



universität  
wien

# DISSERTATION

Titel der Dissertation

„Structural Dynamics of the Intrinsically Disordered  
Protein BASP1“

Verfasser

Mag.rer.nat. Leonhard Geist

angestrebter akademischer Grad

Doktor der Naturwissenschaft (Dr.rer.nat.)

Wien, 2013

Studienkennzahl lt.

A 091 490

Studienblatt:

Dissertationsgebiet lt.

Dr.-Studium der Naturwissenschaften Molekulare Biologie

Studienblatt:

Betreuer:

Univ.-Prof. Dr. Robert Konrat



# ***Acknowledgements***

This work wouldn't have been possible without the scientific, technical and moral support of many people, whom I would like to thank.

I'm especially grateful to my supervisor Robert Konrat for his constant support and infectious enthusiasm, creating an outstanding learning and working environment, supporting me not only in my scientific career but also in my musical ambitions. Moreover, special thanks to Georg Kontaxis, Karin Kloiber and Martin Tollinger for their patient help with NMR-technical and software related issues.

Mostly, I want to thank all the members of the NMR-group, especially Karin Ledolter, Nicolas Coudeville, Gerald Platzner, Gönül Kizilsavas, Thomas Schwarz and Sandra Haiderer, for the enjoyable working atmosphere, the fruitful and crazy discussions during coffee breaks and their willingness to help in all aspects of work and everyday life. Special thanks to Morkos Henen, who is a great friend and finds always time for lifting ones spirits.

I also want to thank our collaborators in Warsaw, the team of Wiktor Koźmiński, and our collaborators in Mainz, Dariush Hinderberger and Dennis Kurzbach, who turned out to be an excellent addition to our team.

Finally I want to thank my parents Maria and Ernst and my sister Anna-Maria and brother Clemens for being as perfect as they are, supporting and encouraging and always there, when you need somebody.

Thanks to all my good friends and thank you, Mahsa, for your support in the final stages of my thesis and for enduring my insanities.



# List of Publications

- 1) **Geist L**, Henen MA, Haiderer S, Schwarz T, Kurzbach D, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Koźmiński W, Hinderberger D, Konrat R (2013)  
*Protein Sci.* 2013 Jul 2. doi: 10.1002/pro.2304. PubMed PMID: 23821606.  
Protonation-dependent Conformational Variability of Intrinsically Disordered Proteins.
- 2) **Geist L**, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Koźmiński W, Konrat R (2012)  
*Biomol NMR Assign.*  
(1)H, (13)C and (15)N resonance assignments of human BASP1.
- 3) Stanek J, Saxena S, **Geist L**, Konrat R, Koźmiński W (2013)  
*Angew Chem Int Ed Engl*, 52(17):4604-6  
Probing local backbone geometries in intrinsically disordered proteins by cross-correlated NMR relaxation.
- 4) Solyom Z, Schwarten M, **Geist L**, Konrat R, Willbold D, Brutscher B (2013)  
*J Biomol NMR*, 55(4):311-21  
BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins.
- 5) Orbán-Németh Z, Henen MA, **Geist L**, Zerko S, Saxena S, Stanek J, Koźmiński W, Propst F, Konrat R (2013)  
*Biomol NMR Assign.*  
Backbone and partial side chain assignment of the microtubule binding domain of the MAP1B light chain.
- 6) Henen MA, Coudevylle N, **Geist L**, Konrat R (2012)  
*J Med Chem*, 55(17):7909-19  
Toward rational fragment-based lead design without 3D structures.
- 7) Coudevylle N, Hoetzing M, **Geist L**, Kontaxis G, Hartl M, Bister K, Konrat R (2011)  
*Biochemistry*, 50(43):9192-9  
Lipocalin Q83 reveals a dual ligand binding mode with potential implications for the functions of siderocalins.
- 8) Coudevylle N, **Geist L**, Hoetzing M, Tollinger M, Konrat R (2011)  
*J Biomol NMR*, 51(1-2):83-8  
Siderocalin Q83 exhibits differential slow dynamics upon ligand binding.

- 9) Coudevylle N, **Geist L**, Hötzing M, Hartl M, Kontaxis G, Bister K, Konrat R (2010)  
*J Biol Chem*, 285(53):41646-52  
The v-myc-induced Q83 lipocalin is a siderocalin.
- 10) Auer R, Kloiber K, Vavrinska A, **Geist L**, Coudevylle N, Konrat R (2010)  
*J Am Chem Soc*, 132(5):1480-1  
Pharmacophore mapping via cross-relaxation during adiabatic fast passage.

# Table of Contents

<b>Acknowledgements</b>	<b>i</b>
<b>List of Publications</b>	<b>iii</b>
<b>Table of Contents</b>	<b>v</b>
<b>Preface</b>	<b>vii</b>
<b>1 Intrinsically Disordered Proteins</b>	<b>1</b>
<b>2 The Meta-Structure concept</b>	<b>7</b>
<b>3 The many facets of BASP1</b>	<b>11</b>
3.1 <i>BASP1 and its role in neuronal cells</i>	13
3.2 <i>BASP1 and its role in transcription regulation</i>	16
<b>4 NMR analysis of BASP1's structural propensities and pH-induced compaction</b>	<b>19</b>
4.1 <i>Secondary Structure Propensities in IDPs</i>	19
4.2 <i>HET-SOFAST spectroscopy</i>	22
4.2.1 HET <sup>noe</sup> -SOFAST:	24
4.2.2 HET <sup>ex</sup> -SOFAST:	25
4.3 <i><sup>13</sup>C-filtered AFP-NOESY-<sup>1</sup>H-<sup>15</sup>N-HSQC</i>	27
4.3.1 The Nuclear Overhauser Effect (NOE)	27
4.3.2 Adiabatic Fast Passage (AFP)	30
4.3.3 1D AFP-NOESY	31
4.3.4 <sup>13</sup> C-filtered AFP-NOESY- <sup>1</sup> H- <sup>15</sup> N-HSQC	35
<b>Discussion</b>	<b>39</b>
<b>Summary - Zusammenfassung</b>	<b>42</b>
<b>References</b>	<b>45</b>
<b>Article 1</b>	<b>51</b>
<b>Article 2</b>	<b>65</b>
<b>CURRICULUM VITAE</b>	<b>71</b>





# ***Preface***

The thesis is based on Article 1 and 2 and is organized as follows:

In Chapter 1 I will shortly introduce Intrinsically Disordered Proteins (IDPs). IDPs have attracted a lot of attention in recent years due to their involvement in many important physiological processes, notably their implication in regulatory pathways (signal transduction, transcriptional control, etc.) and their association with human diseases (e.g. cancer, neurodegenerative diseases, etc.) (Uversky et al, 2008). IDPs contradict the classic structure-function paradigm, as they fulfill their function(s) despite lacking a stably folded tertiary structure, with their intrinsic flexibility actually having a significant impact on their biological functionality. The sampling of a vast and heterogeneous conformational space allows IDPs to interact with, and control, multiple binding partners at once, allowing for unique regulatory functionalities in living organisms (Fink, 2005; Tompa, 2002). As IDPs often constitute interaction hubs in the protein interactomes of organisms, they have attracted considerable scientific interest (Uversky et al, 2009).

It is difficult to get structural information of IDPs. Clearly this protein family is not amenable to X-ray crystallographic structure determination. NMR spectroscopy, which has been developed into a powerful structural biology technique complementing protein X-ray crystallography, offers unique opportunities for structural and dynamic studies of IDPs (Eliezer, 2007; Jensen et al, 2013; Mittag & Forman-Kay, 2007; Schneider et al, 2012; Sibille & Bernado, 2012). As NMR provides ensemble-averaged data it still remains a challenge to interpret this data in terms of an ensemble of distinct conformations. The meta-structure approach provides important insights for the interpretation of NMR data. Chapter 2 will outline the basics of this novel computational approach that views proteins as a network of interacting residues and evaluates – only based on the primary sequence – how certain residues are embedded inside this network (Konrat, 2009). Numerous applications have already demonstrated the applicability of this approach (Hasenohrl et al, 2011; Henen et al,

2012; Mayer et al, 2012; Schedlbauer et al, 2011) and most recently we used the meta-structure approach to assess structural changes in IDPs under acidic conditions (Article 1).

In Chapter 3 I will briefly discuss, what is known about the cellular functions of the IDP BASP1, highlighting its implication in quite diverse cellular processes: On the one hand BASP1 is interacting with membranes, sequestering acidic phospholipids into lipid rafts and influencing actin-cytoskeleton dynamics; on the other hand it was discovered as a tumor suppressor, acting as co-suppressor of the potent transcription regulator Wilms' Tumor Suppressor Protein 1 (WT1). For NMR structural and dynamic studies of BASP1 NMR chemical shift assignment is necessary, which was accomplished with the help of novel 5D NMR experiments in combination with random sampling of the indirect time domain specially designed for the assignment of IDPs (Article 2).

Due to their highly flexible nature IDPs constitute very sensitive probes to changes in environmental conditions, particularly changes in pH. For example, it was shown that a reduction in pH influences local secondary structure propensities especially in the C-terminal acidic region of  $\alpha$ -synuclein and that the conformational ensemble is enriched in more compact structures (Cho et al, 2009). During my thesis I primarily focused on the structural dynamics of BASP1 and the influence of pH on its conformational ensemble. We have established with the help of NMR spectroscopic methods, that BASP1 is similarly affected by changes in the pH, populating significantly more compact conformations at low pH with an increase in  $\alpha$ -helical secondary structure propensities in acidic residue-enriched regions of the protein (Chapter 4 and Article 1). We developed a novel NMR method,  $^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC, for the measurement of  $^1\text{H}$ - $^1\text{H}$  homonuclear cross-relaxation between aliphatic and amide protons during adiabatic fast passage (AFP). AFP pulse strength dependent modulation of the effective cross-relaxation provides information about conformational dynamics and differential compaction along the protein backbone. Chapter 4.3 gives a theoretic background of the method, with a final explanation of the pulse-sequence and results on BASP1 (Chapter 4.3.4), which are in line with the observation of pH-induced compaction. Electron paramagnetic resonance (EPR)-based double electron-electron resonance (DEER) measurements further corroborate meta-structure and NMR-based findings. Interestingly, a large-scale meta-structure analysis of the pH-dependence of IDPs revealed that there is a significant tendency of IDPs in general to adopt more compact conformations with a tendency towards  $\alpha$ -helical secondary structure formation under acidic conditions (Article 1).

# ***Chapter 1***

## ***Intrinsically Disordered Proteins***

For more than 100 years it was believed that function is inherent to structure starting with the lock-and-key proposal of Fischer in 1894 (Fischer, 1894). In the decades to follow, vast amounts of protein 3D-structures provided by X-ray crystallography as well as nuclear magnetic resonance spectroscopy (NMR) has led to a fortification of the classic structure-function paradigm, leaving no space for alternative views.

The key to the function of a protein was believed to lie in the positioning of specific side-chain atoms in an exact three-dimensional arrangement allowing for the catalysis of chemical reactions, specific interaction with other proteins, binding of ligands, etc. Lack of a stable structure was seen as non-functional, probably its only use to flexibly link “functional” folded parts of a polypeptide chain. The availability of fully sequenced genomes, together with gene-based functional analysis, led to the discovery of intrinsically disordered/unstructured proteins (IDPs) as well as disordered polypeptide stretches in otherwise well-folded proteins that fulfill important molecular functions in cells (Dyson & Wright, 2005). Consequently, the structure-function paradigm, quoting that only a well-folded protein is able to perform a function, has been questioned in the last decades. Additionally, advances in methodology especially in the field of NMR spectroscopy, which can provide site-specific dynamic information of a protein covering different time-scales, shifted the view of proteins as static entities to a more dynamic one.

IDPs have been discovered to be implicated in many different cellular processes, their functional repertoire actually exceeding that of structured proteins (Dunker et al, 2008; Xie et al, 2007b). In particular, IDPs are often associated with regulatory functions and play a central role in protein interaction networks, which probably explains their preponderance in

higher eukaryotes. It has been estimated that 10-35% of prokaryotic and about 15-45% of eukaryotic proteins contain significant disorder (i.e. disordered regions longer than 30 residues), with even higher predictions in the subset of human signaling (~67%) and cancer-associated proteins (~80%) (Iakoucheva et al, 2002; Tompa, 2012).

IDPs are highly flexible molecules usually envisaged as random-coil like molecules that only fold upon interaction. As a matter of fact, it becomes more and more clear in recent years, that the terms “unstructured”, “disordered” or “random-coil” are not very convenient to describe a protein family, that comprises proteins with different dynamic behaviors and capabilities to transiently form structure. IDPs exist as an ensemble of rapidly fluctuating conformations in their native state, different regions of the polypeptide exhibiting differential tendencies to form secondary structure or even tertiary contacts. This conformational plasticity confers IDPs the potential to interact with multiple binding partners and it was proposed that this adaptability allows them to efficiently engage in weak regulatory networks (such as transcription regulation) (Tantos et al, 2012). Hence, the emergence of disorder in proteins might have been a crucial step in the evolution of higher organisms and it has been shown that the occurrence of disorder correlates with evolutionary complexity (Dosztanyi et al, 2006).

One of the most intensively studied example of a typical IDP is  $\alpha$ -synuclein, which has a high propensity to aggregate and is associated with neurodegenerative disorders including Parkinson’s disease (PD) and Alzheimer’s disease (AD) (Uversky et al, 2008). Especially the influence of environmental conditions on the conformational ensemble of  $\alpha$ -synuclein has been investigated, showing that  $\alpha$ -synuclein is able to sample different conformations from highly extended to strongly enriched in  $\alpha$ -helical content, highlighting that the choice of conformations in an IDP is strongly dependent on the peculiarities of its environment (Uversky, 2003).

The idea that IDPs are highly adaptable to changing environmental conditions can be simply explained by comparing hypothetical energy landscapes of a typical stably folded protein and an IDP. The energy landscape of a globular protein depicts the typical folding funnel with one deep energy minimum, where the protein adopts its three dimensional fold (Dill & Chan, 1997). Energy barriers are usually too high to allow for large-scale conformational changes. Even though, especially NMR spectroscopic studies showed that even folded proteins are highly dynamic in their native state and populate excited states, which are often associated with the specific function of the protein (Bruschweiler et al, 2013; Bruschweiler et al, 2009). The energy landscape of an IDP rather resembles a highly rugged

surface with a lot of accessible local energy minima and small barriers in between those minima not allowing the formation of a stably folded conformation. The interaction with a particular ligand strongly affects the energy landscape, making some energy minima deeper and/or energy barriers higher, (transiently) trapping an IDP in a certain conformation (Uversky et al, 2008).

One of the most fundamental questions concerning IDPs is, how they interact with other molecules. Several models have been proposed, with the folding upon binding model being the most extensively studied. The disorder to order transition during the binding process is associated with an entropic loss resulting in a rather weak and transient interaction. Still, high specificity can be preserved, as it is rather dependent on the size and complementarity of the interaction surface. Thus, it has been postulated that variation of the length and the degree of preformation of an interaction motif allows for fine-tuning of the interaction affinity and simultaneously preserving high specificity. This type of interaction, combining low affinity (transient interaction) with high specificity, is beneficial in signal transduction networks, where information from different pathways has to be integrated and transmitted quickly (Wright & Dyson, 2009). Still, certain complexes involving intrinsically disordered proteins that interact in a folding upon binding mechanism have reportedly high affinity (p27 binds to a Cdk2-cyclin A complex with a nanomolar affinity) (Lacy et al, 2004). Moreover, it has been noted that electrostatic interactions play a major role in interactions of IDPs. The “polyelectrostatic” interaction is mediated by multiple charges in combination with the rapid interconversion of different conformations leading to a “mean electrostatic field” reflecting the net charge of the whole disordered interaction segment significantly contributing to the free energy of binding (Borg et al, 2007; Mittag et al, 2010). This allows also for tuning of a specific interaction by post-translational modifications (PTMs), as a lot of those modifications will modify the net charge of an IDP. Additionally, the presentation of a charged interaction surface will lead to conformational rearrangements within the conformational ensemble of a highly charged IDP, e.g. bringing negatively charged residues closer together to match a positively charged surface. This polarizability is an additional advantage of an interaction mediated by an IDP and will significantly contribute to the affinity of a “disordered” complex (Mittag et al, 2010). The appreciation of disorder in functional protein complexes led to the creation of the term “fuzziness” in protein complexes, which tries to classify the spectrum of disorder possible in protein complexes. Static disorder describes the possibility of different well-defined binding conformations in a complex, whereas dynamic disorder implies the retention of flexibility and the rapid fluctuation

between different conformers in the bound state (Mittag et al, 2010; Tompa & Fuxreiter, 2008).

