding of both proteins to liposomes. Folch extract from bovine brain was used to prepare “Folch liposomes” that recapitulate the phospholipid composition of neuronal membranes (Folch et al., 1957). As can be seen on Figure 1A, N-myristoylated BASP1 (Myr-BASP1) but also unacylated GAP-43 and BASP1 all bind to Folch liposomes, revealing that the peptidic parts of GAP-43 and BASP1 alone are sufficient to bind membranes. Since calcium is known to modulate the interaction with membranes of many proteins (Arac et al., 2006; Moreno et al., 2010; Mosior and Newton, 1998; Paddock et al., 2008), we examined the influence of Ca^{2+} on the liposome binding properties of GAP-43 and BASP1. Unexpectedly, in the presence of 2 mM Ca^{2+} un-acylated GAP-43 and BASP1, as well as Myr-BASP1, are no longer able to bind to liposomes. In order to conveniently and accurately monitor the binding of BASP1 and GAP-43 to liposomes, fluorescent versions of GAP-43 and BASP1 were designed by fusing the fluorescent tag iLOV (Christie et al., 2012; Gawthorne et al., 2012) to the C-terminal extremity of the protein (as it does not seem to be involved in membrane binding). This permits to measure with improved speed and accuracy (compared to SDS-PAGE based quantification) the amount of protein in both the liposome (pellet) and soluble (supernatant) fractions. As can be seen on Figure 1B, using iLOV fusion proteins allows to measure accurately the amount of protein bound to liposome (64% for BASP1 and 83% for GAP-43) and to show that addition of 2 mM Ca^{2+} almost completely abolishes the binding to Folch liposome for both proteins.

In order to gain site-specific information on the membrane binding mode of BASP1 and GAP-43, we performed ^{1}H-^{15}N HSQC monitored titrations of ^{15}N labeled proteins by Folch liposomes. At low liposome concentration (around 2 $mg.mL^{-1}$ of lipid), no chemical shift change or significant line broadening is observed, however the intensity of many crosspeaks is greatly reduced (Figure 1C-D). This indicates that the lipid bound form of both proteins is effectively NMR-invisible. However, due to the strong structural dynamic heterogeneity of these proteins, the non-uniform decrease of intensity along the protein backbone indicates the location of the phospholipid-binding sites (i.e. the immobilized residues in the liposome bound state). Both proteins clearly interact through basic patches, suggesting that they bind to the membrane *via* electrostatic interactions with negatively charged lipid head-groups (such as PIP2 or PS). However they seem to adopt slightly different binding modes, GAP-43 clearly interacts through its N-terminal extremity (up to residue 10) together with the N-terminal half of the IQ domain (residues 25 to 40). After residue 40 the signal intensity quickly reached

back its unbound values (intensity ratio of 1), revealing that the rest of the protein remains disordered in its liposome-bound state.

BASP1 seems to adopt a more complex binding mode. The very N-terminal is clearly involved (up to residue 10), but the intensity along the rest of the protein is affected in distinct steps. A first segment consists of residues 10 to 50 (intensity ratio of 0.5), followed by a second segment from residue 50 to 100 (intensity ratio of 0.8). This observation is reminiscent of the binding mode of α-synuclein towards small unilamellar vesicles and could be due to coexistence of multiple binding modes (Bodner et al., 2009).

GAP-43 and BASP1 target negative lipid head-groups.

The NMR based titrations by Folch liposomes clearly suggest that the binding of unacylated BASP1 and GAP-43 to membranes is driven by the interaction between basic patches of the proteins and negatively charged lipids of the plasma membrane. In order to determine which lipid head-groups and moieties the proteins are specifically targeting, we performed liposome sedimentation assay with whole PC (neutral) liposomes to which were added increasing amounts of different lipids. As can be seen on Figure 2A-B, both GAP-43 and BASP1 only marginally bind to plain PC liposomes. Increasing the fraction of PE (neutral) or cholesterol in the liposome does not increase protein binding. However, addition of PS (negatively charged) increases the amount of liposome bound protein, which confirms that un-acylated GAP-43 and BASP1 bind to membranes through the interaction between their basic residues and negatively charged lipid head-groups.

A more detailed analysis of the effect of PS on BASP1 and GAP-43 binding to liposomes reveals that protein binding displays a sigmoidal dependence on the mole fraction of PS in PC liposomes (Figure 2C-D). This apparent sigmoidal dependence would imply that BASP1 and GAP-43 both bind to several PS molecules in a cooperative way. However, this apparent cooperativity is a known behaviour for proteins binding to multiple ligands embedded in a lipid membrane. It has first been described for Protein Kinase C's activation by PS (Mosior and Newton, 1998) and mainly originates from the reduction of spatial dimensionality occurring while the protein binds to the membrane (first binding event) and the associated increase in local ligand (PS) concentration that facilitate subsequent binding events (second, possibly third etc.). Consequently, the sigmoidal dependence on the mole fraction of PS does not necessarily imply an allosteric behaviour. However, it clearly shows that both BASP1 and GAP-43 bind multiple PS molecules. Using Hill's equation to fit the experimental data yields Hill coefficients of 3.8 and 5.4 for BASP1 and GAP-43 respectively. Although these values cannot be taken literally, they give an estimate of the number of PS

molecules bound by each protein and reveal that GAP-43 binds more PS molecules than BASP1.

Both BASP1 and GAP-43 have been reported to sequester PIP2 (Laux et al., 2000), therefore we investigated the influence of PIP2 on the binding of BASP1 and GAP-43 to liposomes. Interestingly, when added to plain PC liposomes, PIP2 doesn't seem to affect the amount of membrane-bound protein (Figure 3A-B). However, in the presence of 50% of PS, PIP2 seems to significantly increase protein binding for both BASP1 and GAP-43 in a PIP2 concentration dependent manner (Figure 3C). Altogether these observations suggest that, at physiological concentration, PIP2 alone is not sufficient to recruit neither BASP1 nor GAP-43 to the membrane but will strengthen the interaction between the membrane and the bound protein. Interestingly, neither BASP1 nor GAP-43 shows any significant affinity for soluble IP3, the head-group of PIP2 (data not show).

Characterization of the membrane bound state of GAP-43 and BASP1.

NMR is uniquely suited to examine molecular interactions with an atomic resolution. However, the size of the molecular assembly formed by the IDP (BASP1 or GAP-43) bound to a liposome is considerably too large to be tackled by high-resolution liquid-state NMR. As exemplified on Figure 1B-C, when bound to a liposome the protein becomes effectively NMR-invisible. In order to be able to characterize the membrane bound state of GAP-43 and BASP1 by high-resolution liquid-state NMR, we dramatically reduced the size of the molecular system by 1) using shorter protein constructs; 2) using a considerably smaller membrane system (small bicelle).

We first designed short constructs of BASP1 and GAP-43 that mainly consist of the membrane binding motifs, specifically the first 57 (BASP1 NTD) and 59 (GAP-43 NTD) residues of the proteins. These shorter constructs have the advantages to lead to simpler spectra and to avoid line-width heterogeneities occurring while the full-length protein is bound to the membrane (broad lines for the membrane-bound residues against narrow lines of the free moving segments). We have already showed that the NTD of GAP-43 is unaffected by the absence of the rest of the protein (Flamm et al., 2016b), ^{1}H-^{15}N HSQC comparisons of the short construct with its full length forms reveal that this is also the case for BASP1 (Suppl. Fig. 2). Additionally, liposome sedimentation shows that both NTDs retain their membrane binding activities (Suppl. Fig. 3).

Next, we used small bicelles, instead of liposomes, as a membrane model. Small bicelles are commonly used in biomolecular NMR as alignment medium, they also have been used to study embedded membrane proteins (Durr et al., 2012). Here we attend to use them as

membrane model to study their interaction with peripheral membrane proteins (such as GAP-43 and BASP1). Small bicelles have the advantages to be considerably smaller than liposomes (around 4 nm for a small bicelle against approximately 400 nm for the type of liposome used in that study) and to harbour a flat surface, leading to more native-like protein/membrane interactions. Since GAP-43 and BASP1 appear to bind to negatively charged head-groups, we used a mixture of DMPS/DHPC with a q-factor of 0.25, ensuring that the proteins will bind to the flat surface of the small bicelles (where DMPS is found).

As can be seen on Figure 4, addition of small bicelles to ^{15}N labelled NTDs leads to large chemical shift changes for both BASP1 and GAP-43 NTDs, showing that both short constructs bind to bicelles and that the system formed by the NTD bound to a small bicelle can be tackled by liquid-state NMR as the protein resonances remain visible in the bound state. ^{1}H-^{15}N HSQC based titrations of BASP1 and GAP-43 NTDs by small bicelles (Figure 5-AB) show that both NTDs can be saturated. For BASP1-NTD, saturation is quickly reached at a total lipid concentration (DHPC+DMPS) of about 20 mM. A peculiar observation is made for GAP-43-NTD, at low lipid concentration the protein seems to aggregate (resonances are disappearing and the sample becomes cloudy), until the concentration reaches the CMC of DHPC, about 15 mM (Warschawski et al., 2011), where resonances re-appear and begin to shift upon subsequent addition of lipid. This peculiar behaviour could be explained by the fact that, at sub CMC lipid concentration, GAP-43-NTD will bind to free lipids (which seems to lead to aggregation). However, once the CMC is reached GAP-43 NTD will preferentially bind to small bicelles, no longer aggregate and display a hyperbolic binding curve (even though it could not be saturated).

Comparing the ^{1}H-^{15}N chemical shifts of both NTDs in absence and in presence of small bicelles (Figure 5-CD) shows that BASP1-NTD and GAP-43-NTD display different binding modes towards small bicelles. Only the very N-terminal part of BASP1 seems to be affected by binding (up to residue 11) suggesting that the N-terminal basic patch is the main interacting segment. Whereas GAP-43-NTD experiences moderate chemical shift changes at its N-terminal but large ones for the IQ domain, suggesting that the interaction interface is mainly composed of the IQ domain with a contribution of the N-terminal residues. These observations are redundant to those made for the full-length proteins bound to liposomes (Figure 1-BC), and are consistent with small bicelle titration (Figure 5-AB). Indeed, if the interaction interface for BASP1 is significantly smaller than for GAP-43, a bicelle can accommodate more BASP1 molecules, which explains that BASP1 reaches saturation at lower lipid concentration than GAP-43.

Finally, using ^{13}C, ^{15}N labelled samples allowed us to collect the backbone and side chain chemical shifts in the bicelle bound state. By comparing the secondary backbone and sidechain chemical shift changes upon small bicelle binding, we can calculate the secondary structure changes upon binding and characterize the membrane bound state. As can be seen on Figure 5-C, BASP1-NTD doesn't seem to undergo any change of its secondary structure change, which is consistent with its limited interaction surface. In contrast, the IQ domain of GAP-43 appears to adopt a more helical conformation upon membrane binding, highlighting the importance of this domain to mediate the interaction between GAP-43 and biological membranes.

Calcium releases GAP-43 and BASP1 form the membrane.

Many peripheral membrane proteins see their membrane binding properties modulated by Ca^{2+} (Arac et al., 2006; Moreno et al., 2010; Mosior and Newton, 1998; Paddock et al., 2008). As can be seen on Figure 1-AB, both GAP-43 and BASP1 lose their ability to bind membrane in presence of 2 mM Ca^{2+}. This is a surprising observation as Ca^{2+} is usually promoting protein/membrane interactions by mediating electrostatic interactions between the negatively charged lipid head-group and acidic residues. In order to delineate the mechanism by which Ca^{2+} triggers membrane release, we first examined the Ca^{2+} binding properties of the proteins in solution. ^{1}H-^{15}N HSQC based Ca^{2+} titrations reveal that both proteins indeed seem to bind Ca^{2+} (Suppl. Fig 4). However, the chemical shifts generated by Ca^{2+} binding are of relatively small amplitude and both proteins seem to harbour several binding sites with different affinities. As a consequence, many residues will see their chemical environment affected by several binding sites with different affinity leading to complex titration curves (Suppl. Fig 5). In order to identify the high affinity binding sites, we used Mn^{2+}, a paramagnetic Ca^{2+} mimetic. Due to its paramagnetic properties, Mn^{2+} will lead to intensity decrease at low concentration where chemical shift changes are not observable yet, revealing the high affinity binding sites. As can be seen on Figure 6-AB, addition of 30 µM Mn^{2+} leads to significant intensity losses for both BASP1 and GAP-43. Unsurprisingly Mn^{2+} seems to bind at acidic patches present in both proteins. However it cannot be determined whether several Ca^{2+} ions are binding or a single ion that recruit distance parts of the protein.

Next we performed Ca^{2+} titrations of the liposome release (Figure 6-CD). In order to maximize the amount of bound protein in absence of calcium we used highly negatively charged liposomes (25/70/5% PC/PS/PIP2). The experimental data could only be correctly fitted by assuming a cooperative binding mode. BASP1 displays an apparent affinity for Ca^{2+} of 92 µM with a hill coefficient of 2 ± 0,1, indicating that BASP1 binds at least 2 Ca^{2+} ions

while it is being released from the membrane. GAP-43 exhibits a very similar behaviour, although it seems to be slightly less sensitive to Ca^{2+} since the apparent affinity is 131 μM with a Hill coefficient of 1,38 ± 0,1 indicating that GAP-43 binds at least 2 Ca^{2+} ions.

In this experimental setup both the membrane and the unbound protein will show some affinity towards Ca^{2+}. It is then difficult to estimate the effective Ca^{2+} concentration (concentration of free Ca^{2+}). Consequently our KD values are probably overestimated. However, they clearly show that 100 to 200 μM Ca^{2+} is sufficient to prevent BASP1 and GAP-43 to bind to the membrane.

DISCUSSION

The study of membrane proteins (integral or peripheral) by liquid state NMR faces many technical challenges and limitations, mainly due to the size of the molecular assembly (protein and membrane system). Integral membrane proteins can be reconstituted in detergent micelles, but this type of membrane system is not suitable for the study of peripheral membrane proteins. When using liposome as a membrane system, the molecular assembly formed by the protein bound to the liposome is substantially too large to be tackled by high-resolution techniques. A few methods have been developed for liquid state NMR to tackle this type of systems in which a small/medium protein (transiently) binds to a very large entity (liposome or protein aggregate), methods such as DEST (Fawzi et al., 2011) or CEST (Vallurupalli et al., 2012) are both based on the characterization of an invisible (minor) state in slow exchange with a visible (major) state but require extensive and tedious measurements to provide relatively limited information. Here we chose to tailor the properties of our system in order to make it amenable to high-resolution solution-state NMR, the major innovation being the use of small bicelles as membrane system. Bicelles are routinely used in NMR as alignment media or even as membrane system for embedded and integral membrane proteins (Lee et al., 2008). But to the best of our knowledge, small bicelles have rarely been used to study peripheral membrane proteins by NMR (Khatun et al., 2012; Sommer and Dames, 2014) and never for such a detailed study. Here we showed that small bicelles could be easily be used to achieve a detailed characterisation of the membrane bound forms of GAP-43 and BASP1. Additionally to their small size, they have several advantages that make them potentially very useful membrane system for the study of protein/membrane interaction. They display a flat surface, leading to more native-like protein/membrane interactions. The size of this surface can easily be tuned by varying the DMPS/DHPC ratio (q-value) and therefore adjusted to the type of interaction under scrutiny. Differences of acyl chain lengths determine the distribution of the lipids and allows for strict control of the localisation of a given (target)

lipid type. Additionally, like other artificial membrane system, small bicelles can be doped by the addition of lipids with special properties (PIP2, paramagnetic lipids, chelating lipids etc.). All these factors contribute to make small bicelles and versatile and easy to implement membrane system to study protein/membrane interactions.

In this study we used small bicelles in a combination of liquid state NMR and biophysical methods to delineate the interaction modes of two peripheral IDPs: BASP1 and GAP-43. Whereas the classical membrane interacting modules display spatially well defined binding sites that are complementarily shaped to their target head-group, it is clear that GAP-43 and BASP1 do not follow that mode of recognition. Our data converge to show that the density of positively charged residues along the backbone is the main factor driving the interaction with the negatively charged lipid head-groups. The higher the density of positively charged residues the stronger the overall affinity of the protein for negatively charged membrane surfaces. Indeed we have provided clear evidence that un-acylated BASP1 interacts with the membrane only through a patch of positive residues at its N-terminal extremity (Figure 1C and 5C) and un-acylated GAP-43 through a few basic residues at its N-terminus together with its positively charged IQ domain (Figure 1D and 5F). Specifically, as GAP-43 displays larger binding interfaces than BASP1 it also seems to bind more PS molecules than BASP1 (Figure 2C-D). Additionally, it seems that neither BASP1 nor GAP-43 have a specific affinity towards PIP2 and that PIP2 alone is not able to recruit the proteins at the membrane (Figure 3A-B), which also suggest that un-acylated BASP1 and GAP-43 do not target any particular head-group but are drawn to the membrane through electrostatic attractions. This is also apparent in the liposome binding sigmoidal dependence on the mole fraction of PS in PC liposomes. This common sigmoidal behaviour for peripheral membrane proteins reveals that the density of negative charges at the surface of the membrane is the main factor that will drive the interaction with the membrane. Once bound to the membrane, the protein can interact with additional negatively charged head-groups more easily (hence the apparent cooperativity), which is in accordance with the fact that PIP2 strengthens the interaction between the protein and the membrane (Figure 3C). Of course, since the acylated form of the proteins will be mainly found bound to the membrane, myristoylated BASP1 or palmitoylated GAP-43 will preferentially target PIP2, as it displays the highest density of negative charges, and consequently modulate the availability of the PIP2 pool. The ability of BASP1 and GAP-43 to sequester PIP2 is at the origin of their role in cytoskeleton organisation (Laux et al., 2000), vesicle exocytosis and membrane fusion (Martin, 2012).

We have also shown that in its membrane bound form BASP1 does not undergo a major local conformational change compared to its free form, whereas the IQ domain of GAP-43 displays a slightly higher helical propensity when bound to the membrane. This shows that GAP-43 and BASP1 have very atypical ligand binding properties, even for IDPs, as they do not seem to fold upon binding or assume a more compact conformation in the bound state. Additionally, this slightly higher helical propensity in the IQ domain of GAP-43 upon membrane binding could modulate the interaction with calmodulin, since in the GAP-43/CaM complex the IQ domain is found in a helical conformation. Indeed, one of the two lobes of CaM could bind to the exposed side of the membrane bound IQ domain and co-localize with GAP-43 at the (presynaptic) membrane. This mechanism could be at the origin of GAP-43's ability to regulate the levels of free CaM (Xia and Storm, 2005). However it is unclear whether CaM is able to strip the IQ domain from the membrane, and if it does whether it would strip the protein completely or if GAP-43 would remain attached to the membrane through its N-terminal positive patch and/or its palmitoyl groups.

Finally, we have shown that Ca^{2+} prevent the association of BASP1 and GAP-43 with the membrane, even for Myr-BASP1. Although the exact mechanism by which Ca^{2+} dissociates the proteins from the membrane remains unclear, it seems that both proteins are able to bind multiple Ca^{2+} ions (at least two). This unexpected feature contrasts with the action mode of other prototypical (folded) membrane binding protein where Ca^{2+} is usually found to promote the interaction with the membrane.

All together our data nicely illustrate how a protein devoid of any fold or stable tertiary structure can assume complex physiological functions. The fact that, despite lacking a fold, these proteins are able to harbour multiple binding sites (For PS, PIP2, Ca^{2+} and calmodulin) showcases their modular nature. This modular nature, combined to an extreme sensitivity to external factors, allows IDPs to integrate different physiological signals in order to define a precise mode of action and shows the unique functional plasticity of IDPs.

EXPERIMENTAL PROCEDURES

Cloning

Cloning of the full-length forms of BASP1 and GAP-43 is described elsewhere (Flamm et al., 2016b; Geist et al., 2013). Both proteins were sub-cloned in the pET29b vector yielding to full-length proteins C-terminally fused to a Thrombin cleavage site and a 6*His6-tag. As already described, the palmitoylation sites of GAP-43 (Cys3 and 4) were replaced by glycines in all constructs.

The iLOV fusion proteins (Christie et al., 2012; Gawthorne et al., 2012) were generated as follow: the coding sequence of the protein (BASP1 or GAP-43) was C-terminally fused to a thrombin cleavage site followed to the coding sequence of iLOV, the sequence of the fusion protein was then inserted into the pET 29b vector introducing a cleavable (TEV) 6*His tag at the C-terminal extremity.

The N-terminal domain of GAP-43 (GAP-43-NTD) consists of the first 59 residues of the full-length protein; the construct used for the expression of ^{15}N labeled samples is described elsewhere (Flamm et al., 2016b) and yields to a protein with a N-terminal 6*His tag followed by a TEV cleavage site, the construct used for liposome binding assays consists of GAP-43-NTD followed by a thrombin-cleavage site, the iLOV protein, a TEV cleavage site and a C-terminal 6*His tag. BASP1-NTD was cloned in the pET- 29b vector with a C-terminal 6*His-tag and comprises the first 57 amino acids of the full length human BASP1 followed by a thrombin-cleavage site, the iLOV protein, a C-terminal TEV cleavage site and a 6*His-tag.

Myristoylated BASP1 was obtained by using a bicistronic vector also encoding for the human N-myristoyl transferase, as described elsewhere (Flamm et al., 2016a).

The annotated schematic representation of the constructs used in this study can be found in the supplementary data (Supp. Fig. 1).

Expression and purification of proteins

Unlabeled forms of GAP-43/BASP1 were obtained by expression in rich medium (LB). $^{15}N/^{13}C$ labeled protein were expressed using the *E. coli* strain T7 express (New England BioLab) in minimal medium containing ^{15}N-labeled ammonium chloride and ^{13}C-glucose as sole nitrogen and carbon sources, respectively. In all cases, protein expression was induced at an A_{600} nm of 0.8 by addition of 0.8 mM IPTG. The cells were collected after 16 h of expression at 30 °C by centrifugation at 4,000 rpm for 20 min and resuspended in 30 ml of ice-cold lysis buffer (20 mM NaPi pH 7.4, 50 mM NaCl) per liter of the original bacterial culture. Bacteria were lysed by passing through a French press, and the cell lysate was cleared by centrifugation at 18,000 rpm for 20 min. The supernatant containing the soluble protein fraction was loaded onto a Ni^{2+} loaded HiTrap 5 mL affinity column (GE Healthcare), washed with 2 column volumes of high salt buffer (20 mM NaPi pH 7.4, 1.5 M NaCl, 10 mM imidazole) and eluted with high imidazole buffer (20 mM NaPi pH 7.4, 50 mM NaCl, 0.5 M imidazole) using a linear gradient of 15 column volumes. After either TEV or Thrombin cleavage overnight, the protein was run through Ni^{2+} loaded HiTrap 5 mL affinity column to get rid of the tag and subsequently dialyzed into the measurement buffer and concentrated.

Myristoylated BASP1 was obtained as described elsewhere (Flamm et al., 2016a).

Liposome binding assays

All lipids (Folch fraction, brain PC, brain PE, brain PS, brain PI(4,5)P2, cholesterol, DMPS, DHPC) were purchased from Avanti Polar Lipids, Inc. Stock solutions were prepared with a concentration 20 mg lipid/mL of chloroform. Liposomes were prepared by mixing the different lipid solutions to obtain a final lipid concentration of 5 mg/mL. The chloroform was evaporated under a stream of nitrogen followed by 1 h in a desiccator under vacuum. The measurement buffer was added to the dry residue for a final concentration of 2 mg liposomes per mL. After an incubation time of 30 minutes, the lipids were resuspended and subjected to 3 cycles of freezing (liquid nitrogen) and thawing (hot water) before short sonication in a water bath. The liposome size was restricted by using an extruder with a 0.4 µm filter pore size.

For the sedimentation assays, 30µL of liposomes (2mg/mL) were mixed with 20µg of protein and variable amounts of Ca^{2+} or EDTA for a final volume of 100µL (in 20mM TrisHCl buffer pH 7.4). The mixture was incubated for 30 minutes and centrifuged at 45,000 rpm for 30min at 25°C using the ultracentrifuge rotor TLA-45. The supernatant was collected and the pellet resuspended in the same volume. The protein content of the pellets and supernatants were analyzed using SDS-PAGE. For the iLOV fusion proteins, the fluorescence of the pellets and supernatants were measured with a Tecan plate reader using the iLOV excitation and emission wavelengths (λ_{exc} = 485 nm, λ_{em} = 535 nm).

Small bicelles preparation

DMPS/DHPC bicelles with q = 0.25 were prepared by mixing the appropriate amount of lipids in a glass vial and evaporate the solvent under a stream of nitrogen followed by 1 h in a desiccator under vacuum. The lipids were hydrated with the buffer used for NMR measurements (20mM Bis Tris pH 6.0, 50mM NaCl) to a final concentration of 15%(w/v) of lipids. The solution was incubated for 30 minutes and vortexed. The lipid solution was put at 4 °C for 30 minutes and then at 42 °C for another 30 min, this freeze-thaw cycle was repeated for at least 3 times resulting in a clear solution.