Most IDPs undergo several post-translational modifications (PTMs) such as acylation, phosphorylation, glycosylation or ubiquitination (Iakoucheva et al, 2004; Xie et al, 2007a). As mentioned before, due to their highly flexible nature, IDPs are strongly influenced by PTMs leading to substantial changes in structural dynamics, i.e. the sampling of different conformations, and strongly affecting the protein's binding behavior. Besides its importance, the influence of PTMs in IDPs is still marginally addressed. Interestingly, in the field of Molecular Dynamics simulation there are recent progresses to incorporate PTMs into simulations (Margreitter et al, 2013).

Due to the importance and overrepresentation of IDPs and disordered regions in signaling and major disease pathways, prompting the creation of the term 'disorder in disorders' (Uversky et al, 2008), IDPs have been suggested as potential drug targets and several studies have shown that it is possible to target disordered regions with small molecules (Metallo, 2010).

A lot of proteins do not fit into the dichotomic partitioning structured as opposed to unstructured. Structural features of proteins appear to span a continuum from stably folded to random-coil like: tightly folded single domains, to multidomain proteins with flexible or disordered regions, to compact but disordered molten globules, to highly extended, heterogeneous unstructured states.

Confronted with this structural continuum of functional proteins, the scientific community tries to classify proteins in terms of their conformational state. One of this classifications is the "protein trinity" concept, which tries to conceptualize native proteins as part of three distinct states – ordered, collapsed-disordered (molten globule) and extended-disordered (random coil) analogous to the aggregation states of matter (solid – liquid – gas). Interconversions are comparable to phase transitions (Dunker et al, 2001). Well-folded, ordered proteins correspond to the solid state harboring short-range (secondary structural elements) as well as long-range interactions (tertiary contacts between distant parts of the polypeptide). Extended-disordered proteins are comparable to the gas state, lacking short- and long-range interactions. Proteins primarily showing short-range interactions but being devoid of specific long-range interactions correspond to the liquid state (e.g.: molten globules). The "protein trinity" classification was expanded to four different conformational states incorporating the pre-molten globule state characterized by a considerable amount of residual

secondary structure, as compared to an extended-disordered state, but being significantly less compact than the ordered and molten globule states (Uversky, 2009).

Importantly, transitions between different protein states are blurred and not nearly as sharp as for bulk phases (Dunker et al, 2001). Therefore, it is not always possible to strictly assign a given protein to any of these states, some proteins switching between conformational sub-states as an answer to changing environmental conditions. It is obvious that protein conformational space is very heterogeneous and that proteins adopt conformations on a continuum of possibilities, where function can arise from any of those conformational states or from transitions between them.

The meta-structure concept, introduced in the following chapter, not only provides information about secondary structure propensities on a residue basis, but also how residues are embedded into the three-dimensional network constituted by the protein. Importantly, meta-structure parameters are only calculated from the protein's primary sequence. The additional information about residue-based structural complexity provides important clues about the degree of order or flexibility associated with a certain secondary structural element and hence allows a quick assessment of the structural features of the protein in question. Consequently, it can help the decision-making process in terms of the method of choice for structural investigation of a given protein. Further, it strongly facilitates the interpretation of data derived from NMR spectroscopy as well as other structure-biology methods.





# ***Chapter 2***

## ***The Meta-Structure concept***

Bioinformatics and computational chemical biology provided us with a multitude of different tools that allow prediction of specific protein parameters, including type of secondary structure, degree of disorder, and more. Still, up until now it is impossible to predict a protein's three-dimensional structure and its function only based on its primary sequence. Interestingly, it was believed that the knowledge of the 3D structure of a protein provides the key to understand its function, but even with the tremendous amount of solved protein structures the functionality of a lot of proteins remains elusive.

The protein meta-structure concept introduced by Robert Konrat in 2009 (Konrat, 2009) represents a novel approach for the elucidation of a protein's secondary structure content but also its residue-based level of compaction. Conventional predictors rely on properties of individual amino acids or on statistical evaluations, which amino acids are found in specific secondary structural elements. The meta-structure concept doesn't a priori take into account specific properties of amino acids but is based on a statistical evaluation of 3D structures of proteins. It views the protein structure as a network of interacting residues and evaluates how certain residues are embedded inside this network. Accordingly, the occurrence of specific amino acid types in certain protein regions and their proximity to other amino acids reflects indirectly certain amino acid properties.

3D structural information (coordinate triples for the individual atoms in the protein) of a subset of proteins of the PDB used by Bax and coworkers (Kontaxis et al, 2005) was transformed into topological space, where each node represents the C<sup>α</sup> position of a certain residue and a connection (or edge) between two nodes indicates a C<sup>α</sup>-C<sup>α</sup> distance smaller than 8Å, hence a close neighborhood. These networks of interacting residues were then

statistically evaluated in terms of shortest path lengths  $\theta$  connecting two residues and as a function of the residue type (A,B) and the distance  $l_{AB}$  in the primary sequence between those residues. The resultant distribution of shortest path lengths for a given residue pair with a given primary sequence distance  $l_{AB}$  is then stored as pairwise distribution function  $q(\theta, A, B, l_{AB})$ . These statistical distribution functions can subsequently be used to predict specific topological parameters solely based on the primary sequence.

(1) The compactness value  $C_i$ :

At first, the primary sequence is used to calculate for each residue pair (A, B) an average topology parameter  $d_{ij}$ , which is related to the most probable shortest path length between residues  $i$  and  $j$  and is based on the pairwise distribution functions  $q(\theta, A, B, l_{AB})$ :

$$d_{ij} = \sum \theta \times \rho(\theta, A, B, l_{AB})$$

The compactness value  $C_i$  for a given residue  $i$  is given by:

$$C_i = N_1 \times \sum \frac{1}{d_{ij}} - N_0$$

$N_1$  and  $N_0$  are empirical scaling factors ensuring  $C_i$  values close to 0 for highly exposed residues. As  $C_i$  is inversely proportional to  $d_{ij}$  it follows that residues with small shortest path lengths to neighboring residues give high compactness values and are expected for residues deeply buried inside a protein.

(2) The local secondary structure parameter  $S_i$ :

For the prediction of local secondary structure only next neighbor distribution functions are used ( $l_{AB} \leq 4$ ) and is defined as:

$$\begin{aligned} S_i &= {}^\alpha P_i - {}^\beta P_i \\ {}^\alpha P_i &= N_A \times \gamma_2 \times \gamma_3 \times \gamma_4 \\ {}^\beta P_i &= N_B \times \delta_3 \times \delta_4 \\ \gamma_i &= \frac{\rho(1,i) \times (1 - \rho(2,i))}{\rho(2,i) \times (1 - \rho(1,i))} \end{aligned}$$

$$\delta_i = \gamma_i^{-1}$$

$q(1,i)$  and  $q(2,i)$  are the probabilities to find a shortest path length  $\theta$  of 1 or 2 between two residues separated by  $i$  positions in the primary sequence.  $N_A$  and  $N_B$  are again scaling factors ensuring comparable magnitudes of  $C_i$  and  $S_i$  values.  $S_i$  values are similarly defined as NMR secondary chemical shift values, with positive values indicating  $\alpha$ -structure and negative values indicating extended ( $\beta$ -strand) structure.

A major advantage of the meta-structure approach is that it not only provides information about secondary structural elements, but also how these elements are embedded in the 3D fold. The combination of residue-based secondary structure and compactness is called the meta-structure.

This novel approach offers unique possibilities for chemical (molecular) biology, structural genomics and drug discovery. For example, it has been demonstrated that it is possible based on meta-structure homologies to identify valid starting points for fragment-based lead (drug) discovery (FBLD, FBDD) (Article 6, (Henen et al, 2012)), as it is able to identify protein similarities that are hidden on the primary sequence level and hence will not be detected by conventional sequence analysis tools (BLAST) (Altschul et al, 1990). In this approach, pairwise meta-structure alignment of the target protein to proteins of the DRUGBANK database (biologically relevant proteins with experimentally verified inhibitory ligands) (Wishart et al, 2006) will provide a list of potential ligands that can be further experimentally verified and refined (Henen et al, 2012). Furthermore, meta-structure analysis has proven a very powerful tool in the identification and investigation of IDPs as well as intrinsically disordered regions (IDRs). Compactness values of typical IDPs and IDRs are very low (average  $C_i \approx 230$ ) as compared to well-folded protein regions (average  $C_i \approx 330$ ) (Geist et al, 2013; Konrat, 2009). Additionally, local secondary structure values in regions of low compactness provide information about residual structure in IDPs or IDRs similar to NMR secondary chemical shift analysis, where the deviation from random coil chemical shift values of backbone atoms is used to estimate secondary structure propensities. Additionally, meta-structure compactness values are a valuable source of information for the choice of paramagnetic spin-label attachment for paramagnetic relaxation enhancement (PRE) studies. Spin-labels are introduced preferentially at sites of very low compactness to avoid the interference of the spin-label with compact or compaction-prone regions. Especially in IDPs, introduction of the bulky spin-label into a (transiently) compact region will strongly influence

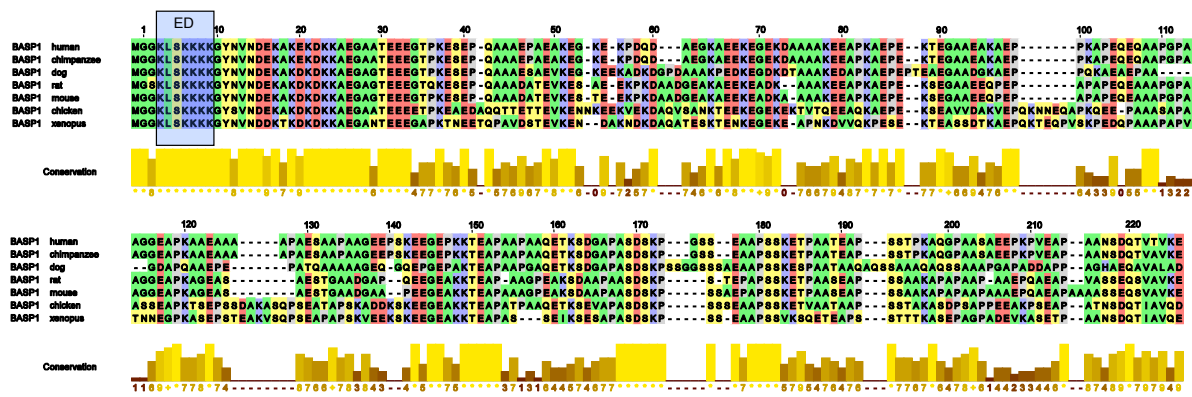
the sampling of certain conformations and hence lead to misleading results concerning residual structure and dynamics of the IDP. Finally, meta-structure analysis could be successfully applied to deduce local secondary structural changes and differential compaction in IDPs upon pH reduction. For this purpose, meta-structure parameters were calculated from the native form as well as a “mutated” form of the IDP, i.e. glutamic acid and aspartic acid residues were substituted in the primary sequence by glutamine and asparagine residues, in order to imitate the protonated Glu/Asp carboxyl-group at low pH. It could be shown that IDPs in general have a tendency to form more compact conformations at low pH with a tendency towards  $\alpha$ -helix formation (Article1, (Geist et al, 2013)).

# ***Chapter 3***

## ***The many facets of BASP1***

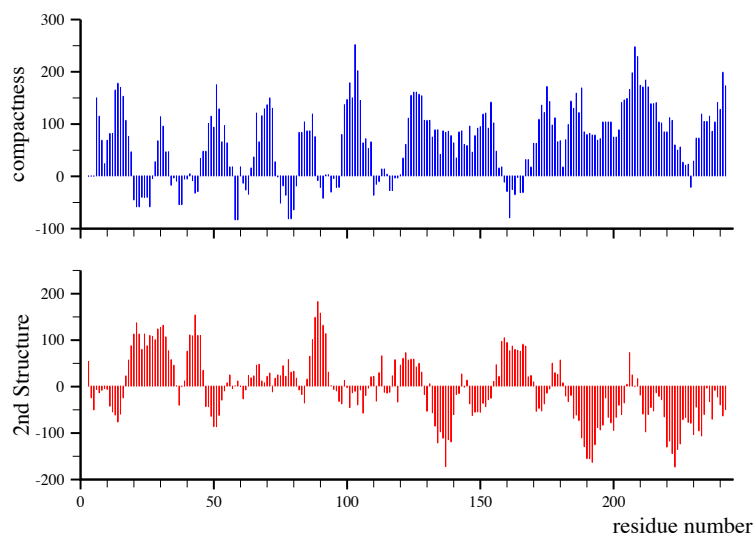
BASP1 [brain acid-soluble protein 1, brain abundant-membrane attached signal protein 1, also known as NAP-22 (neuron-specific acidic protein of 22 kDa) (Maekawa et al, 1993) and CAP-23 (cortical cytoskeleton-associated protein of 23 kDa) (Widmer & Caroni, 1990)] is a fascinating protein, as it appears to fulfill quite diverse tasks in the cell. Numerous studies are investigating its role in plasma membrane organization and neurite outgrowth, whereas others try to shed light on its implication in transcription regulation. Nonetheless, it remains unclear how BASP1 interacts with proteins and exerts its functions.

BASP1 belongs to the class of intrinsically disordered proteins (IDPs). As IDPs are less affected by mutation as compared to globular proteins, where the integrity of a distinct three-dimensional fold needs to be preserved, it is not surprising that BASP1 homologs from different species share only limited sequence homology. Fig. 1 shows a protein sequence alignment of BASP1 from different species. The high conservation of the N-terminal region (residues 1-34), harboring the basic effector domain (residues 3-9), highlights its importance for BASP1 function.



**Fig. 1.** Primary sequence alignment of BASP1 homologues from different species. The colour code outlines roughly amino acid properties: acidic (red), basic (blue), hydrophobic (green), polar (yellow) and proline (grey). The histogram below depicts the level of conservation between the sequences. The blue box indicates the basic effector domain (ED) spanning residues 3-9. The N-terminal methionine is cut off during N-terminal myristoylation of glycine 1.

NMR spectra of BASP1 show the for IDPs typical narrow peak dispersion in the  $^1\text{H}$  dimension, and X-ray crystallographic data exists only of a myristoylated N-terminal fragment (residues 1-9) of human BASP1 in complex with  $\text{Ca}^{2+}$ -CaM (Matsubara et al, 2004). The meta-structure approach corroborates the high degree of disorder in BASP1 predicting very low compactness values ( $< 200$ ) and only small stretches of  $\alpha$ -helical propensity, with most of the protein rather adopting extended conformation (Fig. 2). Still, despite very low compactness values, some parts of BASP1 (residues 20-45 and residues 87-92) are marked by significant  $\alpha$ -helical propensities ( $\text{Si} > 100$ ), highlighting tendencies of these regions to transiently adopt helical conformation.