NMR measurements

All NMR data were recorded on un-tagged proteins (no 6*His or iLOV tag). All samples were concentrated up to 0.3 mM of protein in 20 mM BisTris, 50 mM Nacl, pH 6.0 supplemented with 10 % D_2O. NMR experiments were carried out at 25 °C on Varian Inova spectrometers operating at 500, 600 or 800 MHz. All spectra were processed using NMRPipe/NMRDraw (Delaglio et al., 1995) and analyzed with Sparky and CARA (Keller, 2004). The resonance assignments of full length BASP1 and GAP-43 as well as GAP-43-

NTD, were obtained from the BMRB (Flamm et al., 2016b; Geist et al., 2013). Resonance assignments of the BASP1-NTD constructs weas done using standard triple-resonance experiments such as 3D HNCA, 3D HN- (CO)CA, 3D HNCACB, 3D HNCO, and 3D HN(CA)CO. Secondary structure propensities were calculated using $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$, ^{1}H and ^{15}N chemical shift as inputs for the neighbourhood corrected structural propensity calculator (http://nmr.chem.rug.nl/ncSPC/) (Tamiola and Mulder, 2012).

Chemical shifts changes (Δδ) measured in ^{1}H-^{15}N HSQC based titrations were calculated using the following equation (where δ_H and δ_N are the proton and nitrogen chemical shift changes respectively):

$$\Delta\delta = \sqrt{\delta_H^2 + (\frac{\delta_N}{5})^2}$$

Titrations data fitting

Titrations data were analyzed using QtiPlot assuming a cooperative binding mode, using the following equation:

$$\Delta I = A.\frac{\left(L_T - \left(\frac{P_T + L_T + K_{Dapp} - \sqrt{\left(P_T + L_T + K_{Dapp}\right)^2 - 4.P_T.L_T}}{2}\right)\right)^n}{{K_{Dapp}}^n + \left(Lt - \left(\frac{P_T + L_T + K_{Dapp} - \sqrt{\left(P_T + L_T + K_{Dapp}\right)^2 - 4P_T.L_T}}{2}\right)\right)^n}$$

Where Δ*I*, A, L_T, P_T K_{Dapp} and *n* are the variation of fluorescence intensity, the fluorescence intensity at saturation, the total ligand concentration (PS or Ca^{2+}), the total protein concentration and the Hill coefficient respectively.

AUTHOR CONTRIBUTIONS

AF RK and NC designed the experimental approach. AF carried out the experiments. NC supervised the work analyzed the data and wrote the manuscript

ACKNOWLEDGMENTS

This work was supported by Austrian Science Foundation Grants FWF P24761 and P26317.

REFERENCES

Apel, E.D., Byford, M.F., Au, D., Walsh, K.A., and Storm, D.R. (1990). Identification of the protein kinase C phosphorylation site in neuromodulin. Biochemistry *29*, 2330-2335.

Arac, D., Chen, X., Khant, H.A., Ubach, J., Ludtke, S.J., Kikkawa, M., Johnson, A.E., Chiu, W., Sudhof, T.C., and Rizo, J. (2006). Close membrane-membrane proximity induced by Ca(2+)-dependent multivalent binding of synaptotagmin-1 to phospholipids. Nature structural & molecular biology *13*, 209-217.

Bodner, C.R., Dobson, C.M., and Bax, A. (2009). Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. J Mol Biol *390*, 775-790.

Christie, J.M., Hitomi, K., Arvai, A.S., Hartfield, K.A., Mettlen, M., Pratt, A.J., Tainer, J.A., and Getzoff, E.D. (2012). Structural tuning of the fluorescent protein iLOV for improved photostability. J Biol Chem *287*, 22295-22304.

Coudevylle, N., Montaville, P., Leonov, A., Zweckstetter, M., and Becker, S. (2008). Structural determinants for Ca2+ and phosphatidylinositol 4,5-bisphosphate binding by the C2A domain of rabphilin-3A. J Biol Chem *283*, 35918-35928.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR *6*, 277-293.

Durr, U.H., Gildenberg, M., and Ramamoorthy, A. (2012). The magic of bicelles lights up membrane protein structure. Chemical reviews *112*, 6054-6074.

Fawzi, N.L., Ying, J., Ghirlando, R., Torchia, D.A., and Clore, G.M. (2011). Atomic-resolution dynamics on the surface of amyloid-beta protofibrils probed by solution NMR. Nature *480*, 268-272.

Flamm, A.G., Le Roux, A.L., Mateos, B., Diaz-Lobo, M., Storch, B., Breuker, K., Konrat, R., Pons, M., and Coudevylle, N. (2016a). N-Lauroylation during the Expression of Recombinant N-Myristoylated Proteins: Implications and Solutions. Chembiochem : a European journal of chemical biology *17*, 82-89.

Flamm, A.G., Zerko, S., Zawadzka-Kazimierczuk, A., Kozminski, W., Konrat, R., and Coudevylle, N. (2016b). 1H, 15N, 13C resonance assignment of human GAP-43. Biomolecular NMR assignments *10*, 171-174.

Folch, J., Lees, M., and Sloane Stanley, G.H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem *226*, 497-509.

Forsova, O.S., and Zakharov, V.V. (2016). High-order oligomers of intrinsically disordered brain proteins BASP1 and GAP-43 preserve the structural disorder. The FEBS journal *283*, 1550-1569.

Frey, D., Laux, T., Xu, L., Schneider, C., and Caroni, P. (2000). Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity. J Cell Biol *149*, 1443-1454.

Gawthorne, J.A., Reddick, L.E., Akpunarlieva, S.N., Beckham, K.S., Christie, J.M., Alto, N.M., Gabrielsen, M., and Roe, A.J. (2012). Express your LOV: an engineered flavoprotein as a reporter for protein expression and purification. PloS one *7*, e52962.

Geist, L., Zawadzka-Kazimierczuk, A., Saxena, S., Zerko, S., Kozminski, W., and Konrat, R. (2013). (1)H, (1)(3)C and (1)(5)N resonance assignments of human BASP1. Biomolecular NMR assignments *7*, 315-319.

Keller, R. (2004). The computer aided resonance assignment tutorial (Cantina Verlag: Goldau).

Khatun, U.L., Goswami, S.K., and Mukhopadhyay, C. (2012). Modulation of the neurotensin solution structure in the presence of ganglioside GM1 bicelle. Biophysical chemistry *168-169*, 48-59.

Kumar, V., Chichili, V.P., Zhong, L., Tang, X., Velazquez-Campoy, A., Sheu, F.S., Seetharaman, J., Gerges, N.Z., and Sivaraman, J. (2013). Structural basis for the interaction of unstructured neuron specific substrates neuromodulin and neurogranin with Calmodulin. Scientific reports *3*, 1392.

Laux, T., Fukami, K., Thelen, M., Golub, T., Frey, D., and Caroni, P. (2000). GAP43, MARCKS, and CAP23 modulate PI(4,5)P(2) at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. J Cell Biol *149*, 1455-1472.

Lee, D., Walter, K.F., Bruckner, A.K., Hilty, C., Becker, S., and Griesinger, C. (2008). Bilayer in small bicelles revealed by lipid-protein interactions using NMR spectroscopy. J Am Chem Soc *130*, 13822-13823.

Liu, Y., Fisher, D.A., and Storm, D.R. (1993). Analysis of the palmitoylation and membrane targeting domain of neuromodulin (GAP-43) by site-specific mutagenesis. Biochemistry *32*, 10714-10719.

Maekawa, S., Murofushi, H., and Nakamura, S. (1994). Inhibitory effect of calmodulin on phosphorylation of NAP-22 with protein kinase C. J Biol Chem *269*, 19462-19465.

Martin, T.F. (2012). Role of PI(4,5)P(2) in vesicle exocytosis and membrane fusion. Sub-cellular biochemistry *59*, 111-130.

Matsubara, M., Nakatsu, T., Kato, H., and Taniguchi, H. (2004). Crystal structure of a myristoylated CAP-23/NAP-22 N-terminal domain complexed with Ca2+/calmodulin. EMBO J *23*, 712-718.

Montaville, P., Coudevylle, N., Radhakrishnan, A., Leonov, A., Zweckstetter, M., and Becker, S. (2008). The PIP2 binding mode of the C2 domains of rabphilin-3A. Protein Sci *17*, 1025-1034.

Montaville, P., Neumann, J.M., Russo-Marie, F., Ochsenbein, F., and Sanson, A. (2002). A new consensus sequence for phosphatidylserine recognition by annexins. J Biol Chem *277*, 24684-24693.

Moreno, H., Linford, A.S., Gilchrist, C.A., and Petri, W.A., Jr. (2010). Phospholipid-binding protein EhC2A mediates calcium-dependent translocation of transcription factor URE3-BP to the plasma membrane of Entamoeba histolytica. Eukaryotic cell *9*, 695-704.

Mosior, M., and Newton, A.C. (1998). Mechanism of the apparent cooperativity in the interaction of protein kinase C with phosphatidylserine. Biochemistry *37*, 17271-17279.

Paddock, B.E., Striegel, A.R., Hui, E., Chapman, E.R., and Reist, N.E. (2008). Ca2+-dependent, phospholipid-binding residues of synaptotagmin are critical for excitation-secretion coupling in vivo. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 7458-7466.

Sommer, L.A., and Dames, S.A. (2014). Characterization of residue-dependent differences in the peripheral membrane association of the FATC domain of the kinase 'target of rapamycin' by NMR and CD spectroscopy. FEBS Lett *588*, 1755-1766.

Tamiola, K., and Mulder, F.A. (2012). Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. Biochemical Society transactions *40*, 1014-1020.

Tejero-Diez, P., Rodriguez-Sanchez, P., Martin-Cofreces, N.B., and Diez-Guerra, F.J. (2000). bFGF stimulates GAP-43 phosphorylation at ser41 and modifies its intracellular localization in cultured hippocampal neurons. Molecular and cellular neurosciences *16*, 766-780.

Vallurupalli, P., Bouvignies, G., and Kay, L.E. (2012). Studying "invisible" excited protein states in slow exchange with a major state conformation. J Am Chem Soc *134*, 8148-8161.

Warschawski, D.E., Arnold, A.A., Beaugrand, M., Gravel, A., Chartrand, E., and Marcotte, I. (2011). Choosing membrane mimetics for NMR structural studies of transmembrane proteins. Biochim Biophys Acta *1808*, 1957-1974.

Xia, Z., and Storm, D.R. (2005). The role of calmodulin as a signal integrator for synaptic plasticity. Nature reviews Neuroscience *6*, 267-276.

Zakharov, V.V., Capony, J.P., Derancourt, J., Kropolova, E.S., Novitskaya, V.A., Bogdanova, M.N., and Mosevitsky, M.I. (2003). Natural N-terminal fragments of brain abundant myristoylated protein BASP1. Biochim Biophys Acta *1622*, 14-19.

Zakharov, V.V., and Mosevitsky, M.I. (2010). Oligomeric structure of brain abundant proteins GAP-43 and BASP1. J Struct Biol *170*, 470-483.

FIGURE CAPTIONS

Figure 1. Un-acylated BASP1 and GAP-43 bind to Folch liposomes. (**A**) SDS-PAGE analysis of the protein content of the pellets (P) and supernatants (S) from the liposome sedimentation assays. (**B**) Bar diagram representation of the fraction of membrane bound BASP1 and GAP-43 in absence or presence of 2 mM Ca^{2+} as measured by using the fluorescent forms (iLOV fusion) of BASP1 and GAP-43. Intensity ratio of ^{1}H-^{15}N HSQC resonances of BASP1 (**C**) and GAP-43 (**D**) in presence (I) and absence (I_0) of 2 mg.mL^{-1} of Folch liposomes portrayed as a residue plot. The location of GAP-43's IQ domain is represented by a shaded box, schematic representation of the proteins are present at the bottom of the plot, acidic residues are represented in red, basic residues are represented in blue.

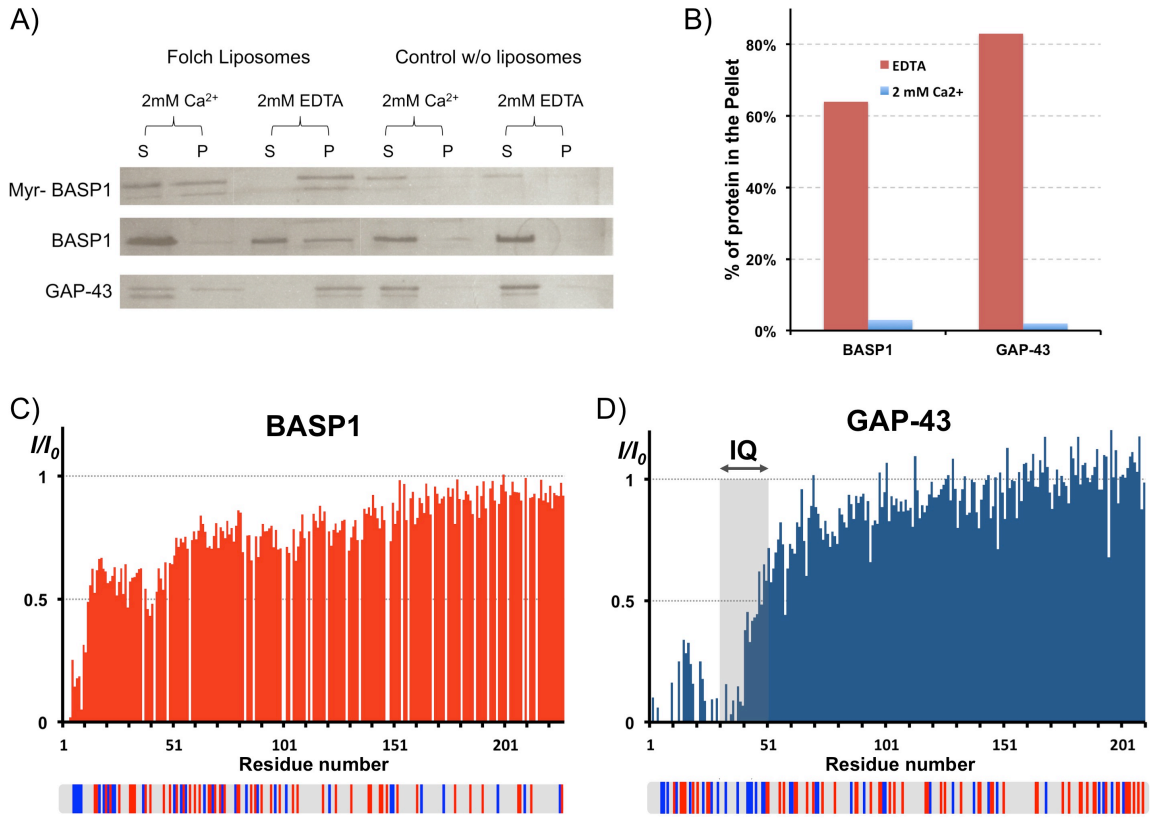


Figure 2. GAP-43 and BASP1 target negative lipid head-groups. Bar diagram representation of the fraction of membrane bound BASP1 (**A**) and GAP-43 (**B**) depending on the liposome composition. Percentage of membrane bound BASP1 (**C**) and GAP-43 (**D**) as a function of the fraction of PS in PC liposomes, the theoretical fit (using the hill's equation as described in material and methods) is depicted as a grey trace.

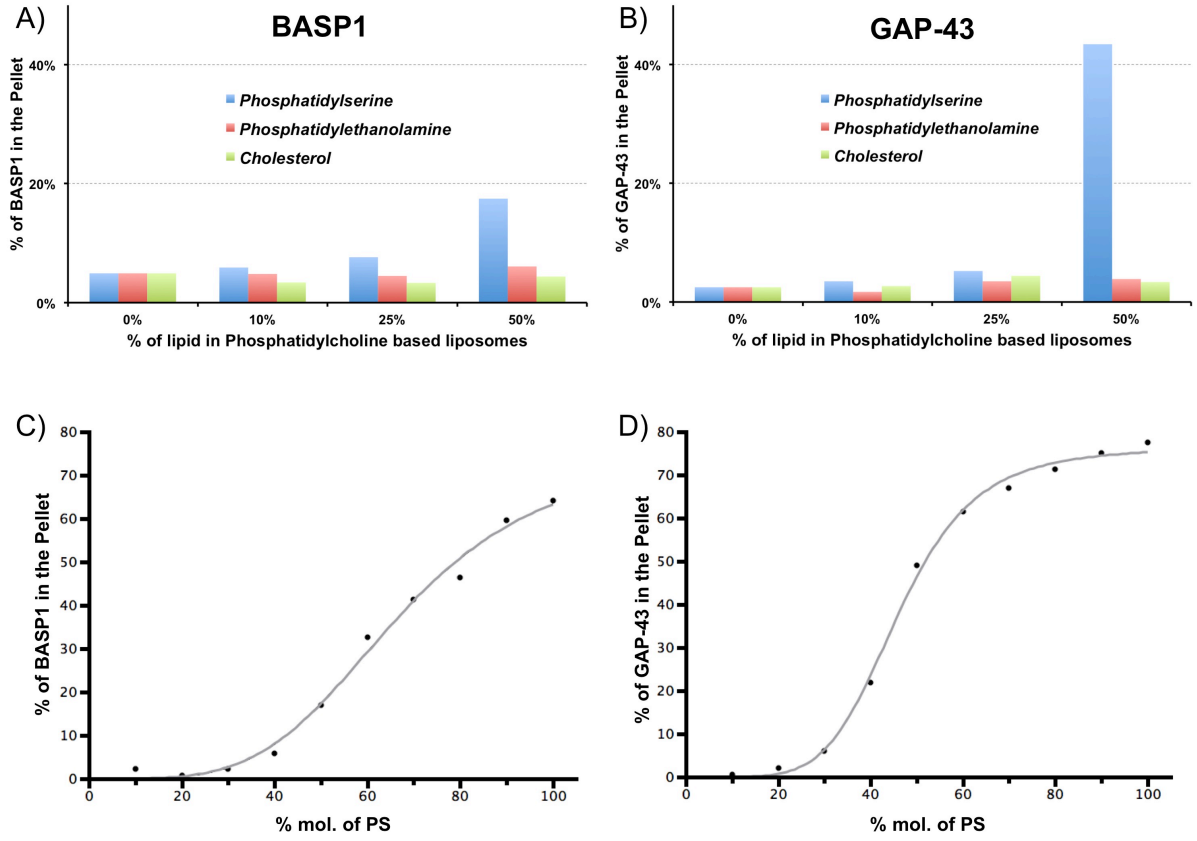


Figure 3. PIP2 cannot recruit BASP1 or GAP-43 at the membrane but strengthen the protein/membrane interaction. Bar diagram representation of the fraction of membrane bound BASP1 (**A**) and GAP-43 (**B**) depending of the presence or absence of 5% PIP2 in PC (orange bars) or 50% PS/PC (blue bars) liposomes. (**C**) Bar diagram representation of the fraction of membrane bound BASP1 (blue bars) and GAP-43 (red bars) for different percentage of PIP2 in 50% PS/PC liposomes.

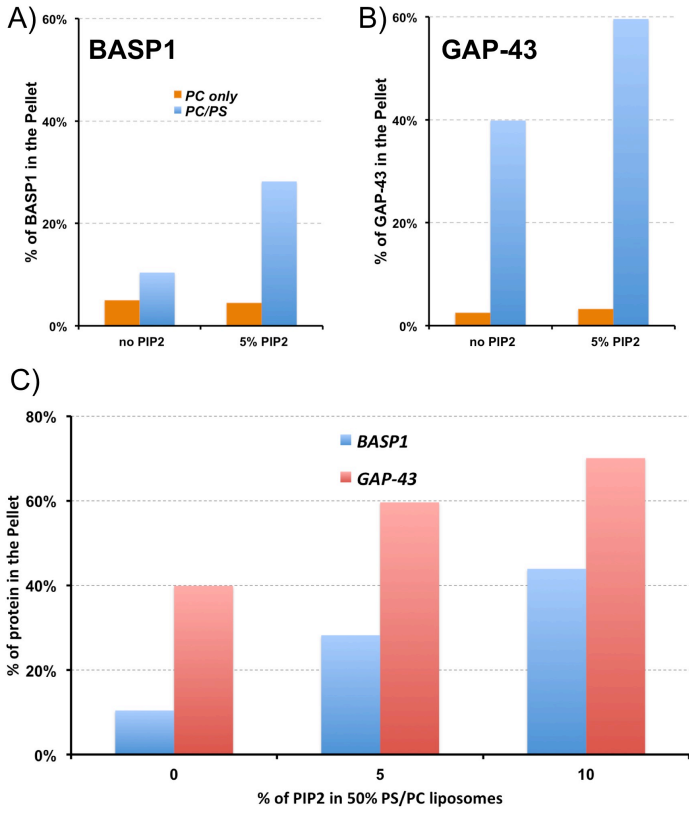


Figure 4. BASP1 and GAP-43 NTDs both bind to small bicelles. Overlay of ^{1}H-^{15}N HSQC spectra in absence (red resonance) and presence (blue resonances) of small bicelles for BASP1 (left panel) and GAP-43 (right panel).

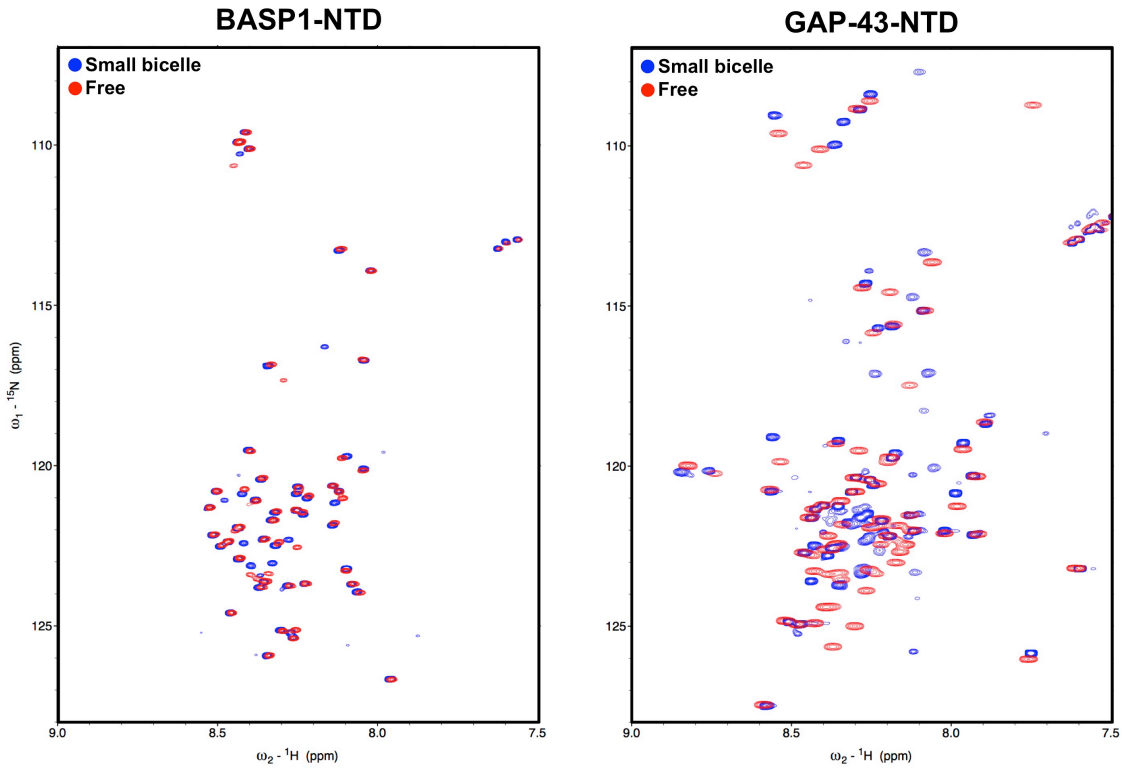


Figure 5. Characterization of the membrane bound state of GAP-43 and BASP1. ^{1}H-^{15}N chemical shift changes Δδ (see Material and Methods) as a function of the total lipid concentration (DHPC+DMPS) for BASP1 Ser6 (**A**) and GAP-43 Gly44 (**B**), the inserts show overlays of the ^{1}H-^{15}N HSQC spectra at different lipid concentration from 0 mM (red resonances) to 10 mM (blue resonances) of Ca^{2+}. ^{1}H-^{15}N chemical shift perturbations δΔ (see Materials and Methods) plotted against the sequence of the BASP1 (**C**) and GAP-43 (**D**) upon the addition of 50 mM DHPC+DMPS. Secondary structure propensity plotted against the sequence of BASP1 (**E**) and GAP-43 (**F**) in the absence (red bars) or presence (blue bars) of DMPS/DHPC bicelles. The location of GAP-43's IQ domain is represented by a shaded box, schematic representation of the proteins are present, acidic residues are represented in red, basic residues are represented in blue.