**Fig. 2.** Meta-structure analysis of chicken BASP1. Low compactness values ( $C_i < 200$ ) are typical for IDPs. Some regions, especially in the N-terminus, show slight  $\alpha$ -helical propensity.

### 3.1 BASP1 and its role in neuronal cells

BASP1 belongs to the family of growth-associated proteins, including GAP-43 (growth-associated protein 43) and MARCKS (myristoylated alanine-rich C kinase substrate), participating in the regulation of subplasmalemmal actin dynamics both in non-neuronal cells (Laux et al, 2000; Wiederkehr et al, 1997) and in axonal growth cones (Frey et al, 2000). In neural tissue they are implicated in neurite outgrowth and synaptic plasticity. Upregulation of BASP1 is also observed in neurons during nerve regeneration (Mosevitsky et al., 1994, 1997; Iino and Maekawa, 1999; Iino et al., 1999), corroborating its essential role in growth cone guidance and actin cytoskeleton organization (Frey et al, 2000; Korshunova et al, 2008).

BASP1, GAP-43 and MARCKS have several physico-chemical and functional properties in common. All three proteins belong to the class of intrinsically disordered proteins characterized by a high content of charged and polar residues as well as a lack of hydrophobic residues (Geist et al, 2012; Kumar et al, 2013; Mancek-Keber et al, 2012). Further they are all subject to N-terminal fatty acylation, facilitating attachment to the plasma membrane, where they are able to sequester the acidic/anionic phospholipid PIP<sub>2</sub> (phosphatidylinositol-4,5-bisphosphate) into lipid rafts. The interaction and sequestration of PIP<sub>2</sub> is mediated by a basic effector domain enriched in lysine residues and constructs lacking the ED lead to inhibition of plasmalemmal PIP<sub>2</sub> modulation. Additionally, the ED is involved in the binding to calmodulin and actin filaments (Laux et al, 2000). These diverse interactions are modulated by Protein Kinase C (PKC)-mediated phosphorylation of conserved serine residues within the ED of these proteins (Mosevitsky, 2005). It has been shown that phosphorylation within the MARCKS-ED leads to the translocation of MARCKS from the plasma membrane to the cytosol by influencing its electrostatic interaction with anionic phospholipids (Ellena et al, 2003; Seykora et al, 1996).

The lipid second messenger PIP<sub>2</sub> is a critical component in promoting filament assembly at the plasma membrane, regulating actin binding as well as focal contact proteins such as gelsolin, profilin, cofilin, and vinculin (Hartwig et al, 1995; Ma et al, 1998; Welch et al, 1997). It has been noted that the overall content of PIP<sub>2</sub> at the inner leaflet of the cell membrane can be comparatively high and constant. Therefore, it was suggested that mechanisms locally concentrating and masking PIP<sub>2</sub> are important for subsequent ligand interaction with demasked concentrated clusters of PIP<sub>2</sub> and downstream signaling to modulate actin cytoskeleton dynamics in response to extracellular signals (Laux et al, 2000). It was noted, that these mechanisms would provide first steps in the formation of signaling

platforms for the concomitant assembly of signaling structures such as focal contacts, adherens junctions, caps and supramolecular activation clusters involved in lymphocyte activation, and synapses in the nervous system (Penninger & Crabtree, 1999).

Interestingly, BASP1 is able to bind  $\text{Ca}^{2+}$  bound calmodulin ( $\text{Ca}^{2+}$ -CaM) despite being devoid of any canonical calmodulin-binding domain (Maekawa et al, 1994). The interaction with  $\text{Ca}^{2+}$ -CaM is mediated by the N-terminal myristoyl moiety together with BASP1 residues 1 to 9 (myr-G<sub>1</sub>GKLSK<sub>9</sub>) (Takasaki et al, 1999). In this way,  $\text{Ca}^{2+}$ -CaM binding to BASP1 disrupts the interaction between the basic ED and PIP<sub>2</sub>. Another function of BASP1-CaM complex formation is the regulation of phosphorylation of BASP1 Ser5 by PKC (Maekawa et al, 1994), although the purpose and implications of BASP1 phosphorylation are still not fully understood.

BASP1 is also expressed in several non-neuronal tissues, including kidney, testis and lymphoid tissues (Carpenter et al, 2004; Mosevitsky et al, 1997) probably related to its role as transcriptional suppressor, but it is particularly abundant in neurons during brain development (Mosevitsky, 2005). It has been shown that overexpression of BASP1 induces neurite outgrowth in neuronal cell lines, where it appears to have overlapping, but not similar, function with GAP-43. Interestingly, this study observed that mutation at the serine-5 position, which is subject to PKC-mediated phosphorylation, had no adverse effect on BASP1 function, whereas the myristoyl moiety of BASP1 is indispensable for the induction of neurite outgrowth (Korshunova et al, 2008).

Furthermore, it has recently been shown that liposomes containing anionic phospholipids induce oligomerization of BASP1, as well as GAP-43. Oligomerization of BASP1 requires its N-terminal myristoylation and interaction with calmodulin leads to the disruption of oligomers and dissociation of BASP1 from the membrane (Zakharov & Mosevitsky, 2010). Additionally, it has been hypothesized that PKC-mediated phosphorylation of Ser5 leads to the disruption of the interaction of the N-terminal positive effector domain of BASP1 (residues 1-9) with anionic phospholipids, which also might interfere with oligomer formation (Laux et al, 2000).

As mentioned before, PIP<sub>2</sub> is implicated in the attachment of the cytoskeleton to the plasma membrane as well as in the regulation of actin dynamics (Golub et al, 2004; Janmey & Lindberg, 2004; Sakisaka et al, 1997; Shibasaki et al, 1997). It was suggested that *in vivo* oligomerization of BASP1, as well as GAP-43, would provide a mechanism to actively sequester and locally concentrate PIP<sub>2</sub> into functional clusters. Subsequent dissociation of



the oligomers and consequent demasking of those clusters induced by CaM-binding as well as phosphorylation by PKC will make those clusters accessible for other PIP<sub>2</sub>-binding proteins (Zakharov & Mosevitsky, 2010). Interestingly, BASP1 was found in a recent study to interact with the PIP<sub>2</sub> phosphatase synaptojanin-1 (SJ-1) (Takaichi et al, 2012). SJ-1 is a presynaptic protein that is implicated in the endocytosis of synaptic vesicles (Perera et al, 2006). After endocytosis, SJ-1 hydrolyses PIP<sub>2</sub>, which leads to the release of PIP<sub>2</sub>-bound proteins and the detachment of the clathrin coat (Royle & Lagnado, 2010; Shupliakov & Brodin, 2010). BASP1 was shown to specifically inhibit SJ-1 by masking of PIP<sub>2</sub> via binding by the N-terminal part of BASP1 enriched in basic residues (residues 6-19 + residues 22-24). It is hypothesized that the influx of Ca<sup>2+</sup> at the presynaptic region and the concomitant activation of PKC and/or CaM leads to the release of BASP1 from PIP<sub>2</sub>. Subsequent binding of the receptor adaptor protein (AP<sub>2</sub>) to PIP<sub>2</sub> then could trigger the formation of the clathrin coat. After clathrin-dependent endocytosis, endophilin mediated recruitment of SJ-1 would lead to the dephosphorylation of PIP<sub>2</sub> and disassembly of the clathrin coat (Takaichi et al, 2012).

A recent study discovered ion channel activity of BASP1 in negatively-charged planar lipid bilayers, which is related to its ability to form oligomers. The observed channels were voltage-independent and cation-selective, with a K<sup>+</sup> to Cl<sup>-</sup> permeability ratio of approximately 14. Interestingly, the neuronal protein GAP-43, which also has been shown to form oligomers, elicited no ion channel currents under the conditions used in the study. BASP1 channels appear to be similar to amyloid protein channels. Although BASP1 is not known to be associated with any amyloid disease, it was suggested that the investigation of the physiological or pathological roles of ion channel activity of BASP1 oligomers would give insight into the mechanism of amyloid protein toxicity (Ostroumova et al, 2011). Soluble oligomers of amyloid proteins are now considered to be the main neurotoxic species in neurodegenerative diseases (e.g. Alzheimer's, Parkinson's, Huntington's diseases, etc.). Interestingly, many amyloid proteins, such as amyloid  $\beta$ -protein, islet amyloid polypeptide, polyglutamine and  $\alpha$ -synuclein, have been shown to form ion channels in lipid membranes (Hirakura et al, 2000; Quist et al, 2005; Zakharov et al, 2007). One way of exerting their toxicity is by provoking membrane permeabilization, which perturbs cellular calcium ion homeostasis (Glabe, 2006; Kagan et al, 2004). Nevertheless, it remains elusive, if ion channel activity of BASP1 oligomers is of functional relevance *in vivo*.

Summing up these findings, a main function of BASP1 is the modulation of the PIP<sub>2</sub> concentration and accessibility in the inner leaflet of the plasma membrane in early steps of

signaling platform formation, eventually regulating actin cytoskeleton dynamics in processes as neurite outgrowth in the developing brain, neuron regeneration as well as presynaptic clathrin-mediated endocytosis.

## **3.2 BASP1 and its role in transcription regulation**

BASP1 was discovered as transcriptional cosuppressor of the potent transcription factor Wilms' tumor suppressor protein 1 (WT1), which regulates target genes including those for growth factors and regulators of cell division and plays a key role in kidney development. Mutations in WT1 as well as aberrant WT1 expression have been associated with Wilms' tumor, a pediatric malignancy of the kidneys. The N-terminal basic effector domain together with a downstream basic cluster are not only serving as epitopes for PIP<sub>2</sub> binding, but also constitute a highly conserved bipartite nuclear localization sequence (NLS). Accordingly, myristoylated BASP1 has been shown to be transported to the nucleus and localizes to the promoters of several WT1 target genes (Carpenter et al, 2004; Essafi et al, 2011; Goodfellow et al, 2011; Green et al, 2009).

BASP1 interacts with an N-terminal suppression region of WT1 (WT1 residues 92-101), which regulates the transcriptional activation domain of WT1 converting it from a transcriptional activator to a repressor. Conversely, elimination of endogenous BASP1 expression enhances transcriptional activation by WT1 (Carpenter et al, 2004). Surprisingly, it has been recently discovered that the N-terminal myristoylation of BASP1 is required for transcriptional repression by WT1-BASP1 (Toska et al, 2012). Further, it has been established that myristoylation as well as the interaction with nuclear PIP<sub>2</sub> are necessary for the recruitment of histone deacetylase 1 (HDAC1) to WT1 target gene promoters. The cosuppressor activity of BASP1 is therefore explained by its ability to tether chromatin-remodeling enzymes to target promoters, where H3K9 deacetylation by HDAC1 leads to a repressive chromatin state. BASP1 mutants, either lacking the N-terminal myristoyl-moiety (BASP1 G1A) or unable to interact with PIP<sub>2</sub> (BASP1 S5D), showed an enhanced nuclear localization as well as an increase in promoter localization, but were both defective in transcriptional repression compared with wild-type BASP1 (Toska et al, 2012). As the BASP1 mutant S5D is deficient in recruiting HDAC1 and hence is not able to act as cosuppressor, it might be possible that PKC-mediated phosphorylation of serine 5 in wildtype

BASP1 plays a major role in transcriptional regulation by inhibiting BASP1's cosuppressor activity.

Another study supporting BASP1's important function in transcriptional control showed that in HeLa cells BASP1 is cleaved and transported into the cytoplasm after induction of apoptosis in a caspase-dependent manner (Ohsawa et al, 2008). These findings, together with the identified cosuppressor activity of BASP1 on WT1 function, raise the possibility that apoptosis induced changes of subcellular localization of BASP1 are a prerequisite for apoptosis-promoting changes in the transcriptional machinery.



# ***Chapter 4***

## ***NMR analysis of BASP1's structural propensities and pH-induced compaction***

### **4.1 Secondary Structure Propensities in IDPs**

Since proteins cannot be strictly separated in either structured or unstructured and the realization that even highly disordered proteins are able to transiently adopt folded conformations in their native state, upon interaction or as a response to environmental changes, IDPs are rather evaluated in terms of structural propensities, which indicates the tendency of a residue or polypeptide stretch to adopt transient structure. This tendency to form certain secondary structure is of course strongly influenced by the environment. One could envisage a highly dynamic interchanging conformational ensemble that samples a vast conformational space including folded conformations, which might be stabilized by certain interactions by conformational selection.

A common method for assessing structural propensities in IDPs is secondary chemical shift analysis, where the measured NMR chemical shifts of the protein backbone atoms are compared to tabulated NMR random coil chemical shifts. The chemical shifts of backbone atoms depend on the dihedral angles  $\phi$  and  $\psi$  and are thus very sensitive on the type of secondary structure they are incorporated in (Spera & Bax, 1991). For folded proteins deviations from the random coil value are quite substantial and are used to identify secondary structural elements by calculating the chemical shift index (CSI) (Wishart & Sykes, 1994). As the CSI is intended to identify fully formed secondary structure, it is not appropriate for

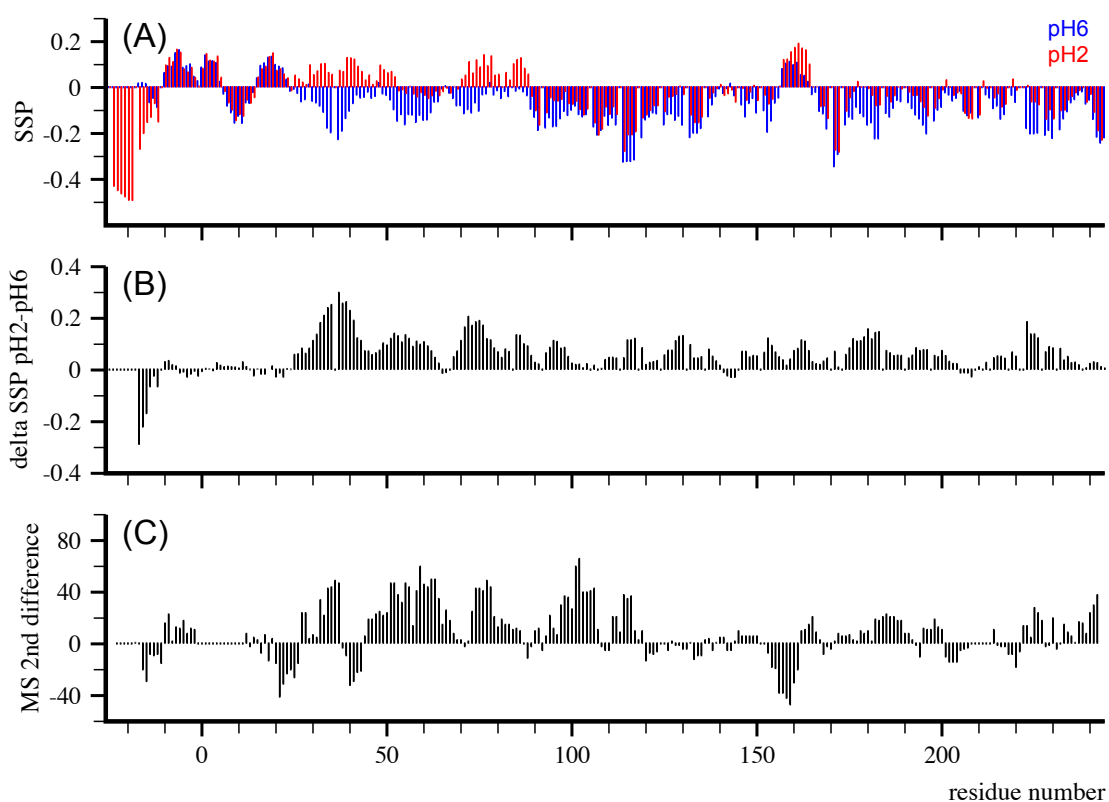
secondary structure evaluation in IDPs. In IDPs secondary structural elements are formed only transiently and therefore chemical shift deviations from the random coil value are only marginal. Secondary chemical shifts,  $\Delta\delta$ , are used to identify the location and the population of transiently structured regions. These regions are of biological interest, given that they are often involved in molecular recognition events (Mohan et al, 2006).