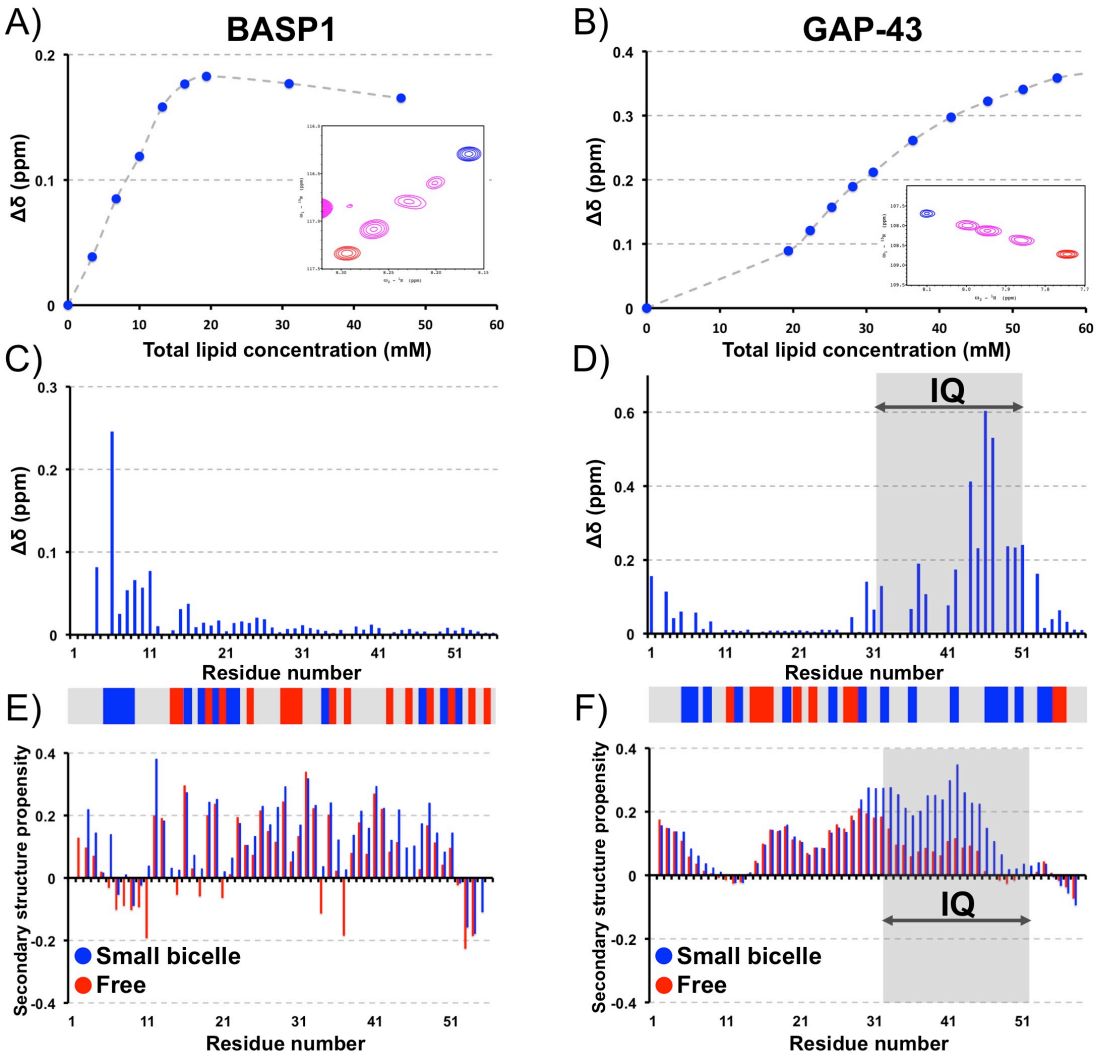
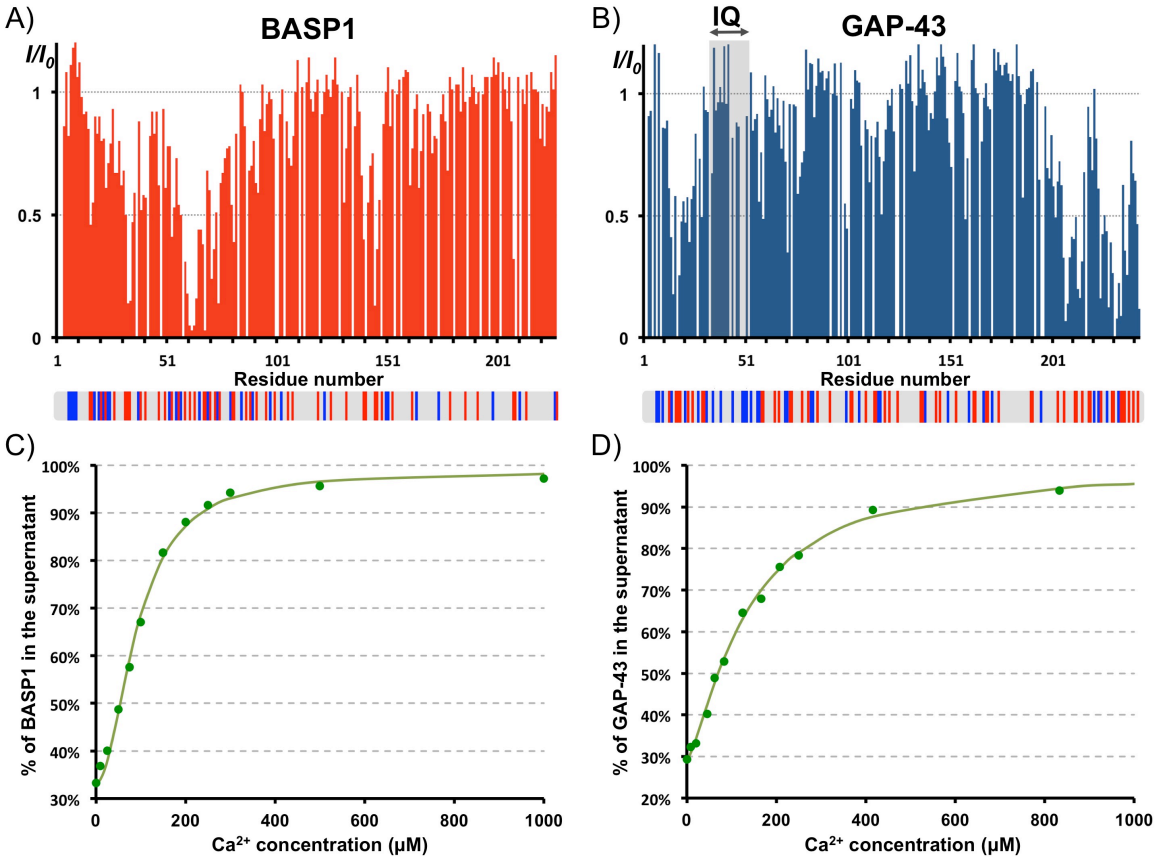


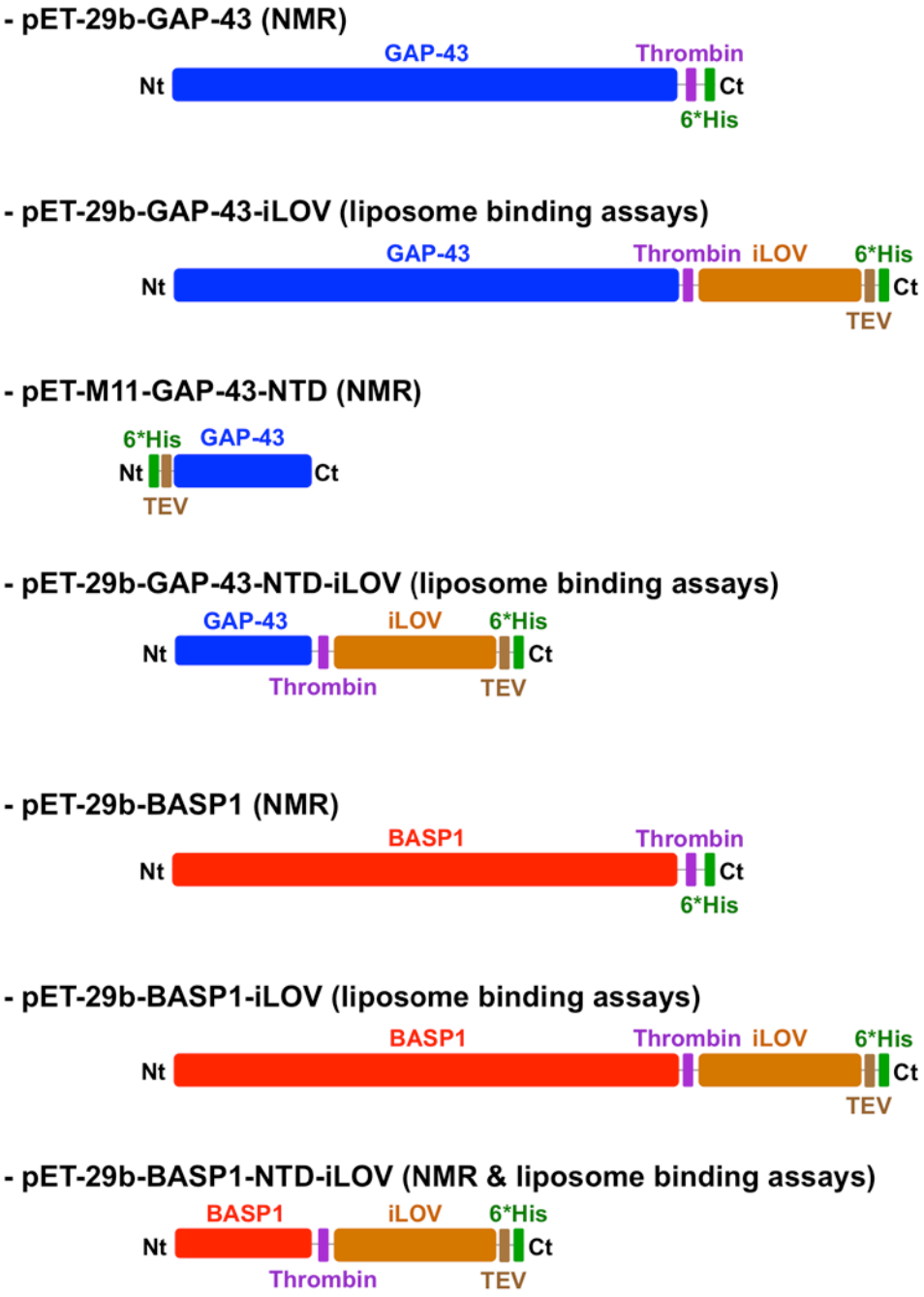
Figure 6. Calcium binding releases GAP-43 and BASP1 form the membrane. Intensity ratio of ^{1}H-^{15}N HSQC resonances of BASP1 (**A**) and GAP-43 (**B**) in presence (I) and absence (I_0) of 30 µM Mn^{2+}, the location of GAP-43's IQ domain is represented by a shaded box, schematic representation of the proteins are present at the bottom of the plot, acidic residues are represented in red, basic residues are represented in blue. Percentage of BASP1 (**C**) or GAP-43 (**D**) found in the supernatant as a function of the Ca^{2+} concentration in presence of 2 $mg.mL^{-1}$ of 25/70/5% PC/PS/PIP2 liposomes, the theoretical fits (using the hill's equation as described in material and methods) are depicted as green traces.



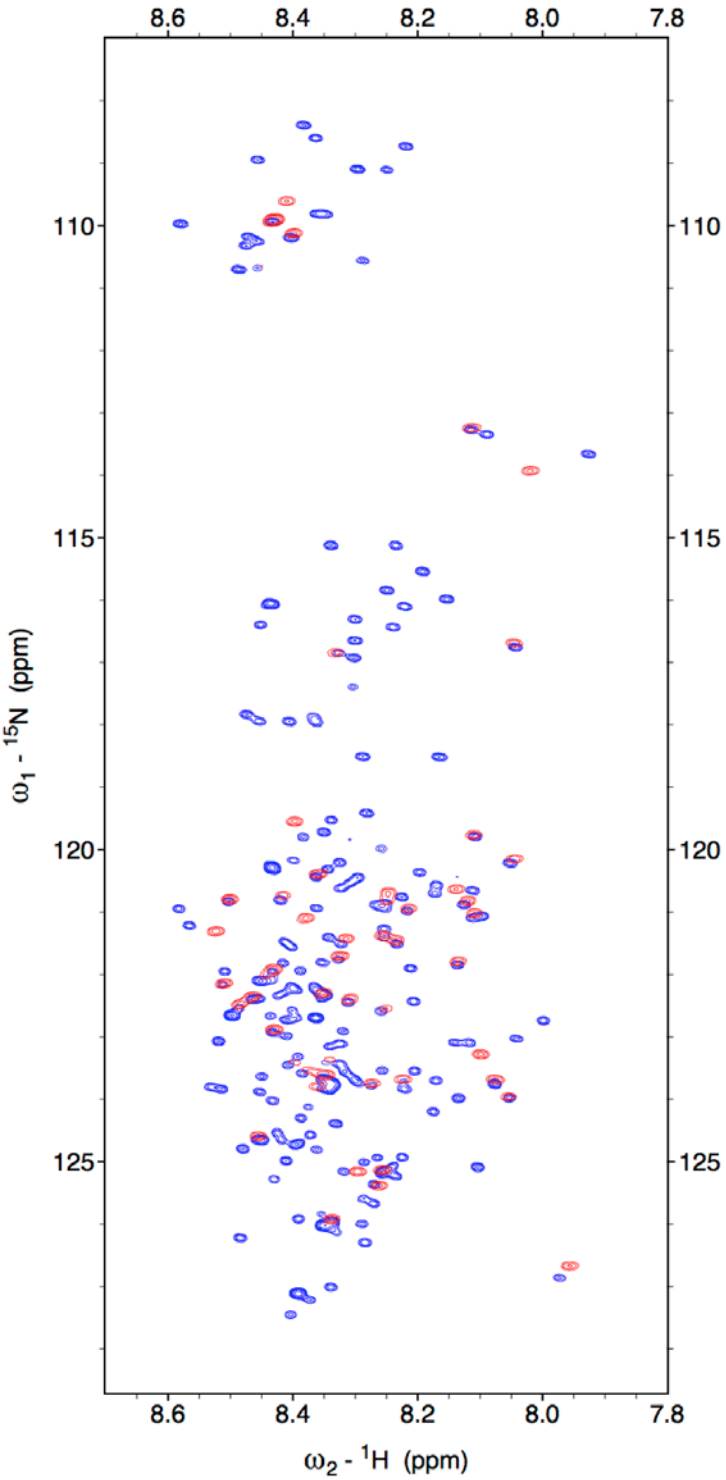
Supplementary data to:

The membrane binding properties of two neuronal intrinsically disordered proteins: BASP1 and GAP-43

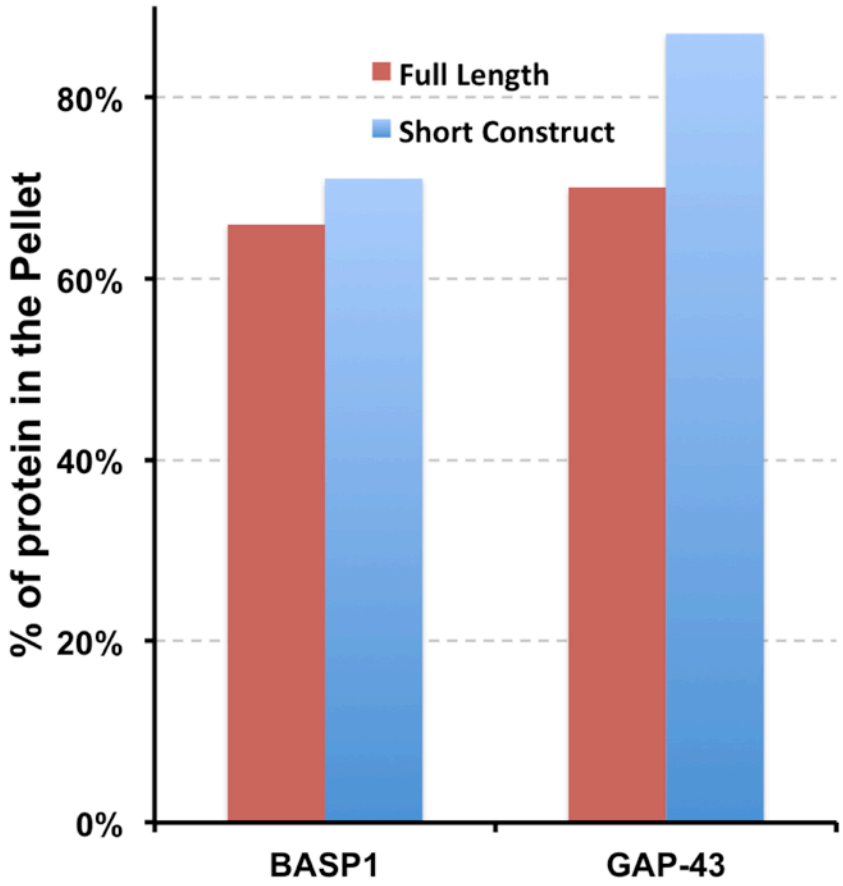
Andrea G. Flamm, Robert Konrat and Nicolas Coudevylle*



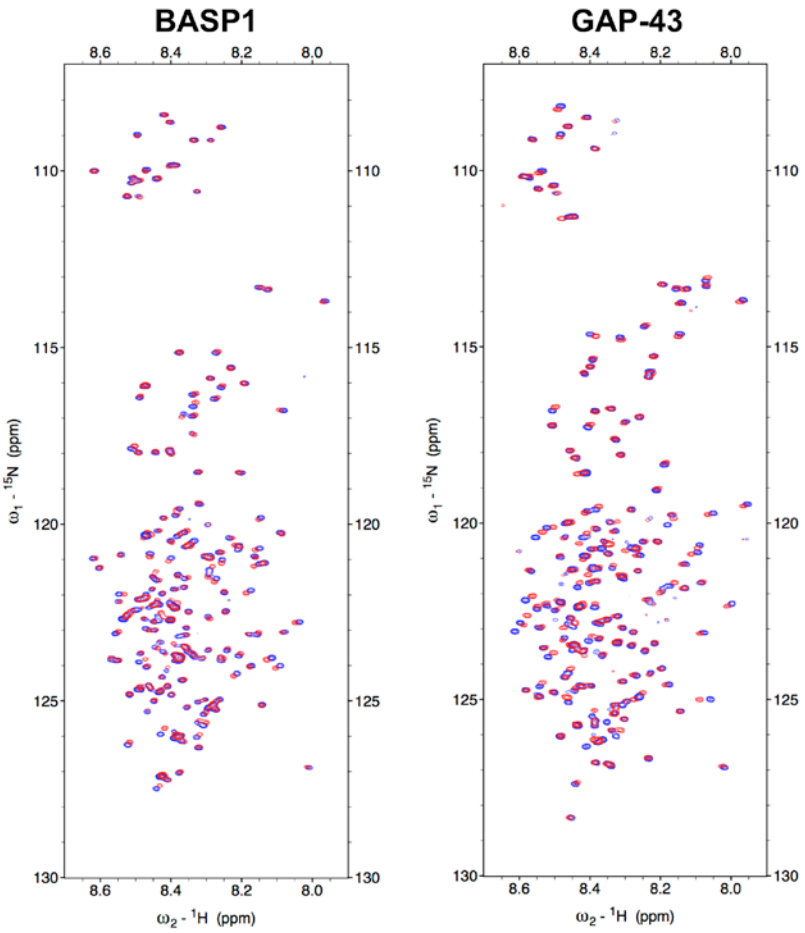
Supp. Figure 1. Schematic representation of the constructs used in the study



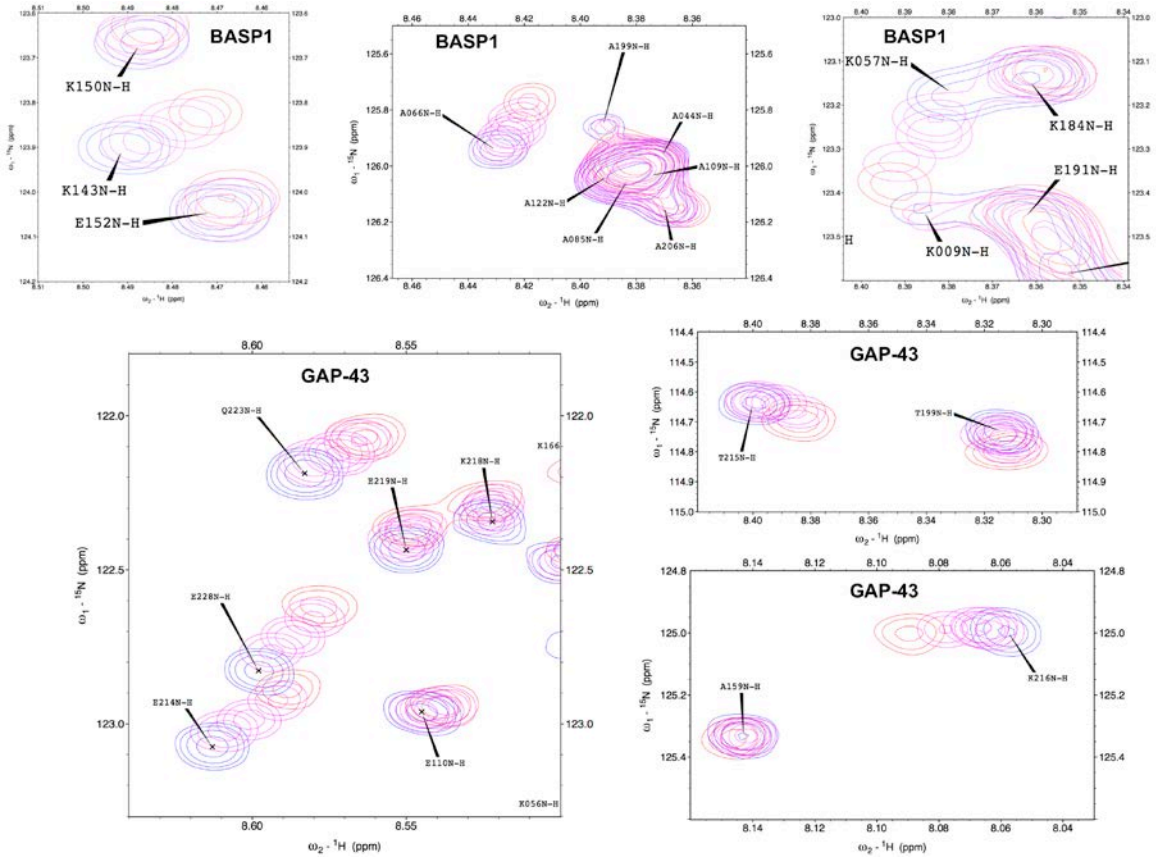
Supp. Figure 2. Overlay of the ^{1}H-^{15}N HSQC spectra of full length BASP1 (blue resonances) and BASP1-NTD (red resonances).



Supp. Figure 3. Bar diagram representation of the fraction of membrane bound of full length and short construct forms of BASP1 and GAP-43 to 25/70/5% PC/PS/PIP2 liposomes.



Supp. Figure 4. Overlays of the ^{1}H-^{15}N HSQC spectra of BASP1 and GAP-43 in absence (blue resonances) and presence (red resonances) of 12 mM Ca^{2+}.



Supp. Figure 5. Close ups of the overlays of the ^{1}H-^{15}N HSQC spectra of BASP1 and GAP-43 in absence (blue resonances) and presence (red resonances) of 12 mM Ca^{2+}. Spectra at intermediate Ca^{2+} concentrations are shown as magenta resonances.

3. Discussion

The topic of this thesis is the interaction and accompanying structural implication of IDPs studied by means of NMR. The IDPs under investigations are the full length human BASP1 and GAP-43. The major interaction studied is the binding of these proteins to membranes and the specificity of these processes.

3.1. Chapter 1

This manuscript resulted from collaboration with Prof. Miquel Pons and Anabel-Lise Le Roux from the University of Barcelona, Spain. Both Prof. Pons' group and ours are working with proteins that are found to be N-terminally myristoylated in the cell. In order to study the impact of this co-translational modification by NMR, we wanted to produce N-myristoylated isotopically labeled BASP1. The common organism used to produce isotopically labeled proteins in large quantities is *E.coli* that, however, lacks the enzyme N-myristoyltransferase, which is responsible for the myristoylation process. It is possible to produce myristoylated and isotopically labeled proteins in *E.coli* by co-expressing the protein of interest and the N-myristoyltransferase.[100] Additionally, the growth medium has to be supplemented with myristic acid. For the case at hand, we chose a strategy developed by Glück et al.[100] who use a bicistronic vector. Surprisingly the expression in minimal medium yields a fully modified protein, however, not homogenously myristoylated (14 carbon fatty acid).[101] Instead, the expression product was a mixture of lauroylated (12 carbon fatty acid) and myristoylated proteins. This was not observed in LB medium, where BASP1 was found to be uniformly myristoylated.[101] A similar problem was independently observed in the group of Prof. Pons for the N-terminal (SH4) region of c-Src. As this appeared as a quite common problem, we decided to engage in collaboration to find a solution. Our collaborative efforts led to the finding that the separation of these two species is very important for biophysical studies. We studied the binding of the two different N-terminal acylated versions of the proteins by SPR. To this end, we chose to coat SPR chips with liposomes (PC/PS) and study the differences in liposome affinities of the two species. These experiments clearly showed the importance of this modification for the binding process: both proteins show a difference in their binding properties to liposomes in dependence of the length of the attached fatty acid. Through this finding we were motivated to develop different purification strategies for the separation of the two species. It is important to note that the purification strategy for the fully disordered BASP1 using the RP- HPLC (reversed phase high

performance liquid chromatography) is possible because the protein can withstand harsh conditions and might not be suitable for folded proteins.