As minor deviations from the random coil value are detected, the analysis is strongly dependent on the accuracy of the random coil chemical shift library. However, it is still unclear how the appropriate values should be defined. A variety of different methods by different groups have been proposed to extract random coil values, which are either based on the evaluation of chemical shifts from a protein chemical shift database (Tamiola et al, 2010; Zhang et al, 2003), or the measurement of chemical shift values from a set of unstructured model peptides (Kjaergaard et al, 2011; Schwarzingner et al, 2001; Schwarzingner et al, 2000). For example, Mulder and coworkers used a set of 14 IDPs of unrelated sequence and function to extract 20 amino acid-specific random coil chemical shifts. Realizing that neighboring residues contribute significantly to the backbone chemical shifts, they provided 40 sequence-dependent left- and right-neighbor correction factors (Tamiola 2010).

The group of Julie Forman-Kay devised a method that combines different backbone atom chemical shifts ( $C^\beta$ ,  $C^\alpha$ ,  $C^O$ , N,  $H^N$ ,  $H^\alpha$ ) into a single residue-specific secondary structure propensity (SSP) score by weighting the contribution of different chemical shifts by their sensitivity to  $\alpha$ - or  $\beta$ -structure. The SSP score represents the residual tendency to form  $\alpha$ - or  $\beta$ -secondary structure at a given residue and scales from -1 to +1, the extremities representing fully formed  $\beta$ - or  $\alpha$ -structure, respectively. This method allowed them to assess significant differences in secondary structure propensities between two homologues of the synuclein family,  $\alpha$ -synuclein and  $\gamma$ -synuclein, associated with Parkinson's disease (Bennett, 2005) and late-stage breast cancers (Liu et al, 2005), respectively. Despite high sequence-similarity in the aggregation prone region of the two synucleins, the differences in secondary structure propensities explained well the observed differences in the tendency to aggregate, which is high in  $\alpha$ -synuclein und low in  $\gamma$ -synuclein (Marsh et al, 2006).

We used the SSP approach to assess changes in transient secondary structure formation in H6-cBASP1 upon pH reduction (Article 1). There are some marked differences between the SSP profiles of H6-cBASP1 at pH6 and pH2 (**FIG. 3**, A and B). Overall H6-cBASP1 seems to adopt rather extended conformations at both pHs, with only 3 small areas showing helical SSP scores at pH6: His-tag residues -10-0, harboring the TEV-site, and the

first 5 residues of cBASP1; cBASP1 residues 15-23 and cBASP1 residues 157-164. It has to be noted that SSP values are in general very low ( $< 0.2$ ), indicating that there are no rigid elements of secondary structure formed along the protein at both pHs. Rather, it appears that the conformational ensemble of BASP1, constituting preferentially extended conformations, has a tendency to populate helical conformations in the aforementioned regions. Interestingly, pH reduction leads to a global increase in SSP scores highlighting a shift in the conformational ensemble towards  $\alpha$ -helical structures, with strongest increases spanning residues 25-88. Residues 25-52 and 70-88 even change in their propensities from an extended to a helical conformation. The regions spanning the TEV-site of the His-tag as well as cBASP1 residues 1-23 seem not to be affected by the pH decrease concerning to their tendency to form secondary structure. Important to notice is that Meta-Structure predictions of secondary structural changes induced by pH reduction (**FIG. 3, C**) are in good agreement with NMR-derived SSP score differences, highlighting the predictive power of this method based solely on primary structure information.



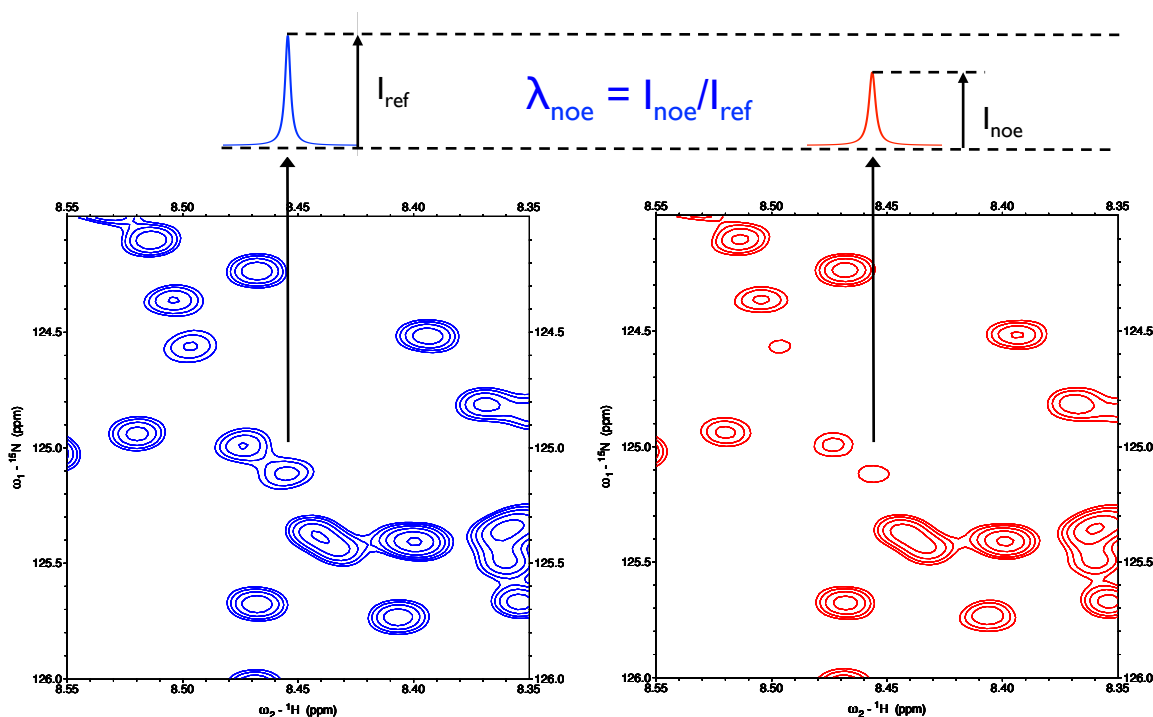
**Fig. 3** Comparison between NMR derived Secondary Structure Propensity (SSP) changes and Meta-Structured derived secondary structure changes. (A) SSP scores for H6-cBASP1 at pH6 (blue) and pH2 (red). (B) Difference in SSP scores (pH2-pH6). (C) Difference in Meta-Structure derived secondary structure propensities.

## 4.2 HET-SOFAST spectroscopy

The HETerogeneity-SOFAST experiment ( $^1\text{H}$ - $^{15}\text{N}$  HET-SOFAST-HMQC) is based on the band-Selective Optimized-Flip-Angle Short-Transient (SOFAST)-HMQC experiment, designed for the fast acquisition of heteronuclear correlation spectra (Schanda & Brutscher, 2005; Schanda et al, 2005). The HET extension to the classical SOFAST-HMQC was designed to assess global as well as site-specific heterogeneity along the protein backbone. Two topological parameters – compaction and solvent accessibility – provide complementary information about a protein and can be used to quickly deduce the degree of structure in a new protein sample. Especially if backbone amides are assigned, residue-specific information will give valuable information about flexible as well as folded regions along the polypeptide chain. Still, even for unassigned proteins, the distribution of the magnitudes of the parameters will already provide interesting insights into the overall topology of a protein.

After perturbing either the water  $^1\text{H}$  or aliphatic  $^1\text{H}$  spin polarization, the  $^1\text{H}$ - $^{15}\text{N}$  spectrum is differentially affected, providing measures of the solvent accessibility ( $\lambda_{\text{ex}}$ ) and proton density ( $\lambda_{\text{noe}}$ ) at the amide proton site. Perturbation is accomplished by applying a band-selective proton inversion pulse at the beginning of a delay  $t_c$  (contact time) before the actual SOFAST pulse sequence. This perturbation will affect the spin-polarization of the amide protons after the contact time  $t_c$ . After recording two data sets, with or without the inversion pulse, the peak intensity ratio ( $\lambda = I^{\text{sat}}/I^{\text{ref}}$ ) provides insight into the interaction between amide protons and saturated protons. **FIG. 4** shows an example of a typical set of  $^1\text{H}$ - $^{15}\text{N}$  HET<sup>noe</sup>-SOFAST-HMQC spectra recorded on a 0.5 mM sample of H6-cBASP1 at pH2, illustrating how the initial aliphatic proton specific inversion pulse affects different resonances.





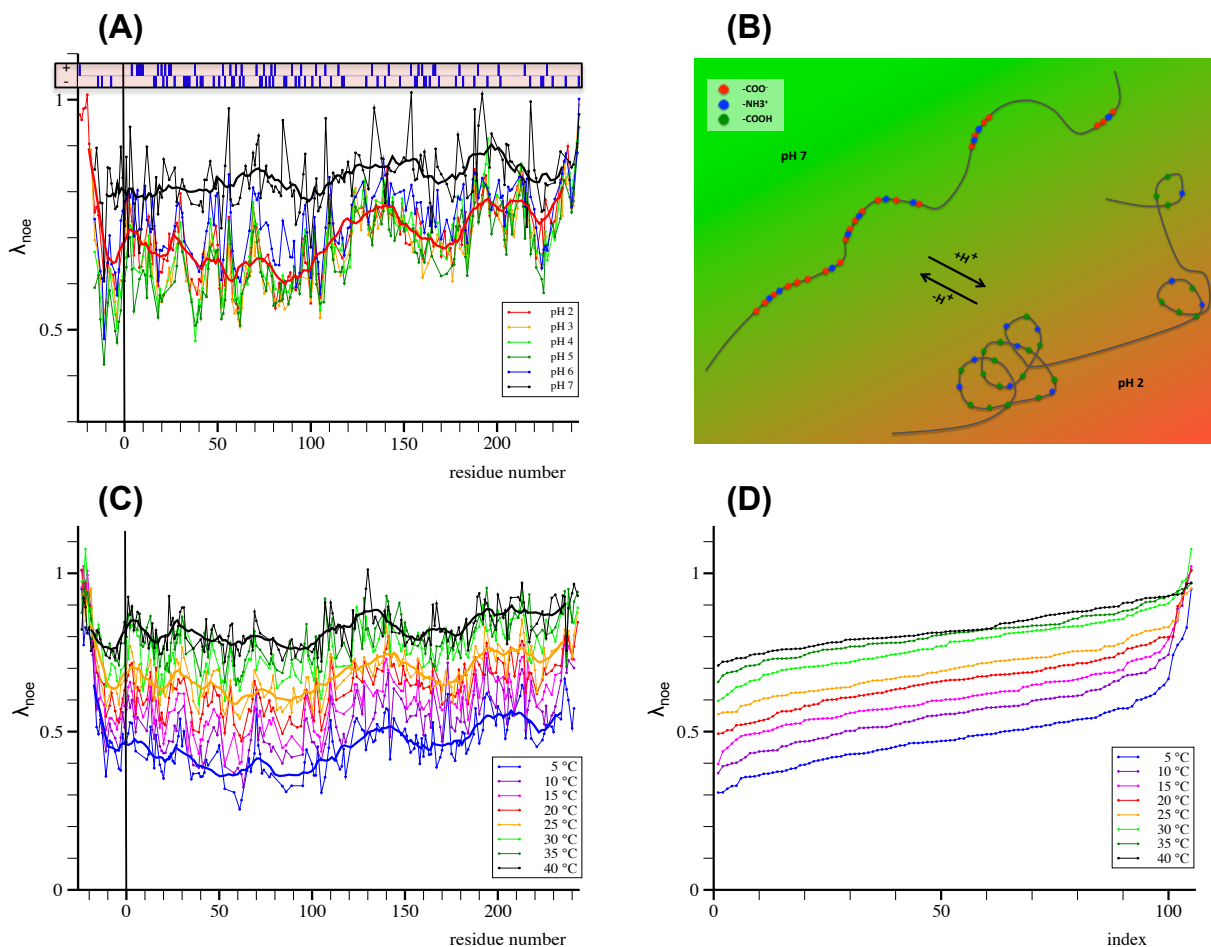
**Fig. 4.** Set of sections of HET<sup>noe</sup>-SOFAST HMQC spectra recorded without (reference) (left spectrum) and with the aliphatic inversion pulse (right spectrum). Spectra were recorded on a 0.5mM sample of H6-cBASP1 at pH2 and at 5°C. This region of the spectrum shows unaffected resonances as well as resonances strongly reduced in intensity after aliphatic proton specific inversion. The intensity ratio  $\lambda_{\text{noe}}$  provides insight into the degree of spin-diffusion at the respective backbone amide.

The two observables  $\lambda_{\text{noe}}$  and  $\lambda_{\text{ex}}$  yield complementary information on the structure and mobility of the protein.  $\lambda_{\text{noe}}$  values tell us about the proton density around the amide proton via the  $^1\text{H}$ - $^1\text{H}$  spin-diffusion or NOE-effect, whereas  $\lambda_{\text{ex}}$  values are indicative of the amide-solvent exchange, i.e. solvent accessibility. If an amide proton is embedded in a compact structure, spin-diffusion effects will dominate, whereas the same proton won't be solvent accessible ( $\lambda_{\text{noe}} < 1$ ,  $\lambda_{\text{ex}} \approx 1$ ). An amide proton in a flexible, solvent-exposed part of a protein will exchange with the solvent water, but will be less affected by spin-diffusion ( $\lambda_{\text{noe}} \approx 1$ ,  $\lambda_{\text{ex}} < 1$ ) (Schanda et al, 2006).

### 4.2.1 $HET^{noe}$ -SOFAST:

In the reference experiment, which is a conventional SOFAST experiment, the application of an amide  $^1H$  selective pulse excites only a few  $^1H$  spins. Energy transfer via spin diffusion ( $^1H$ - $^1H$  NOE) within the dipolar-coupled network constituted by the protein provides the main mechanism of spin-lattice relaxation and energy is released at sites of high mobility (e.g. methyl groups). In the saturation experiment the efficiency of this energy dissipation is reduced, as the aliphatic  $^1H$  specific inversion pulse at the beginning of the contact time  $t_c$  puts most of the proton spins out of their equilibrium. Hence,  $HET^{noe}$ -SOFAST NMR compares effective spin-lattice relaxation between the reference and the saturation experiment.

$\lambda_{noe}$  values of H6-cBASP1 are very high, as expected for a highly unfolded protein, and lie in the range of 0.5 to 0.9. Interestingly, comparison of the  $\lambda_{noe}$  profiles of H6-cBASP1 at different pHs shows that pH reduction differentially affects distinct parts of the polypeptide (Fig. 5A). While at pH7  $\lambda_{noe}$  values are rather high and uniformly distributed along the polypeptide chain ( $\lambda_{noe} \approx 0.8$ ), indicating that the protein samples only highly extended conformations, pH reduction appears to induce transient compaction in different parts of the protein. Especially the N-terminal half of BASP1, which is highly enriched in acidic residues (see Fig. 5A top, delineating the charge distribution at pH7), is marked by a stronger decrease in  $\lambda_{noe}$  values. This supports the idea of enhanced compaction of this region at low pH induced by the protonation of glutamic acid and aspartic acid side chains (Fig. 5A, B). As the  $^1H$ - $^1H$  spin diffusion effect and thus the  $\lambda_{noe}$  value depend on the effective tumbling correlation time ( $\tau_c$ ) we measured  $\lambda_{noe}$  values at pH2 at different temperatures. For a folded protein a spread of  $\lambda_{noe}$  values of about 0.1 has been observed with temperatures ranging from 15-45°C, corresponding to  $\tau_c$  values between 3-7ns (Schanda et al, 2006). For BASP1 at pH2 the observed spread in  $\lambda_{noe}$  values approaches 0.4 for temperatures ranging from 5-40°C (Fig. 5C, D). The more pronounced spread in  $\lambda_{noe}$  values is due to a stronger effect of the temperature on the local correlation times and internal mobility of the backbone of an IDP. At lower temperatures the differential compaction of the N-terminal and C-terminal half of BASP1 at pH2 is much more pronounced than at higher temperatures. Increasing the temperature appears to lead to a melting of the compaction-prone parts of BASP1 or to a slight shift towards less compact conformers, respectively.