In a next step, we performed NMR experiments of myr-BASP1. These revealed myr-BASP1 oligomers in solution. Their formation dependents on the presence of the myristoyl moiety and, thus, involves the N-terminus of the IDP. As can be seen in the spectrum (see Chapter 1, Supporting Information, Figure S1), signals in the vicinity of the myristoylation site disappear. In the case of BASP1, we already showed in earlier studies based on PRE measurements that the unacylated protein does not show any intermolecular interactions on the N-terminus.[102] We may therefore conclude that the oligomerization is a direct result of the acylation. This exemplifies the major impact of a simple modification on the interaction properties of an IDP.

The major conclusion from Chapter 1 is that it is possible to produce large amounts of isotopically labeled myristoylated proteins by using a bicistronic vector in minimal medium. Here, we observed a heterogeneous modification of the proteins of interest. This led to the discovery of the dependence of the acylation process on the choice of growth medium. The difference in hydrophobicity that results from the myristoylation or lauroylation can be exploited for the separation of the two differently modified species of the protein. This separation turned out to be a necessary prerequisite for biophysical studies of myr-BASP1 as both species have different liposome binding properties.

3.2. Chapter 2

The first step in the characterization of a protein by NMR is the corresponding resonance assignment (cf. section NMR Resonance Assignment). This important achievement is a prerequisite for studying the residue resolved formation of complexes. Our manuscript is concerned with the resonance assignment of GAP-43 by using modern high-dimensionality assignment strategies in cooperation with the group of Wiktor Koźmiński from the University of Warsaw, Poland. The majority of the signals could be assigned by 4D and 5D NMR experiments. Unfortunately, one functionally important part of the protein, the IQ domain, could not be assigned by this procedure. The relevant cross peaks of the IQ domain in the conventional ^{1}H-^{15}N HSQC spectrum are already very weak as a result of fast relaxation. Consequently, they were invisible in the higher dimensional spectra. Due to this and the additional strong signal overlap in the spectra of the full protein, we decided to clone only the first 59 residues of GAP-43 that span the whole IQ domain and the N-terminus of the protein into an appropriate expression vector. The resulting short construct allows assigning GAP-43 completely by simplifying the spectrum of the protein and that enabled the identification of the missing resonances and their assignment.

Naturally, it is not advisable to remove a certain part of a protein without prior structural knowledge. Before designing this construct we therefore performed a PRE analysis of GAP-43 (with a sparse assignment) to look at long range contacts within the protein (not shown in Chapter 2). These might influence the chemical shifts of affected amino acids impeding a resonance assignment that can be transferred to the full- length protein. We found that the N-terminus does sample compact conformers, which bury the residues of the IQ domain in transient compact structures affecting their resonance frequencies. For exactly this reason we chose to include the N-terminal part of the protein in the abovementioned shorter construct to maintain these contacts. Through this procedure the peaks of the N-terminal domain of GAP-43 (referred to as GAP-43 NTD, spanning residues 1-59) in the short construct overlap with the spectrum of the full length protein. Therefore, the assignment of the GAP-43 NTD could be employed for the assignment of the full-length construct. Using these strategies, we have been able to almost fully assign the resonances of human GAP-43.

In a next step, secondary chemical shifts have been used to calculate the secondary structure propensities using neighborhood corrected chemical shift data (cf. Structural Information from NMR). From the secondary chemical shifts obtained from the assignment, we were able to calculate secondary structure propensities. We conclude that GAP-43 is a fully disordered protein. A small alpha helical propensity could be attributed to the IQ

domain. As already described earlier (cf. section Principles of intrinsically disordered protein complexes), regions with increased secondary structure propensity can indicate interaction motifs. This is also true for the case of GAP-43 as was shown by our membrane binding studies (see Chapter 3). From the crystal structure of the CaM-IQ domain complex one can deduce that the IQ domain constitutes the CaM interaction epitope.[68] In this case, this region adopts a helical conformation.[68]

The major conclusion of Chapter 2 is that GAP-43 can be characterized as an intrinsically disordered protein as can be shown by chemical shift analysis from the assignment, but its IQ domain samples pre-structured (alpha helical) binding motifs, which are important for protein (CaM) and membrane interactions.

3.3. Chapter 3

This manuscript concerns the membrane binding properties of the two intrinsically disordered proteins BASP1 and GAP-43. In an initial step we checked for a suitable lipid composition for the binding of the unacylated forms of both proteins. Both proteins bind to Folch liposomes via their basic N-terminal regions. As both proteins are found to be important for neuronal development, we considered the choice of Folch liposomes suitable because the lipids are extracted from brain. We used different liposome composition and performed liposome co-sedimentation assays (cf. section Membrane mimics for studying protein- lipid interactions) to further deepen our knowledge about the influence of the liposomes' compositions on their protein binding properties. These experiments showed a dependence of BASP1 and GAP-43 liposome binding on the presence of negatively charged lipids (phosphatidyserine).

A major question about the interaction of the IDPs and membranes concerns specificity. In a next step, we focused on PIP2- containing liposomes. According to our data the interaction between GAP-43/BASP1 and PIP2 does not rely on a PIP2 binding domain but on the negative charge of PIP2. It seems, hence, plausible that this interaction is based on mutual attraction between PIP2 and positively charged residues of the two IDPs. In this respect, our results show that the mere negative charge of PIP2 laterally sequesters these proteins within a lipid bilayer.

Furthermore, the manuscript sheds light onto the structural changes that occur upon binding of BASP1 and GAP-43 to bicelles. Chemical shift changes were used to calculate the secondary structure differences between the apo and holo forms of both IDPs. Structural rearrangements take place in GAP-43, which can be observed through an increased alpha helical propensity in the IQ domain (yet not in the outermost N-terminal part). In the case of BASP1, we have not been able to detect major structural rearrangements when studying the membrane binding properties. Interestingly though, we also conducted chemical shift analysis of the BASP1-NTD interaction with SDS micelles (not shown here). This leads to an increased helical propensity in the outermost N-terminus that binds the lipids accompanied by major chemical shift changes of the involved residues (upon reaching the critical micelle concentration of SDS). But as we did not consider SDS an appropriate lipid mimetic, we did not include this data in the manuscript. This dependence on the observed structural adaptions on the type of membrane exemplifies the importance of the choice of *in vitro* study conditions. This circumstance has been already described by other groups that compared different lipid mimics.[103] The choice of the biomembrane mimics is crucial for the interpretation and reliability of the *in vitro* results. An *in vitro* experiment will always be limited

by its experimental condition and never be able to incorporate all parameters that have to be considered in order to perfectly reproduce a physiological environment. Yet, it can give valuable information on the possible behavior of a protein in a certain cellular situation and lead to a model of a protein interaction as was demonstrated in this work.

The major conclusion of Chapter 3 is that both unacylated IDPs studied here are drawn to the membrane by electrostatic interaction with negatively charged lipid head groups by interaction with their positively charged N-terminal region. In the case of GAP-43, the binding to bicelles leads to a higher helical propensity of the IQ-domain, while BASP1 does not seem to undergo major local structural rearrangements at the binding site.

4. Abbreviations

BAR	Bin, Amohiohysin and Rvs
BASP1	Brain acid soluble protein
BEST	band selective excitation short transient)
CaM	Calmodulin
CAP-23	cortical associated protein 23
CD	Circular dichroism
Cdc4	Cell division control protein 4
CMC	Critical micelle concentration
CREB	cAMP response element-binding protein
CS	Chemical shift
DAG	Diacylglycerole
DEER	Double electron electron resonance
DPC	dodecylphosphocholine
ENTH	Epsin N-terminal homology domain
EPR	Electron paramagnetic resonance
FRET	Förster Resonance energy transfer
FTIR	Fourier-Transform Infrared Spectroscopy
FYVE	Fab1, YOTB, Vac1 and EEA1
GAP-43	Growth associated protein 43
HMQC	Heteronuclear Multiple Quantum Coherence
HSQC	Heteronuclear Single Quantum Coherence
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
KID	kinase inducible domain
KIX	kinase inducible domain interacting

	domain
LB	Lysogeny broth
MARCKS	Myristoylated alanine-rich C-kinase substrate
MTSL	S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
NAP-22	neuronal tissue-enriched acidic membrane protein
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser enhancement
NTD	N-terminal domain
NUS	Non uniform sampling
ORD	optical rotatory dispersion
PC	phosphatidylcholine
PDB	Protein database
PDDF	Pairwise distance distribution function
PIP2	phosphatidylinositol-4,5-bisphosphate
PKC	Protein Kinase C
PRE	Paramagnetic relaxation enhancement
PS	Phosphatidylserine
PTM	Post-translational modification
PX	Phox homology
RDC	Residual dipolar coupling
SAXS	Small angle Xray scattering
SDS	sodium dodecyl sulphate
Sic1	Substrate, inhibitor of cyclin dependent protein kinase Cdc28
SOFAST	band selective optimized flip angle short transient
SPR	Surface plasmon resonance

5. References

1. Ward, J.J., et al., *Prediction and functional analysis of native disorder in proteins from the three kingdoms of life.* J Mol Biol, 2004. **337**(3): p. 635-45.
2. Uversky, V.N., *A decade and a half of protein intrinsic disorder: biology still waits for physics.* Protein Sci, 2013. **22**(6): p. 693-724.
3. 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.
4. Haynes, C., et al., *Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes.* PLoS Comput Biol, 2006. **2**(8): p. e100.
5. Iakoucheva, L.M., et al., *The importance of intrinsic disorder for protein phosphorylation.* Nucleic Acids Res, 2004. **32**(3): p. 1037-49.
6. Gsponer, J., et al., *Tight regulation of unstructured proteins: from transcript synthesis to protein degradation.* Science, 2008. **322**(5906): p. 1365-8.
7. Uversky, V.N., C.J. Oldfield, and A.K. Dunker, *Intrinsically disordered proteins in human diseases: introducing the D2 concept.* Annu Rev Biophys, 2008. **37**: p. 215-46.
8. Dunker, A.K., et al., *Intrinsic protein disorder in complete genomes.* Genome Inform Ser Workshop Genome Inform, 2000. **11**: p. 161-71.
9. 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.
10. He, B., et al., *Predicting intrinsic disorder in proteins: an overview.* Cell Res, 2009. **19**(8): p. 929-49.
11. Xue, B., et al., *Analysis of structured and intrinsically disordered regions of transmembrane proteins.* Mol Biosyst, 2009. **5**(12): p. 1688-1702.
12. Xue, B., A.K. Dunker, and V.N. Uversky, *Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life.* J Biomol Struct Dyn, 2012. **30**(2): p. 137-49.
13. Oates, M.E., et al., *D(2)P(2): database of disordered protein predictions.* Nucleic Acids Res, 2013. **41**(Database issue): p. D508-16.
14. Frauenfelder, H., S.G. Sligar, and P.G. Wolynes, *The energy landscapes and motions of proteins.* Science, 1991. **254**(5038): p. 1598-603.
15. Berman, H.M., et al., *The archiving and dissemination of biological structure data.* Curr Opin Struct Biol, 2016. **40**: p. 17-22.
16. Fischer, E., *Einfluss der Configuration auf die Wirkung der Enzyme.* Berichte der Deutschen chemischen Gesellschaft zu Berlin 1894. **27**(3): p. 2985–93.
17. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
18. Radivojac, P., et al., *Protein flexibility and intrinsic disorder.* Protein Sci, 2004. **13**(1): p. 71-80.
19. Eliezer, D., *Characterizing residual structure in disordered protein States using nuclear magnetic resonance.* Methods Mol Biol, 2007. **350**: p. 49-67.
20. Fuxreiter, M., et al., *Preformed structural elements feature in partner recognition by intrinsically unstructured proteins.* J Mol Biol, 2004. **338**(5): p. 1015-26.
21. Wilkins, D.K., et al., *Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques.* Biochemistry, 1999. **38**(50): p. 16424-31.
22. Kurzbach, D., et al., *Compensatory adaptations of structural dynamics in an intrinsically disordered protein complex.* Angew Chem Int Ed Engl, 2014. **53**(15): p. 3840-3.
23. Kurzbach, D., et al., *Cooperative unfolding of compact conformations of the intrinsically disordered protein osteopontin.* Biochemistry, 2013. **52**(31): p. 5167-75.