**Fig. 5.** HET<sup>noe</sup>-SOFAST measurements on H6-cBASP1 ( $t_c = 300\text{ms}$ ). Intensity-ratios ( $\lambda_{\text{noe}}$ ) as a function of the residue at different pHs (A) and a schematic depiction of protonation-induced compaction of the IDP BASP1 (B). Experimental evaluation of the influence of the temperature on the measured  $\lambda_{\text{noe}}$  values as a function of the residue (C) or sorted for increasing  $\lambda_{\text{noe}}$  value (D). The inlay at the top of the graph in (A) shows the charge distribution of BASP1 at pH7. Residues -25 to 0 are His-tag residues.

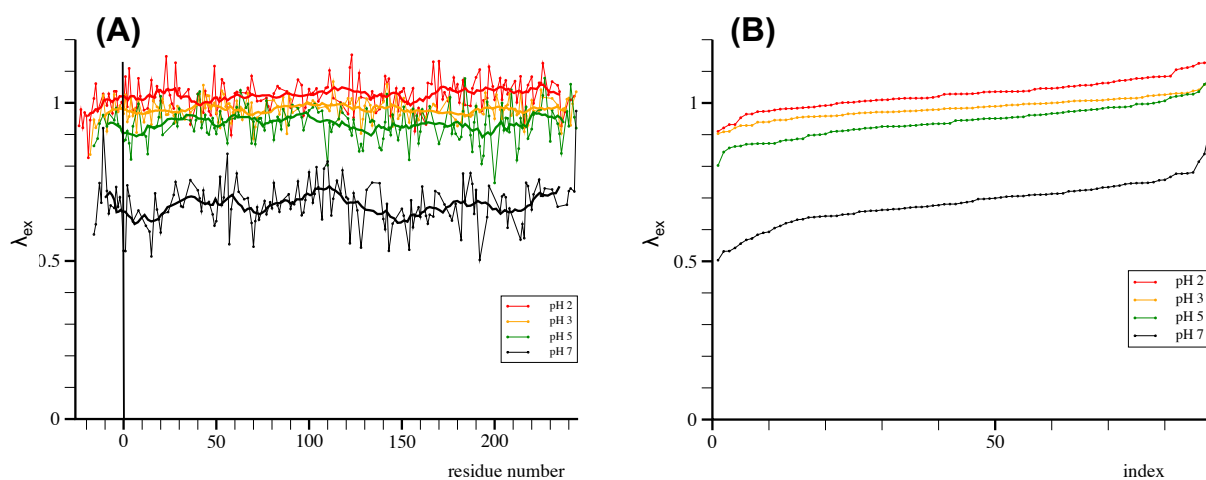
#### 4.2.2 HET<sup>ex</sup>-SOFAST:

The hydrogen atoms of amide groups or other polar groups in a protein are in continuous exchange with solvent hydrogens, unless they are protected through hydrogen-bonding. Therefore, measurement of hydrogen exchange allows the evaluation of protected, folded parts or exchanging, flexible parts of a protein, respectively. Due to differential protection in a protein, exchange times can range from milliseconds for highly disordered parts to days or months in stably folded regions of the protein.

In a HET<sup>ex</sup>-SOFAST experiment the water resonance is selectively perturbed and consequently affects the spin-polarization of exchangeable amide hydrogens during the

contact time  $t_c$ . As the contact time is in the order of a few hundred milliseconds HET<sup>ex</sup>-SOFASST spectroscopy probes fast hydrogen exchange (sub-second time scale). Accordingly, it has been suggested that the experiment is most suitable for samples at neutral or basic pH. As the exchange rates are strongly dependent on the solvent conditions (pH, temperature, ...), care has to be taken when comparing different protein samples (Schanda et al, 2006).

As anticipated for an IDP, given that amide protons in disordered parts of the polypeptide are not protected through hydrogen bonding, residue-based  $\lambda_{ex}$  values for H6-cBASP1 are quite homogeneously distributed along the protein backbone (Fig. 6A, B). At neutral pH  $\lambda_{ex}$  values are distributed around 0.7 with no obvious signs for compact regions. Due to the quench of amide hydrogen-solvent exchange with decreasing pH, where amide proton residence times are getting too long for quantitative measurement of exchange with HET<sup>ex</sup>-SOFASST spectroscopy,  $\lambda_{ex}$  values are increasing and are distributed around  $\lambda_{ex}=1$  at pH2.



**Fig. 6.** HET<sup>ex</sup>-SOFASST measurements on H6-cBASP1 ( $t_c = 300$ ms). Intensity-ratios ( $\lambda_{ex}$ ) as a function of the residue at different pHs (A) or sorted for increasing  $\lambda_{ex}$  values (B). Residues -25 to 0 are His-tag residues.

## 4.3 $^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC

### 4.3.1 *The Nuclear Overhauser Effect (NOE)*

The NOE (Nuclear Overhauser Effect) describes the transfer of magnetization via longitudinal cross-relaxation between NMR-active nuclei through space and allows the correlation of nuclei up to a distance of  $\sim 5\text{\AA}$ . This relaxation mechanism is also referred to as dipole-dipole relaxation, as it involves neighboring dipoles. Cross-peak intensity can be directly translated into distance information and therefore is used to get distance restraints for structure calculation. A simple 1D NOE experiment involves an initial selective  $180^\circ$  pulse followed by a mixing time  $\tau_m$  and a  $90^\circ$  read-out pulse. During the mixing time ( $\tau_m = 100\text{--}500$  ms, depending on the molecular weight of the molecule of interest) magnetization is transferred from the selectively inverted spin to neighboring spins close in space. NOESY experiments are usually set up as difference experiments, where the reference experiment (without the initial  $180^\circ$  pulse) is subtracted from the NOE experiment, which finally leaves only the signal from the source spin and NOE-enhanced signals in the final spectrum.

The 2D-NOESY experiment is a straightforward extension of the 1D version, incorporating an initial  $90^\circ$  pulse, an evolution time delay  $t_1$  – providing  $^1\text{H}$  chemical shift information – and a second  $90^\circ$  pulse that brings the transverse magnetization that is orthogonal to the pulse direction along  $z$ , followed by the aforementioned NOESY mixing time and the final read-out pulse.

The intensity of a cross peak is proportional to the cross-relaxation rate  $\sigma_{\text{NOE}}$  between two dipolar-coupled spins, which is given by the difference between the double quantum and zero quantum transition rate constants  $W_2$  and  $W_0$ :

$$\sigma_{\text{NOE}} = W_2 - W_0 = b^2 \left[ \frac{3}{10} j(2\omega_0) - \frac{1}{20} j(0) \right]$$
$$b = \frac{\mu_0 \gamma_H^2 \hbar}{4\pi r^3}$$

with vacuum permeability  $\mu_0$ ,  $^1\text{H}$  gyromagnetic ratio  $\gamma_H$ , reduced Planck constant  $\hbar$ ,  $^1\text{H}$  Larmor frequency  $\omega_0$ , and the distance between two spins  $r$ . The reduced spectral density function  $j(\omega)$  is the Fourier transform of the reduced correlation function  $g(\tau)$ , which is independent of the size of the local fields and is one fifth of the spectral density  $J(\omega)$  as

described by the Lipari-Szabo model free formalism (Keeler, 2007; Lipari & Szabo, 1982a; Lipari & Szabo, 1982b).

$$J(\omega) = \frac{1}{5} \frac{2\tau_c}{1 + \omega^2\tau_c^2} = \frac{1}{5} j(\omega)$$

It follows that for **fast tumbling molecules** (extreme narrowing limit:  $\omega_0\tau_c \ll 1$ ) the cross-relaxation rate is positive, hence cross peaks are negative:

$$j(2\omega_0) = j(0) = 2\tau_c$$

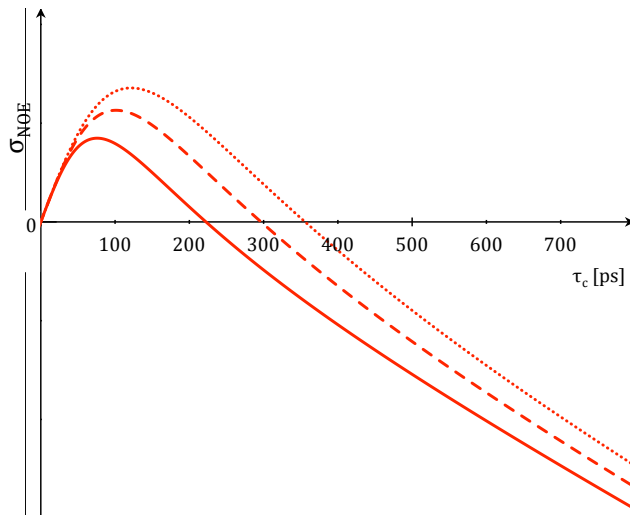
$$\sigma_{NOE} = b^2 \frac{1}{2} \tau_c$$

For **slowly tumbling molecules** (spin diffusion limit:  $\omega_0\tau_c \gg 1$ )  $j(2\omega_0)$  is approaching zero, therefore  $\sigma_{NOE}$  is negative and cross peaks are positive:

$$j(2\omega_0) = 0$$

$$\sigma_{NOE} = -b^2 \frac{1}{10} \tau_c$$

Consequently,  $\sigma_{NOE}$  is zero for  $\omega_0\tau_c = 1.12$ . Fig. 7 depicts the behavior of the longitudinal cross-relaxation rate  $\sigma_{NOE}$  of protons at different Larmor frequencies as a function of the correlation time  $\tau_c$ .



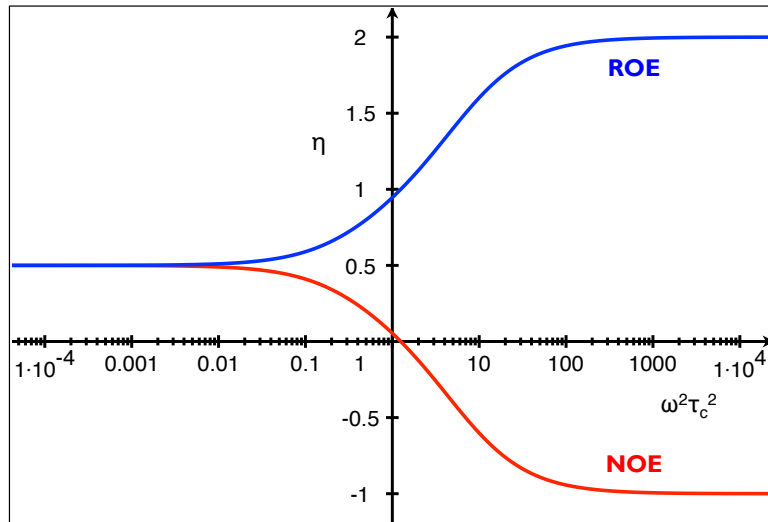
**Fig. 7.** Illustration of the behavior of the longitudinal cross-relaxation rate  $\sigma_{NOE}$  as a function of the correlation time  $\tau_c$  at different magnetic fields (500 MHz: dotted line, 600 MHz: dashed line, 800 MHz: solid line). Cross-over points are at  $\tau_c \approx 360$ ps (500 MHz),  $\tau_c \approx 300$ ps (600 MHz) and  $\tau_c \approx 220$ ps (800 MHz).

Due to limitations of NOESY spectroscopy, concerning cross-relaxation rates close to zero at certain correlation times, ROESY (rotating-frame nuclear Overhauser effect spectroscopy) was developed, where cross-relaxation is measured in the transverse plane, employing a spin-lock field during the mixing time. The transverse cross-relaxation rate  $\sigma_{ROE}$  is given by

$$\sigma_{ROE} = b^2 \left[ \frac{1}{10} j(0) + \frac{3}{20} j(\omega_0) \right]$$

where  $\omega_0$  depicts the Larmor frequency of both spins. Important to notice here is that the transverse cross-relaxation rate is positive for all values of  $\tau_c$ .

Thus, NOE and ROE enhancements differ in sign for large molecules (spin diffusion limit:  $\omega_0\tau_c \gg 1$ ), with  $\eta_{NOE}$  reaching a maximum of -100% and  $\eta_{ROE}$  reaching a maximum of 200%. In contrast, the NOE and ROE enhancements are the same for small molecules, i.e. small correlation times (extreme narrowing limit:  $\omega_0\tau_c \ll 1$ )  $\eta_{NOE} = \eta_{ROE} = 50\%$  (Fig. 8).

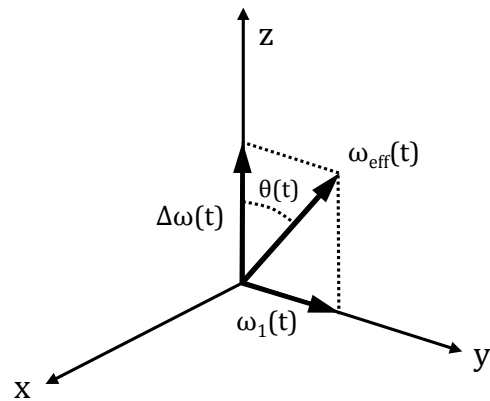


**Fig. 8.** NOE and ROE enhancements  $\eta = \sigma_{NOE/ROE}/R_1$  vs  $\omega^2 \tau_c^2$ .

### 4.3.2 Adiabatic Fast Passage (AFP)

Adiabatic fast passage (AFP) is a highly effective method for spin inversion over a large spectral window and therefore is commonly applied for broadband spin inversion and heteronuclear decoupling. Apart from covering broad bandwidths with relatively low radiofrequency power, AFP has the additional advantage that the inversion efficiency is independent of the intensity of the radiofrequency field, provided that the adiabatic condition is satisfied  $|d\theta/dt| \ll \omega_{\text{eff}}$ , i.e. that the magnetization follows the time-dependent effective field  $\omega_{\text{eff}}$ . Additionally, careful pulse calibration is no longer required and inversion efficiency is insensitive to radiofrequency field inhomogeneities (Auer, 2010; Bohlen & Bodenhausen, 1993; Kupce & Freeman, 2007).  $\omega_{\text{eff}}(t) = \gamma B_{\text{eff}}(t)$  is the effective field expressed in angular frequency units and is given by the vector sum of the time-dependent applied radiofrequency field  $\omega_1(t) = \gamma B_1(t)$  and the resonance offset  $\Delta\omega(t) = \omega_{\text{RF}}(t) - \Omega_I$ , with the Larmor frequency  $\Omega_I$  of spin I (Fig. 9):

**Fig. 9.** Representation of the adiabatic spin-lock frame with the rf field  $\omega_1(t)$ , the offset  $\Delta\omega(t)$ , the effective field  $\omega_{\text{eff}}(t)$  and the angle  $\theta(t)$  between the offset and the effective field. During the time course of the adiabatic fast passage the spin-locked magnetization of spin I follows the effective magnetic field and passes through y exactly at on-resonance (adapted from Konrat & Tollinger, 1999 (Konrat & Tollinger, 1999)).

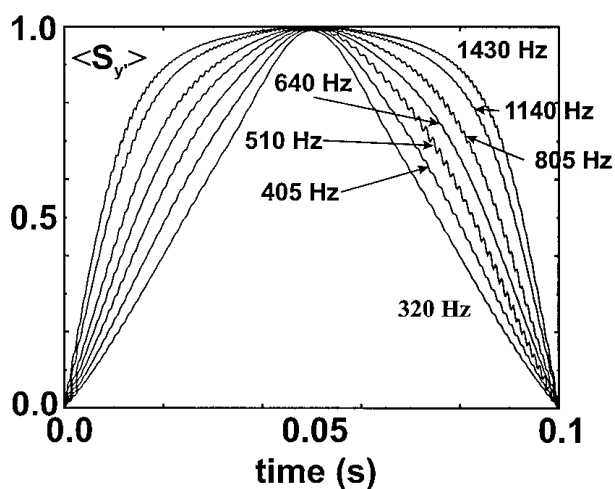


The AFP pulse is a relatively long pulse (hundreds of ms) with a frequency sweep over a defined spectral window and ramps at the beginning and the end of the pulse, which will smoothly decrease the amplitude to zero at the two extremities of the pulse. This is necessary as the effective field  $B_{\text{eff}}$  should start on the  $+z$  axis and end on the  $-z$  axis. A simple constant-amplitude frequency sweep would leave  $B_{\text{eff}}$  inclined with respect to the  $z$  axis as a finite offset  $\Delta\omega$  remains at the extremities of the sweep (Kupce & Freeman, 1995). In order to efficiently invert all the frequencies, the field strength has to be strong enough to effectively spin-lock the magnetization and the sweep rate has to fulfill the adiabatic condition (Auer, 2010; Kupce & Freeman, 1995; Tannus & Garwood, 1997). For the



purposes of broadband spin inversion  $\omega_1(t)$  is relatively weak (the adiabatic frequency sweep is very large compared with the radio-frequency field intensity), with the result that the effective field  $\omega_{\text{eff}}(t)$  will be weakest at on-resonance, hence, a spin with the Larmor frequency  $\Omega_I$  is aligned mostly along the  $+z$  axis, flips quickly when  $\omega_1(t)$  is on-resonance with  $\Omega_I$  ( $\Delta\omega = \omega_{\text{RF}} - \Omega_I = 0$ ) and aligns along the  $-z$  axis as the frequency sweep proceeds. This also means, that relaxation during the adiabatic fast passage is mainly longitudinal.

Interestingly, if the radio-frequency field intensity is not small compared to the frequency sweep range but of comparable strength, a significant amount of transverse magnetization will be created. It has been demonstrated that it is possible to measure transverse ( $R_2$ ) relaxation rates based on measurements of the adiabatic spin-lock relaxation rate  $R_{1Q}$  at different amplitudes of the radiofrequency field  $\omega_1$ . Fig. 10 shows a simulation of the time-evolution of transverse magnetization during adiabatic fast passage at variable adiabaticity (i.e. variable rf spin-lock pulse strength). Pulse parameters as mentioned in the figure are adopted for the measurement of  $^{15}\text{N}$ - $T_{1Q}$  relaxation times (Konrat & Tollinger, 1999).



**Fig. 10.** Time evolution of the transverse magnetization (expectation value  $\langle S_y \rangle$ ) during the adiabatic fast passage. Increasing the rf field amplitude from 320 Hz to 1430 Hz leads to an increase of transverse magnetization during the frequency sweep. In this simulation the center of the sweep coincides with the resonance frequency of spin S. Pulse parameters are as follows: adiabatic sweep duration = 100 ms, sweep range = 2.0 kHz, sweep rate =  $2 \times 10^4 \text{ s}^{-2}$ . Sine-shaped ramps apodize the first and last 30% of the pulse (adapted from Konrat & Tollinger, 1999 (Konrat & Tollinger, 1999)).

### 4.3.3 1D AFP-NOESY

Another application of adiabatic fast passage with strong radio-frequency field intensities is the measurement of homonuclear cross-relaxation under spin-lock conditions. Application of an AFP pulse during the NOESY mixing time of a conventional 1D NOESY

experiment leads to a time-dependent weighted combination of NOE and ROE effects. For weak AFP pulses mostly longitudinal cross-relaxation (NOE) contributes to the effective cross relaxation rate. Increasing the strength of the AFP pulse in successive experiments, will increase the transverse (ROE) contribution to the effective spin-lock cross-relaxation rate  $\sigma_{\text{eff}}$  between two interacting spins  $i$  and  $j$ :

$$\sigma_{\text{eff}}^{i,j} = \sigma_{\text{NOE}} \cos^2 \theta_{\text{eff}} + \sigma_{\text{ROE}} \sin^2 \theta_{\text{eff}}$$

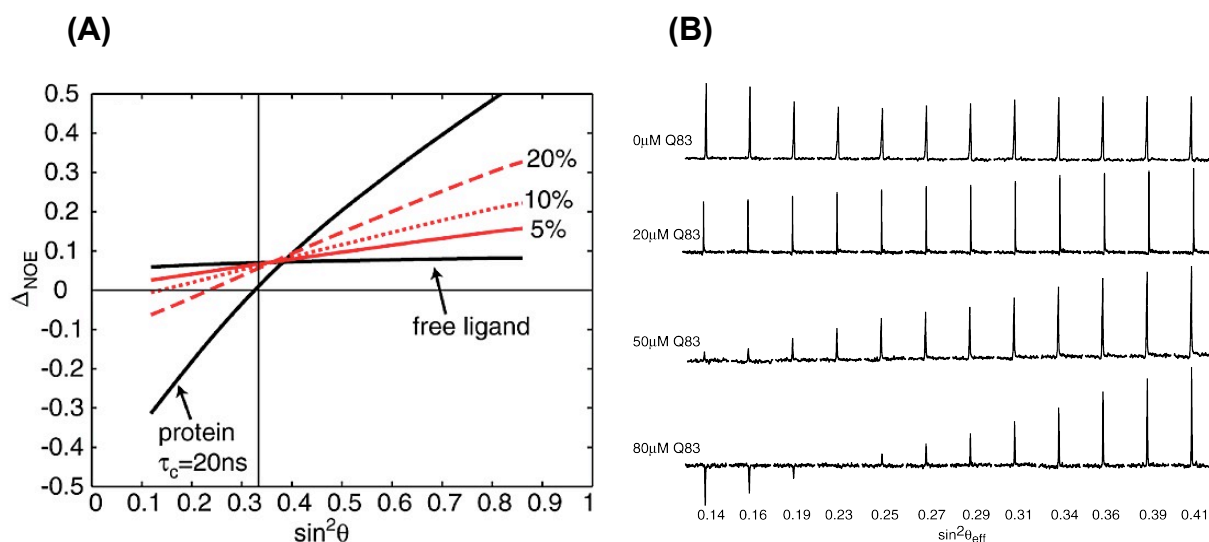
with the longitudinal and transverse cross-relaxation rates  $\sigma_{\text{NOE}}$  and  $\sigma_{\text{ROE}}$  and the effective tilt-angle  $\theta_{\text{eff}}$  of the effective spin-lock field. The effective tilt-angle  $\theta_{\text{eff}}$  can be envisaged by assuming a static spin-lock field with the magnetization remaining at a constant angle  $\theta_{\text{eff}}$ . A spin in such a static field would show the same relaxation behavior as a spin that follows the time-dependent spin-lock field from  $+z$  to  $-z$  (Auer, 2010). As mentioned before, for small molecules NOE and ROE enhancements are the same (+50%). Therefore, the measured spin-lock cross-relaxation rate  $\sigma_{\text{eff}}$  is independent of the radiofrequency amplitude  $\omega_1$ . In contrast, NOE and ROE enhancements are of different sign and strength for large molecules, i.e. proteins (NOE -100%, ROE +200%) (Fig. 8). Consequently,  $\sigma_{\text{eff}}$  will change sign dependent on the applied radiofrequency field. At weak  $\omega_1$  longitudinal cross-relaxation (NOE) will dominate the effective spin-lock cross-relaxation rate, leading to negative enhancements. Whereas at strong  $\omega_1$  the magnetization will stay predominantly in the transverse plane and transverse cross-relaxation (ROE) will lead to positive enhancements. This also means, that there is a characteristic zero crossing of the spin-lock cross-relaxation rate for large molecules (zero passage tilt angle), where the NOE and ROE cross-relaxation rates cancel each other out.

$$\sigma_{\text{eff}}^{i,j} = \sigma_{\text{NOE}} \cos^2 \theta_{\text{eff}} + \sigma_{\text{ROE}} \sin^2 \theta_{\text{eff}} = 0$$

$$\frac{\sigma_{\text{NOE}}}{\sigma_{\text{ROE}}} = -\tan^2 \theta_{\text{eff}} = -\frac{1}{2}$$

$$\theta_{\text{eff}} = \arctan \sqrt{\frac{1}{2}} = 35.26^\circ$$

Measuring cross-relaxation during adiabatic fast passage has proven as a powerful tool for the analysis of protein-ligand interaction. A novel NMR experiment was introduced – 1D AFP-NOESY – that is based on a simple 1D NOESY experiment incorporating an AFP pulse into the NOESY mixing time. This new pulse sequence allows the measurement of homonuclear  $^1\text{H}$ - $^1\text{H}$  cross-relaxation rates during adiabatic fast passage. By measuring the NOE/ROE between ligand protons as a function of the spin-lock rf amplitude and protein concentration, it was possible not only to prove ligand binding but also to map the binding epitope (pharmacophore) of ligands reversibly binding to their target protein. Fig. 11 shows the dependence of the NOE/ROE-enhanced signal on the rf field intensity expected for a small molecule, i.e. a free ligand, or a rigid macromolecule, i.e. a protein, and what happens upon ligand-protein interaction. As mentioned before, the NOE/ROE enhancement ( $\Delta_{\text{NOE}}$ ) of the free ligand is independent of the applied radio-frequency amplitude  $\omega_1$  (extreme narrowing limit) and therefore leads to a flat profile (no  $\sin^2\theta_{\text{eff}}$  dependence). Increasing the concentration of the interacting protein leads to an increase of the effective cross-relaxation rate between ligand protons as well as to a clear dependence of the cross-relaxation rate on the radio-frequency amplitude  $\omega_1$ . Hence the typical profile for macromolecules with a zero-crossing at a tilt angle of  $\theta_{\text{eff}} = 35.26^\circ$  (pronounced  $\sin^2\theta_{\text{eff}}$  dependence) will be observed (Auer, 2010; Auer et al, 2010).

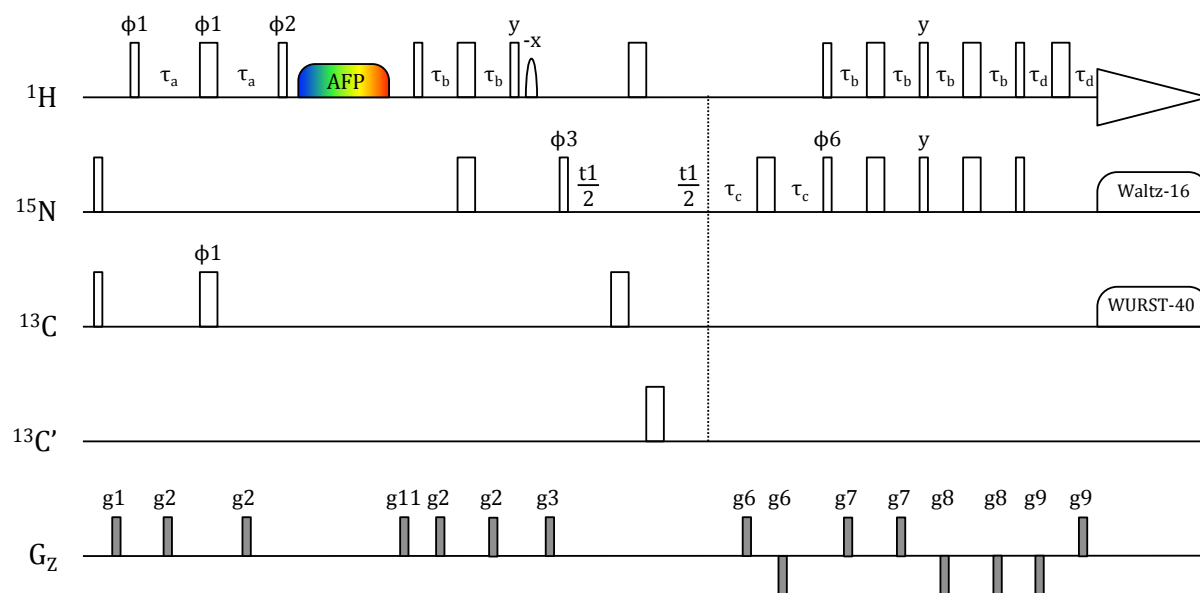


**Fig. 11.** (A) Simulation of the NOE/ROE-enhanced signal intensity changes (due to cross-relaxation from a selectively inverted signal) for a free ligand (small molecule) or a protein (rigid macromolecule) as a function of the effective tilt-angle, which corresponds to rf field intensity. The correlation time of the ligand increases upon binding to the protein, i.e. the free ligand (after dissociation) exhibits protein-like behavior. Simulations are with 5% (solid), 10% (dotted), and 20% (dashed) of free ligand with protein-like behavior. (B) 1D AFP-NOESY spectra monitoring the complex formation between vanillic acid (VA = 4-hydroxy-3-methoxybenzoic acid) and the lipocalin Q83. The NOE/ROE between 3-OCH<sub>3</sub> and H2 of VA is measured at increasing rf field strengths and at different protein concentrations (adapted from Auer et al, 2010 (Auer et al, 2010)).

Additionally, it has been shown that altered tilt-angle profiles (zero-passages deviating from the macromolecular behavior) are a rich source of information about the binding mode of a ligand. Enhanced internal mobility leads to zero crossings at smaller tilt-angles. Spin-diffusion leads to zero passages at larger tilt angles (or no zero passage at all), thus telling us about ligand protons that are embedded into a hydrophobic pocket of the protein. These indirect relaxation pathways via spin-diffusion have been successfully exploited for the investigation of the binding mode of AMP and NAD<sup>+</sup>, respectively, to ADH (Alcohol Dehydrogenase from *S. cerevisiae*). Tilt-angle profiles of the effective cross-relaxation rate for different ligand protons could be successfully used to map parts of the ligands embedded into a hydrophobic pocket of the protein (clear signs of spin diffusion) or exposed to the solvent (Auer et al, 2010).

### 4.3.4 $^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC

The  $^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC pulse sequence (Fig. 12) is based on the classical 3D NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC experiment with the exception that the AFP pulse replaces the NOESY mixing time. Additionally, the initial element recording  $^1\text{H}$  chemical shift evolution is replaced by a  $^{13}\text{C}$ -filter element to restrict NOE/ROE measurements to the dipole-interaction between aliphatic  $^{13}\text{C}$ -attached protons and amide protons. In contrast to a conventional INEPT element, the delay  $\tau_A$  is chosen that  $2\tau_A = 1/J_{\text{CH}}$  and thus leads to a selective inversion of protons bound to  $^{13}\text{C}$ -labelled carbons. Experiments are performed twice, with and without  $J_{\text{CH}}$  scalar coupling evolution. Signals stemming from  $^{13}\text{C}$ -bound protons are selected by proper combination of sub-spectra. All other contributions, amide proton to amide proton as well as solvent to amide protons are thus filtered out.