24. Fasman, G.D., *Circular dichroism and the conformational analysis of biomolecules*. The language of science. 1996, New York: Plenum Press. ix, 738 p.
25. Maltsev, A.S., J. Ying, and A. Bax, *Impact of N-terminal acetylation of alpha-synuclein on its random coil and lipid binding properties.* Biochemistry, 2012. **51**(25): p. 5004-13.
26. Uversky, V.N., J. Li, and A.L. Fink, *Evidence for a partially folded intermediate in alpha-synuclein fibril formation.* J Biol Chem, 2001. **276**(14): p. 10737-44.
27. Galea, C.A., et al., *Role of intrinsic flexibility in signal transduction mediated by the cell cycle regulator, p27 Kip1.* J Mol Biol, 2008. **376**(3): p. 827-38.
28. Lee, J.C., et al., *Alpha-synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's disease.* Proc Natl Acad Sci U S A, 2004. **101**(47): p. 16466-71.
29. Mukhopadhyay, S., et al., *A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2649-54.
30. Balasubramaniam, D. and E.A. Komives, *Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins.* Biochim Biophys Acta, 2013. **1834**(6): p. 1202-9.
31. Sinz, A., *Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions.* Mass Spectrom Rev, 2006. **25**(4): p. 663-82.
32. Hubbard, S.J., *The structural aspects of limited proteolysis of native proteins.* Biochim Biophys Acta, 1998. **1382**(2): p. 191-206.
33. Maekawa, S., et al., *Purification and molecular cloning of a novel acidic calmodulin binding protein from rat brain.* J Biol Chem, 1993. **268**(18): p. 13703-9.
34. Ikura, M., L.E. Kay, and A. Bax, *A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin.* Biochemistry, 1990. **29**(19): p. 4659-67.
35. Geist, L., et al., *(1)H, (1)(3)C and (1)(5)N resonance assignments of human BASP1.* Biomol NMR Assign, 2013. **7**(2): p. 315-9.
36. Flamm, A.G., et al., *(1)H, (15)N, (13)C resonance assignment of human GAP-43.* Biomol NMR Assign, 2016. **10**(1): p. 171-4.
37. Kizilsavas, G., et al., *(1)H, (1)(3)C, and (1)(5)N backbone and side chain resonance assignments of the C-terminal DNA binding and dimerization domain of v-Myc.* Biomol NMR Assign, 2013. **7**(2): p. 321-4.
38. Platzer, G., et al., *(1)H, (15)N, (13)C resonance assignment of human osteopontin.* Biomol NMR Assign, 2015. **9**(2): p. 289-92.
39. Kazimierczuk, K., et al., *High-dimensional NMR spectra for structural studies of biomolecules.* Chemphyschem, 2013. **14**(13): p. 3015-25.
40. Maciejewski, M.W., et al., *Data sampling in multidimensional NMR: fundamentals and strategies.* Top Curr Chem, 2012. **316**: p. 49-77.
41. Schanda, P., H. Van Melckebeke, and B. Brutscher, *Speeding up three-dimensional protein NMR experiments to a few minutes.* J Am Chem Soc, 2006. **128**(28): p. 9042-3.
42. Schanda, P. and B. Brutscher, *Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds.* J Am Chem Soc, 2005. **127**(22): p. 8014-5.
43. Kurzbach, D., A.G. Flamm, and T. Sara, *Network representation of protein interactions - experimental results.* Protein Sci, 2016.
44. Wishart, D.S., et al., *1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J Biomol NMR, 1995. **5**(1): p. 67-81.
45. Kjaergaard, M., S. Brander, and F.M. Poulsen, *Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH.* J Biomol NMR, 2011. **49**(2): p. 139-49.
46. Tamiola, K., B. Acar, and F.A. Mulder, *Sequence-specific random coil chemical shifts of intrinsically disordered proteins.* J Am Chem Soc, 2010. **132**(51): p. 18000-3.

47. Smith, L.J., et al., *Analysis of main chain torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations.* J Mol Biol, 1996. **255**(3): p. 494-506.
48. Rule, G.S. and T.K. Hitchens, *Fundamentals of protein NMR spectroscopy.* Springer Science & Business Media, 2006. **5**.
49. Mohana-Borges, R., et al., *Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings.* J Mol Biol, 2004. **340**(5): p. 1131-42.
50. Baker, J.M., et al., *CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices.* Nat Struct Mol Biol, 2007. **14**(8): p. 738-45.
51. Felli, I.C.E. and R.E. Pierattelli, *Intrinsically Disordered Proteins Studied by NMR Spectroscopy.* Springer, 2015. **870**.
52. Kurzbach, D., *Network representation of protein interactions: Theory of graph description and analysis.* Protein Sci, 2016. **25**(9): p. 1617-27.
53. Clore, G.M. and J. Iwahara, *Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes.* Chem Rev, 2009. **109**(9): p. 4108-39.
54. Schorghuber, J., et al., *Novel approaches in selective tryptophan isotope labeling by using Escherichia coli overexpression media.* Chembiochem, 2015. **16**(5): p. 746-51.
55. Felli, I.C. and R. Pierattelli, *Novel methods based on (13)C detection to study intrinsically disordered proteins.* J Magn Reson, 2014. **241**: p. 115-25.
56. Bah, A. and J.D. Forman-Kay, *Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications.* J Biol Chem, 2016. **291**(13): p. 6696-705.
57. Alberts, B., J.H. Wilson, and T. Hunt, *Molecular biology of the cell*. 5th ed. 2008, New York: Garland Science. xxxiii, 1601, 90 p.
58. Mitrea, D.M. and R.W. Kriwacki, *Regulated unfolding of proteins in signaling.* FEBS Lett, 2013. **587**(8): p. 1081-8.
59. Bah, A., et al., *Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch.* Nature, 2015. **519**(7541): p. 106-9.
60. Selenko, P. and G. Wagner, *Looking into live cells with in-cell NMR spectroscopy.* J Struct Biol, 2007. **158**(2): p. 244-53.
61. Ito, Y. and P. Selenko, *Cellular structural biology.* Curr Opin Struct Biol, 2010. **20**(5): p. 640-8.
62. Luchinat, E. and L. Banci, *A Unique Tool for Cellular Structural Biology: In-cell NMR.* J Biol Chem, 2016. **291**(8): p. 3776-84.
63. Amata, I., et al., *Multi-phosphorylation of the intrinsically disordered unique domain of c-Src studied by in-cell and real-time NMR spectroscopy.* Chembiochem, 2013. **14**(14): p. 1820-7.
64. Theillet, F.X., et al., *Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs).* Chem Rev, 2014. **114**(13): p. 6661-714.
65. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Current Opinion in Structural Biology, 2009. **19**(1): p. 31-38.
66. Radhakrishnan, I., et al., *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions.* Cell, 1997. **91**(6): p. 741-52.
67. Radhakrishnan, I., et al., *Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB.* FEBS Lett, 1998. **430**(3): p. 317-22.
68. Kumar, V., et al., *Structural basis for the interaction of unstructured neuron specific substrates neuromodulin and neurogranin with Calmodulin.* Sci Rep, 2013. **3**: p. 1392.
69. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions.* Trends Biochem Sci, 2008. **33**(1): p. 2-8.
70. Mittag, T., et al., *Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17772-7.

71. Davey, N.E., et al., *Attributes of short linear motifs.* Mol Biosyst, 2012. **8**(1): p. 268-81.
72. Widmer, F. and P. Caroni, *Identification, localization, and primary structure of CAP-23, a particle-bound cytosolic protein of early development.* J Cell Biol, 1990. **111**(6 Pt 2): p. 3035-47.
73. Maekawa, S., H. Murofushi, and S. Nakamura, *Inhibitory effect of calmodulin on phosphorylation of NAP-22 with protein kinase C.* J Biol Chem, 1994. **269**(30): p. 19462-5.
74. Zakharov, V.V., et al., *Natural N-terminal fragments of brain abundant myristoylated protein BASP1.* Biochim Biophys Acta, 2003. **1622**(1): p. 14-9.
75. Takasaki, A., et al., *Identification of the calmodulin-binding domain of neuron-specific protein kinase C substrate protein CAP-22/NAP-22. Direct involvement of protein myristoylation in calmodulin-target protein interaction.* J Biol Chem, 1999. **274**(17): p. 11848-53.
76. Matsubara, M., et al., *Crystal structure of a myristoylated CAP-23/NAP-22 N-terminal domain complexed with Ca2+/calmodulin.* EMBO J, 2004. **23**(4): p. 712-8.
77. Laux, T., et al., *GAP43, MARCKS, and CAP23 modulate PI(4,5)P(2) at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism.* J Cell Biol, 2000. **149**(7): p. 1455-72.
78. Korshunova, I., et al., *Characterization of BASP1-mediated neurite outgrowth.* J Neurosci Res, 2008. **86**(10): p. 2201-13.
79. Frey, D., et al., *Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity.* J Cell Biol, 2000. **149**(7): p. 1443-54.
80. Caroni, P., L. Aigner, and C. Schneider, *Intrinsic neuronal determinants locally regulate extrasynaptic and synaptic growth at the adult neuromuscular junction.* J Cell Biol, 1997. **136**(3): p. 679-92.
81. Wiederkehr, A., J. Staple, and P. Caroni, *The motility-associated proteins GAP-43, MARCKS, and CAP-23 share unique targeting and surface activity-inducing properties.* Exp Cell Res, 1997. **236**(1): p. 103-16.
82. Iino, S., et al., *Motor, sensory and autonomic nerve terminals containing NAP-22 immunoreactivity in the rat muscle.* Brain Res, 2004. **1002**(1-2): p. 142-50.
83. Yamamoto, Y., Y. Sokawa, and S. Maekawa, *Biochemical evidence for the presence of NAP-22, a novel acidic calmodulin binding protein, in the synaptic vesicles of rat brain.* Neurosci Lett, 1997. **224**(2): p. 127-30.
84. Green, L.M., et al., *Dynamic interaction between WT1 and BASP1 in transcriptional regulation during differentiation.* Nucleic Acids Res, 2009. **37**(2): p. 431-40.
85. Mosevitsky, M.I., et al., *The BASP1 family of myristoylated proteins abundant in axonal termini. Primary structure analysis and physico-chemical properties.* Biochimie, 1997. **79**(6): p. 373-84.
86. Hartl, M., et al., *Inhibition of Myc-induced cell transformation by brain acid-soluble protein 1 (BASP1).* Proc Natl Acad Sci U S A, 2009. **106**(14): p. 5604-9.
87. Bahler, M. and A. Rhoads, *Calmodulin signaling via the IQ motif.* FEBS Lett, 2002. **513**(1): p. 107-13.
88. Apel, E.D., et al., *Identification of the protein kinase C phosphorylation site in neuromodulin.* Biochemistry, 1990. **29**(9): p. 2330-5.
89. Alexander, K.A., et al., *Regulation of calmodulin binding to P-57. A neurospecific calmodulin binding protein.* J Biol Chem, 1987. **262**(13): p. 6108-13.
90. Tejero-Diez, P., et al., *bFGF stimulates GAP-43 phosphorylation at ser41 and modifies its intracellular localization in cultured hippocampal neurons.* Mol Cell Neurosci, 2000. **16**(6): p. 766-80.
91. Liu, Y., D.A. Fisher, and D.R. Storm, *Analysis of the palmitoylation and membrane targeting domain of neuromodulin (GAP-43) by site-specific mutagenesis.* Biochemistry, 1993. **32**(40): p. 10714-9.

92. Liang, X., et al., *Mass spectrometric analysis of GAP-43/neuromodulin reveals the presence of a variety of fatty acylated species.* J Biol Chem, 2002. **277**(36): p. 33032-40.
93. Shen, Y., et al., *Growth-associated protein-43 is required for commissural axon guidance in the developing vertebrate nervous system.* J Neurosci, 2002. **22**(1): p. 239-47.
94. He, Q., E.W. Dent, and K.F. Meiri, *Modulation of actin filament behavior by GAP-43 (neuromodulin) is dependent on the phosphorylation status of serine 41, the protein kinase C site.* J Neurosci, 1997. **17**(10): p. 3515-24.
95. Hurley, J.H., *Membrane binding domains.* Biochim Biophys Acta, 2006. **1761**(8): p. 805-11.
96. McLaughlin, S. and D. Murray, *Plasma membrane phosphoinositide organization by protein electrostatics.* Nature, 2005. **438**(7068): p. 605-11.
97. Antonny, B., *Mechanisms of membrane curvature sensing.* Annu Rev Biochem, 2011. **80**: p. 101-23.
98. Warschawski, D.E., et al., *Choosing membrane mimetics for NMR structural studies of transmembrane proteins.* Biochim Biophys Acta, 2011. **1808**(8): p. 1957-74.
99. Kielec, J.M., K.G. Valentine, and A.J. Wand, *A method for solution NMR structural studies of large integral membrane proteins: reverse micelle encapsulation.* Biochim Biophys Acta, 2010. **1798**(2): p. 150-60.
100. Gluck, J.M., et al., *Single vector system for efficient N-myristoylation of recombinant proteins in E. coli.* PLoS One, 2010. **5**(4): p. e10081.
101. Flamm, A.G., *NMR studies of the intrinsically disordered protein BASP1.* Masterarbeit, Universität Wien. Zentrum für Molekulare Biologie. BetreuerIn: Konrat, Robert, 2013.
102. Kurzbach, D., et al., *Detection of correlated conformational fluctuations in intrinsically disordered proteins through paramagnetic relaxation interference.* Phys Chem Chem Phys, 2016. **18**(8): p. 5753-8.
103. Cross, T.A., et al., *Influence of solubilizing environments on membrane protein structures.* Trends Biochem Sci, 2011. **36**(2): p. 117-25.