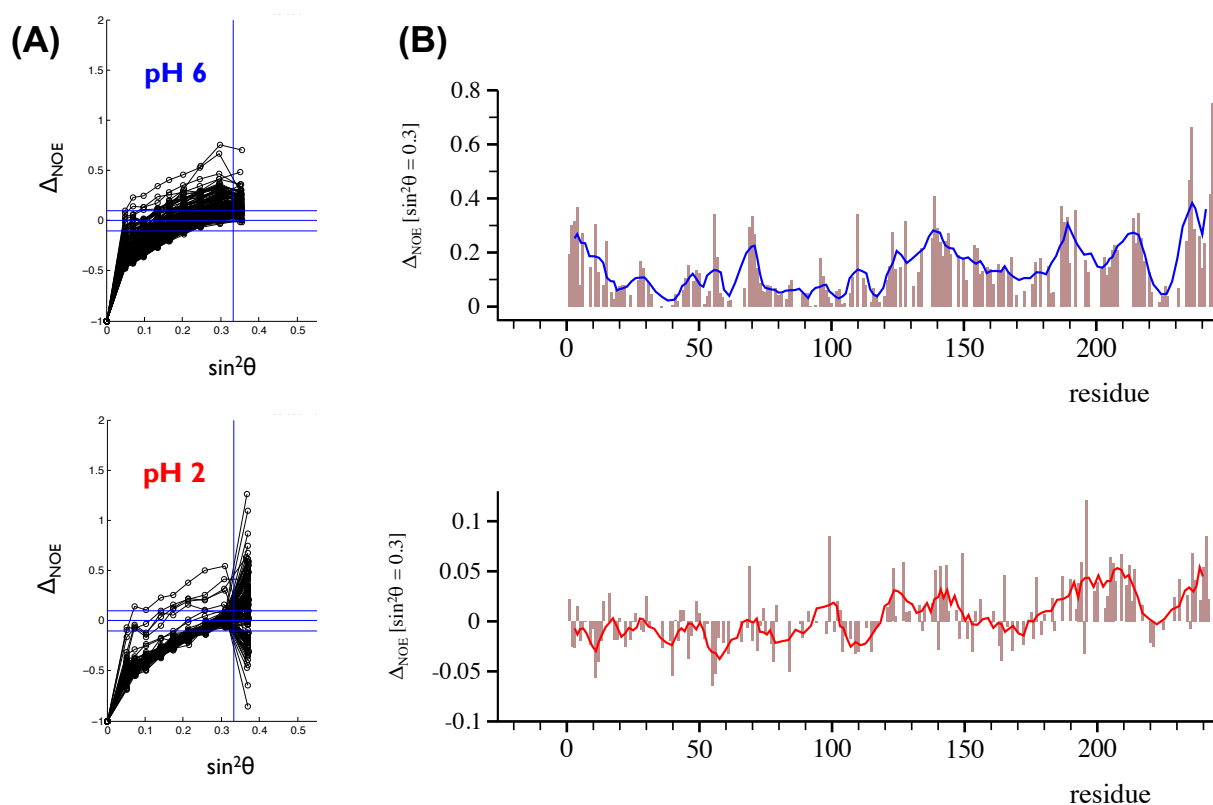
**Fig. 12.**  $^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC pulse scheme for the measurement of homonuclear  $^1\text{H}$ - $^1\text{H}$  cross-relaxation (NOE/ROE) between aliphatic protons and amide protons during adiabatic fast passage. Narrow and wide pulses indicate  $90^\circ$  and  $180^\circ$  pulses, respectively and all pulses are applied along the x-axis unless indicated otherwise.

Gradient levels were as follows:  $g1=(1\text{ ms}; 5\text{ G/cm})$ ;  $g2=(0.8\text{ ms}; 3\text{ G/cm})$ ;  $g3=(1.5\text{ ms}; 10\text{ G/cm})$ ;  $g6=(2\text{ ms}; \pm 20\text{ G/cm})$ ;  $g7=(0.5\text{ ms}; 5\text{ G/cm})$ ;  $g8=(0.7\text{ ms}; -8\text{ G/cm})$ ;  $g9=(0.2\text{ ms}; \pm 19.91\text{ G/cm})$ ;  $g11=(0.4\text{ ms}; 15\text{ G/cm})$ .

The phase cycling was  $\phi1=4(x), 8(-x), 4(x), 4(y), 8(-y), 4(y)$ ;  $\phi2=8(x), 8(-x), 8(y), 8(-y)$ ;  $\phi3=2(x), 2(-x)$ ;  $\phi6=x$ ; and the receiver was set to  $\text{rec}=x, -x, -x, x, -x, x, x, -x$ .

$^{13}\text{C}$ -filtered AFP-NOESY- $^1\text{H}$ - $^{15}\text{N}$ -HSQC experiments are used to probe differential structural compaction along the backbone via  $^1\text{H}$ - $^1\text{H}$  cross-relaxation dynamics. Comparison of the tilt-angle profiles of the effective cross-relaxation rates of H6-cBASP1 at pH6 and pH2 reveal a markedly different dynamic behavior of the conformational ensemble at different pH (Fig. 13). The distribution of zero-passages is much broader at pH6 and is shifted towards smaller tilt-angles, which is a sign for higher internal dynamics at pH6 (Fig. 13A, top). It is important to notice that amide-solvent hydrogen exchange will also affect the measurement of cross-relaxation rates. Enhanced hydrogen exchange of the amide groups at pH6 therefore also contributes to the observed shift of zero-passages towards smaller tilt-angles. Nonetheless, the distribution of  $\Delta_{\text{NOE}}$  intensities at  $\sin^2\theta_{\text{eff}} = 0.3$  (Fig. 13B, top) appears to depict some regions in the C-terminus with markedly higher internal dynamics/exchange contributions as well as regions with reduced internal dynamics/exchange predominantly at the N-terminus. Therefore, already at pH6 the protein appears to show differential dynamics along the protein, indicating different degrees of flexibility.

At pH2 zero-passages approach the typical tilt-angle  $\theta_{\text{eff}} = 35.26^\circ$  expected for rigid macromolecules, indicating less internal dynamics at this pH (Fig. 13A, bottom). Importantly, amide-solvent hydrogen exchange is strongly reduced at pH2 as indicated by the  $\text{HET}^{\text{ex}}$ -SOFAS experiments ( $\lambda_{\text{ex}} \approx 1$ , Fig. 6), therefore the distribution of zero-passages reflects more clearly the differential dynamics along the backbone. Interestingly, residue-based  $\Delta_{\text{NOE}}$  intensities at  $\sin^2\theta_{\text{eff}} = 0.3$  (Fig. 13B, bottom) reveal the differential dynamics of the protein backbone, indicating varying degrees of residue-based compaction. Residues in the N-terminal half of BASP1, which has been shown to undergo compaction at low pH, show negative and zero  $\Delta_{\text{NOE}}$  intensities, indicating zero-passage tilt-angles typical for a rigid macromolecule. On the contrary, positive intensities, hence zero-passages at smaller tilt-angles, are found in the C-terminal half and are indicative of residual internal dynamics in this part of the protein. This data is in line with results from Meta-Structure and Secondary Structure Propensity calculations (Fig. 3) as well as  $\text{HET}^{\text{noe}}$ -SOFAS experiments (Fig. 5) suggesting a pH-induced compaction and increase in the population of  $\alpha$ -helical conformation especially in the N-terminal Glu/Asp-rich region of cBASP1.



**Fig. 13.**  $^1\text{H}$ - $^1\text{H}$  homonuclear cross-relaxation rates during adiabatic fast passage as a probe for differential compaction along the protein backbone. (A) Tilt-angle profiles of the NOE/ROE-enhanced signal intensities ( $\Delta_{\text{NOE}}$ ) of H6-cBASP1 at pH6 (top) and pH2 (bottom). Mark the narrow distribution of zero-passages at pH2 as well as a shift of the zero-passages towards the theoretical tilt-angle  $\theta = 35.26^\circ$  ( $\sin^2\theta \approx 0.333$ ) found for a rigid macromolecule. (B)  $\Delta_{\text{NOE}}$  at  $\sin^2\theta = 0.3$  as a function of the residue indicating differential dynamics along the protein backbone.





# Discussion

It has clearly been shown that the conformational ensemble of chicken BASP1 undergoes substantial pH-induced rearrangements, with an increase in  $\alpha$ -helical propensities especially in protein regions enriched in acidic residues. This increase in population of helical structure is accompanied by an overall compaction of the protein. Interestingly, pH-induced compaction has been observed for other extended IDPs. Far-UV CD spectra of  $\alpha$ -synuclein and Prothymosin- $\alpha$  show the typical “turned out” response to changes in the pH, with a strong increase in negative intensity of the molar ellipticity at 222 nm  $[\theta]_{222}$  from pH5.5 to pH3, which was correlated with the formation of a partially folded pre-molten globule-like conformation (Uversky, 2009). The pH-induced protonation of acidic residues reduces the net-charge of IDPs, thereby decreasing electrostatic intramolecular repulsion. Additionally, the protonation-induced increase in hydrophobicity of these residues permits hydrophobic collapse to partially folded conformations. Specific intracellular environments might induce the partial protonation of acidic residues in BASP1 and a shift towards more compact conformations in its conformational ensemble.

Several IDPs have been shown to bind to artificial or natural membranes, with the binding accompanied by structural changes involving  $\alpha$ -helix formation.  $\alpha$ -synuclein has been proposed to exist in two structurally different isoforms in vivo: a disordered, cytosolic form and a  $\alpha$ -helix-rich membrane-bound form, which was suggested to generate nuclei that seed the aggregation of the more abundant cytosolic disordered isoform (Lee et al, 2002). The details of the interaction of  $\alpha$ -synuclein with lipid vesicles of different sizes and composition have been studied thoroughly by CD-spectroscopy, with some important results: lipids containing negatively charged phospholipids more strongly induced helix-formation and  $\alpha$ -synuclein bound preferentially to vesicles with smaller diameter (Zhu & Fink, 2003; Zhu et al, 2003). Interestingly, results on the implication of membrane-bound  $\alpha$ -synuclein in fibril

formation are quite controversial, some findings rather promoting the idea that membrane interaction impedes fibril formation, whereas others correlating membrane interaction with enhanced fibril formation. A study using NMR spectroscopy shed new light on the differences in the mode of interaction of wildtype  $\alpha$ -synuclein as compared to different mutants with membranes. Interestingly, three familial variants of  $\alpha$ -synuclein (A30P, E46K and A53T) associated with rare inherited Parkinson's Disease (PD) showed significantly reduced helix-formation in the center of the protein harboring the aggregation-prone NAC sequence upon interaction with a membrane, despite the quite different kinetics of interaction between all those mutants. Membrane interaction, but reduced helix formation, therefore might promote the intermolecular  $\beta$ -sheet formation of close monomers, enhancing fibril formation (Bodner et al, 2009).

These highly diverse observations show that the interaction of a disordered protein with a membrane is highly dependent on the cellular context, in the case of  $\alpha$ -synuclein depending on the plasma membrane composition, charge distribution, minor deviations in the primary sequence, etc. This again highlights the notion that IDPs constitute very sensitive probes for environmental changes.

BASP1 was shown to interact with the acidic phospholipid PIP<sub>2</sub>. Partial protonation of side chains as a consequence of a localized low pH might induce conformational changes in BASP1 that are needed for interaction with other proteins. Protonation-induced conformational changes might also play a role in oligomer formation at the lipid bilayer. One might imagine a scenario, where local sequestration of the acidic phospholipid PIP<sub>2</sub> by BASP1 leads to modulation of the local pH in proximity to the plasma membrane inducing changes in BASP1's structural preferences initiating oligomer formation. Ca<sup>2+</sup>-CaM binding as well as phosphorylation by PKC was shown to disrupt oligomers and leads to the demasking of PIP<sub>2</sub> clusters, enabling binding of other PIP<sub>2</sub> binding proteins (Zakharov & Mosevitsky, 2010). As Ca<sup>2+</sup>-CaM binds to the N-terminal myristoyl-group and the basic effector domain of BASP1, it interferes not only with the membrane anchoring of BASP1 but also with the PIP<sub>2</sub> interaction. The concomitant dissociation of BASP1 from the membrane and concentrated PIP<sub>2</sub> clusters will influence its structural propensities due to a change in the local environment, inducing oligomer dissociation.

Additionally, electrostatics might play an important role in the structural changes observed in BASP1. BASP1 is highly enriched in positively as well as negatively charged residues (see **FIG. 1** and inlay in **FIG. 5A**, top). Especially the N-terminus of BASP1 is enriched in lysine residues, which are indispensable for the interaction with PIP<sub>2</sub>. It has been

mentioned, that the “polyelectrostatic” effect in combination with the polarizability is of major importance in the interaction mediated by intrinsically disordered regions (Mittag et al, 2010). The electrostatic interaction of lysine residues with the anionic PIP<sub>2</sub> will also significantly contribute to structural rearrangements in BASP1 and influence its binding behavior.

Differences in structural propensities induced by subtle environmental changes (a special property of IDPs) in combination with oligomer formation offer the possibility of a quick and concerted response to external signals. In the case of BASP1, sequestration of PIP<sub>2</sub> will provide the necessary environment for oligomer formation and increasing oligomerization in turn will lead to a quicker accumulation of PIP<sub>2</sub> molecules. On the other hand, Ca<sup>2+</sup>-influx and the subsequent interference with membrane-anchoring and PIP<sub>2</sub> binding of BASP1 will lead to a quick destabilization of BASP1 oligomers due to the change in structural preference induced by the change in environmental conditions.

## ***Summary - Zusammenfassung***

**English:** Intrinsically Disordered Proteins (IDPs) are characterized by a heterogeneous and highly dynamic conformational ensemble. The lack of a stable well-folded structure in combination with fast interconversion of a multitude of accessible sub-states leads to reduced chemical shift dispersion and extensive signal overlap in 2D & 3D NMR experiments, aggravating chemical shift assignment by conventional methods. Random sampling of indirect time domains in higher dimensionality (5D) spectra together with Sparse Matrix Fourier Transformation (SMFT) is a novel and efficient methodology for the assignment of even large IDPs and allowed a nearly complete backbone assignment of human BASP1 (Brain Acid-Soluble Protein 1) (Article 2) and its chicken homolog cBASP1 in a short amount of time.

The possibility to quickly and fully assign IDPs offers new options for the study of the structural and dynamic behavior of this challenging protein family by NMR. The flexible nature of IDPs makes them sensitive probes for changes in environmental conditions (e.g. ionic strength, pH, molecular crowding, etc.), which can lead to substantial structural rearrangements within their conformational ensemble. Performing a large-scale analysis of the pH-dependence of IDPs with the recently introduced meta-structure approach, it was shown that IDPs in general tend to populate more compact conformations with a propensity towards  $\alpha$ -helix formation at low pH. The predictive validity of this approach was illustrated by its application to two well-studied examples of pH-induced compaction, the IDPs  $\alpha$ -synuclein and prothymosin  $\alpha$ , and was further substantiated by measurements on two of our model IDPs, the tumor suppressor BASP1 and the transcription factor Tcf4. The NMR chemical shift assignment of cBASP1 at two different pHs allowed the application of NMR spectroscopic methods monitoring residue-based Secondary Structure Propensity (SSP) changes as well as differential compaction along the protein backbone upon pH decrease, employing already established HET-SOFAST spectroscopy as well as a novel NMR spectroscopic method measuring  $^1\text{H}$ - $^1\text{H}$  homonuclear cross-relaxation during adiabatic fast passage. In addition, electron paramagnetic resonance (EPR)-based double electron-electron resonance (DEER) spectroscopy corroborated the finding of pH-induced compaction as monitored by meta-structure calculations and NMR spectroscopy (Article 1).

**Deutsch:** Das Struktur-Ensemble von Intrinsisch Ungeordneten Proteinen (IUPs) ist sehr heterogen und dynamisch. Das Fehlen einer stabilen Struktur und der schnelle Wechsel zwischen einer Vielzahl an möglichen Sub-Strukturen führt zu einer schmalen Verteilung der chemischen Verschiebungen und einer daraus resultierenden Überlappung von Resonanzen in 2D & 3D NMR Experimenten. Dies erschwert die Zuweisung von chemischen Verschiebungen zu einzelnen Atomen im Protein durch konventionelle Methoden. Random Sampling indirekter Dimensionen in Spektren höherer Dimensionalität (5D) in Kombination mit Sparse Matrix Fourier Transformation (SMFT) ist eine neue und effiziente Methode für die Zuordnung chemischer Verschiebungen selbst in großen IUPs. Mithilfe dieser 5D Experimente war es möglich in relativ kurzer Zeit fast das gesamte Protein-Rückgrat von humanem BASP1 (Brain Acid-Soluble Protein 1) (Artikel 2) und dessen Hühner-Homolog cBASP1 zuzuordnen.

Die Möglichkeit einer schnellen und umfassenden Zuordnung von IUPs eröffnet neue Optionen die Struktur und Dynamik dieser faszinierenden Protein-Familie zu erforschen. Die Flexibilität von IUPs macht sie sehr empfänglich für Änderungen in ihrer Umgebung (z.B. Ionenstärke, pH, hohe molekulare Dichte, etc.). Diese Änderungen können zu erheblichen strukturellen Umordnungen innerhalb ihres Struktur-Ensembles führen. Mithilfe einer umfassenden Meta-Struktur Analyse der pH-Abhängigkeit von IUPs konnte gezeigt werden, dass IUPs im Allgemeinen dazu tendieren bei niedrigem pH kompaktere Konformationen anzunehmen und sie dazu neigen  $\alpha$ -helikale Sekundär-Struktur auszubilden. Die Voraussagekraft des Meta-Struktur Ansatzes konnte durch seine Anwendung auf zwei gut erforschte Beispiele für pH-induzierte Kompaktierung veranschaulicht werden – die IUPs  $\alpha$ -Synuclein und Prothymosin  $\alpha$ . Messungen an zwei weiteren IUPs, dem Tumor Suppressor BASP1 und dem Transkriptionsfaktor Tcf4, untermauern zusätzlich die Validität der Meta-Struktur Analyse. Die Zuordnung der chemischen Verschiebungen von cBASP1 bei zwei unterschiedlichen pH Werten erlaubte die Anwendung NMR-spektroskopischer Methoden, die zum einen pH-abhängige Unterschiede in der Tendenz Sekundär-Struktur (Secondary Structure Propensity) auszubilden überprüfen, und zum anderen differentielle Kompaktierung entlang des Protein-Rückgrats aufgrund der pH-Änderung messen. Zu diesem Zweck wurden sowohl bereits etablierte HET-SOFAST Experimente angewendet, als auch eine neu entwickelte NMR-spektroskopische Methode, die  $^1\text{H}$ - $^1\text{H}$  homonukleare Cross-Relaxation während schneller adiabatischer Passage (adiabatic fast passage) misst. Weiters bestätigt Elektronenspinresonanz (ESR)-basierte Doppel-Elektron-Elektron-Resonanz (DEER)

Spektroskopie die Ergebnisse der Meta-Struktur Analyse und NMR Spektroskopie, die eine pH-induzierte Kompaktierung des Struktur-Ensembles von cBASP1 feststellen (Artikel 1).

# References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410
- Auer R (2010) NMR Relaxation under Spin-Locking Conditions. *Doctoral Thesis Universität Wien*
- Auer R, Kloiber K, Vavrinska A, Geist L, Coudeville N, Konrat R (2010) Pharmacophore mapping via cross-relaxation during adiabatic fast passage. *J Am Chem Soc* **132**: 1480-1481
- Bennett MC (2005) The role of alpha-synuclein in neurodegenerative diseases. *Pharmacol Ther* **105**: 311-331
- Bodner CR, Dobson CM, Bax A (2009) Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. *J Mol Biol* **390**: 775-790
- Bohlen JM, Bodenhausen G (1993) Experimental Aspects of Chirp NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* **102**: 293-301
- Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, Chan HS (2007) Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc Natl Acad Sci U S A* **104**: 9650-9655
- Bruschweiler S, Konrat R, Tollinger M (2013) Allosteric Communication in the KIX Domain Proceeds through Dynamic Repacking of the Hydrophobic Core. *ACS chemical biology*
- Bruschweiler S, Schanda P, Kloiber K, Brutscher B, Kontaxis G, Konrat R, Tollinger M (2009) Direct observation of the dynamic process underlying allosteric signal transmission. *J Am Chem Soc* **131**: 3063-3068
- Carpenter B, Hill KJ, Charalambous M, Wagner KJ, Lahiri D, James DI, Andersen JS, Schumacher V, Royer-Pokora B, Mann M, Ward A, Roberts SG (2004) BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1. *Mol Cell Biol* **24**: 537-549
- Cho MK, Nodet G, Kim HY, Jensen MR, Bernado P, Fernandez CO, Becker S, Blackledge M, Zweckstetter M (2009) Structural characterization of alpha-synuclein in an aggregation prone state. *Protein Science* **18**: 1840-1846
- Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nature structural biology* **4**: 10-19
- Dosztanyi Z, Chen J, Dunker AK, Simon I, Tompa P (2006) Disorder and sequence repeats in hub proteins and their implications for network evolution. *Journal of proteome research* **5**: 2985-2995
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *Journal of molecular graphics & modelling* **19**: 26-59
- Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* **18**: 756-764
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**: 197-208
- Eliezer D (2007) Characterizing residual structure in disordered protein States using nuclear magnetic resonance. *Methods Mol Biol* **350**: 49-67
- Ellena JF, Burnitz MC, Cafiso DS (2003) Location of the myristoylated alanine-rich C-kinase substrate (MARCKS) effector domain in negatively charged phospholipid bicelles. *Biophys J* **85**: 2442-2448
- Essafi A, Webb A, Berry RL, Slight J, Burn SF, Spraggon L, Velecela V, Martinez-Estrada OM, Wiltshire JH, Roberts SG, Brownstein D, Davies JA, Hastie ND, Hohenstein P (2011) A wt1-controlled chromatin switching mechanism underpins tissue-specific wnt4 activation and repression. *Developmental cell* **21**: 559-574
- Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* **15**: 35-41
- Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**: 2985-2993
- 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**: 1443-1454

- Geist L, Henen MA, Haiderer S, Schwarz TC, Kurzbach D, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Kozminski W, Hinderberger D, Konrat R (2013) Protonation-dependent conformational variability of intrinsically disordered proteins. *Protein science : a publication of the Protein Society* **22**: 1196-1205
- Geist L, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Kozminski W, Konrat R (2012) (1)H, (13)C and (15)N resonance assignments of human BASP1. *Biomolecular NMR assignments*
- Glabe CG (2006) Common mechanisms of amyloid oligomer pathogenesis in degenerative disease. *Neurobiology of aging* **27**: 570-575
- Golub T, Wacha S, Caroni P (2004) Spatial and temporal control of signaling through lipid rafts. *Current opinion in neurobiology* **14**: 542-550
- Goodfellow SJ, Rebello MR, Toska E, Zeef LA, Rudd SG, Medler KF, Roberts SG (2011) WT1 and its transcriptional cofactor BASP1 redirect the differentiation pathway of an established blood cell line. *Biochem J* **435**: 113-125
- Green LM, Wagner KJ, Campbell HA, Addison K, Roberts SG (2009) Dynamic interaction between WT1 and BASP1 in transcriptional regulation during differentiation. *Nucleic Acids Res* **37**: 431-440
- Hartwig JH, Bokoch GM, Carpenter CL, Janmey PA, Taylor LA, Toker A, Stossel TP (1995) Thrombin receptor ligation and activated Rac uncap actin filament barbed ends through phosphoinositide synthesis in permeabilized human platelets. *Cell* **82**: 643-653
- Hasenohrl D, Konrat R, Blasi U (2011) Identification of an RNase J ortholog in *Sulfolobus solfataricus*: implications for 5'-to-3' directional decay and 5'-end protection of mRNA in Crenarchaeota. *RNA* **17**: 99-107
- Henen MA, Coudeville N, Geist L, Konrat R (2012) Toward rational fragment-based lead design without 3D structures. *Journal of medicinal chemistry* **55**: 7909-7919
- Hirakura Y, Azimov R, Azimova R, Kagan BL (2000) Polyglutamine-induced ion channels: a possible mechanism for the neurotoxicity of Huntington and other CAG repeat diseases. *J Neurosci Res* **60**: 490-494
- Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol* **323**: 573-584
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* **32**: 1037-1049
- Janmey PA, Lindberg U (2004) Cytoskeletal regulation: rich in lipids. *Nat Rev Mol Cell Biol* **5**: 658-666
- Jensen MR, Ruigrok RW, Blackledge M (2013) Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr Opin Struct Biol*
- Kagan BL, Azimov R, Azimova R (2004) Amyloid peptide channels. *The Journal of membrane biology* **202**: 1-10
- Keeler J (2007) *Understanding NMR spectroscopy*, Chichester: Wiley.
- Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J Biomol NMR* **49**: 139-149
- Konrat R (2009) The protein meta-structure: a novel concept for chemical and molecular biology. *Cell Mol Life Sci* **66**: 3625-3639
- Konrat R, Tollinger M (1999) Heteronuclear relaxation in time-dependent spin systems: (15)N-T1 (rho) dispersion during adiabatic fast passage. *J Biomol NMR* **13**: 213-221
- Kontaxis G, Delaglio F, Bax A (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol* **394**: 42-78
- Korshunova I, Caroni P, Kolkova K, Berezin V, Bock E, Walmod PS (2008) Characterization of BASP1-mediated neurite outgrowth. *J Neurosci Res* **86**: 2201-2213
- 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. *Scientific reports* **3**: 1392
- Kupce E, Freeman R (1995) Adiabatic Pulses for Wideband Inversion and Broadband Decoupling. *Journal of Magnetic Resonance, Series A* **115**: 273-276



- Kupce E, Freeman R (2007) Compensated adiabatic inversion pulses: broadband INEPT and HSQC. *J Magn Reson* **187**: 258-265
- Lacy ER, Filippov I, Lewis WS, Otieno S, Xiao L, Weiss S, Hengst L, Kriwacki RW (2004) p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nature structural & molecular biology* **11**: 358-364
- 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**: 1455-1472
- Lee HJ, Choi C, Lee SJ (2002) Membrane-bound alpha-synuclein has a high aggregation propensity and the ability to seed the aggregation of the cytosolic form. *J Biol Chem* **277**: 671-678
- Lipari G, Szabo A (1982a) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* **104**: 4546-4559
- Lipari G, Szabo A (1982b) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc* **104**: 4559-4570
- Liu H, Liu W, Wu Y, Zhou Y, Xue R, Luo C, Wang L, Zhao W, Jiang JD, Liu J (2005) Loss of epigenetic control of synuclein-gamma gene as a molecular indicator of metastasis in a wide range of human cancers. *Cancer research* **65**: 7635-7643
- Ma L, Cantley LC, Janmey PA, Kirschner MW (1998) Corequirement of specific phosphoinositides and small GTP-binding protein Cdc42 in inducing actin assembly in Xenopus egg extracts. *J Cell Biol* **140**: 1125-1136
- Maekawa S, Maekawa M, Hattori S, Nakamura S (1993) Purification and molecular cloning of a novel acidic calmodulin binding protein from rat brain. *J Biol Chem* **268**: 13703-13709
- Maekawa S, Murofushi H, Nakamura S (1994) Inhibitory effect of calmodulin on phosphorylation of NAP-22 with protein kinase C. *J Biol Chem* **269**: 19462-19465
- Mancek-Keber M, Bencina M, Japelj B, Panter G, Andra J, Brandenburg K, Triantafilou M, Triantafilou K, Jerala R (2012) MARCKS as a negative regulator of lipopolysaccharide signaling. *J Immunol* **188**: 3893-3902
- Margreitter C, Petrov D, Zagrovic B (2013) Vienna-PTM web server: a toolkit for MD simulations of protein post-translational modifications. *Nucleic Acids Res* **41**: W422-426
- Marsh JA, Singh VK, Jia Z, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein science : a publication of the Protein Society* **15**: 2795-2804
- Matsubara M, Nakatsu T, Kato H, Taniguchi H (2004) Crystal structure of a myristoylated CAP-23/NAP-22 N-terminal domain complexed with Ca<sup>2+</sup>/calmodulin. *Embo J* **23**: 712-718
- Mayer C, Slater L, Erat MC, Konrat R, Vakonakis I (2012) Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J Biol Chem* **287**: 7182-7189
- Metallo SJ (2010) Intrinsically disordered proteins are potential drug targets. *Current opinion in chemical biology* **14**: 481-488
- Mittag T, Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol* **17**: 3-14
- Mittag T, Kay LE, Forman-Kay JD (2010) Protein dynamics and conformational disorder in molecular recognition. *Journal of molecular recognition : JMR* **23**: 105-116
- Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK, Uversky VN (2006) Analysis of molecular recognition features (MoRFs). *J Mol Biol* **362**: 1043-1059
- Mosevitsky MI (2005) Nerve ending "signal" proteins GAP-43, MARCKS, and BASP1. *Int Rev Cytol* **245**: 245-325
- Mosevitsky MI, Capony JP, Skladchikova G, Novitskaya VA, Plekhanov A, Zakharov VV (1997) The BASP1 family of myristoylated proteins abundant in axonal termini. Primary structure analysis and physico-chemical properties. *Biochimie* **79**: 373-384
- Ohsawa S, Watanabe T, Katada T, Nishina H, Miura M (2008) Novel antibody to human BASP1 labels apoptotic cells post-caspase activation. *Biochem Biophys Res Commun* **371**: 639-643

- Ostroumova OS, Schagina LV, Mosevitsky MI, Zakharov VV (2011) Ion channel activity of brain abundant protein BASP1 in planar lipid bilayers. *The FEBS journal* **278**: 461-469
- Penninger JM, Crabtree GR (1999) The actin cytoskeleton and lymphocyte activation. *Cell* **96**: 9-12
- Perera RM, Zoncu R, Lucast L, De Camilli P, Toomre D (2006) Two synaptojanin 1 isoforms are recruited to clathrin-coated pits at different stages. *Proc Natl Acad Sci U S A* **103**: 19332-19337
- Quist A, Doudevski I, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R (2005) Amyloid ion channels: a common structural link for protein-misfolding disease. *Proc Natl Acad Sci U S A* **102**: 10427-10432
- Royle SJ, Lagnado L (2010) Clathrin-mediated endocytosis at the synaptic terminal: bridging the gap between physiology and molecules. *Traffic* **11**: 1489-1497
- Sakisaka T, Itoh T, Miura K, Takenawa T (1997) Phosphatidylinositol 4,5-bisphosphate phosphatase regulates the rearrangement of actin filaments. *Mol Cell Biol* **17**: 3841-3849
- Schanda P, Brutscher B (2005) 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* **127**: 8014-8015
- Schanda P, Forge V, Brutscher B (2006) HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magnetic resonance in chemistry : MRC* **44 Spec No**: S177-184
- Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* **33**: 199-211
- Schedlbauer A, Gandini R, Kontaxis G, Paulmichl M, Furst J, Konrat R (2011) The C-terminus of ICLn is natively disordered but displays local structural preformation. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* **28**: 1203-1210
- Schneider R, Huang JR, Yao M, Communie G, Ozenne V, Mollica L, Salmon L, Ringkjøbing Jensen M, Blackledge M (2012) Towards a robust description of intrinsic protein disorder using nuclear magnetic resonance spectroscopy. *Mol Biosyst* **8**: 58-68
- Schwarzinger S, Kroon GJ, Foss TR, Chung J, Wright PE, Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J Am Chem Soc* **123**: 2970-2978
- Schwarzinger S, Kroon GJ, Foss TR, Wright PE, Dyson HJ (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J Biomol NMR* **18**: 43-48
- Seykora JT, Myat MM, Allen LA, Ravetch JV, Aderem A (1996) Molecular determinants of the myristoyl-electrostatic switch of MARCKS. *J Biol Chem* **271**: 18797-18802
- Shibasaki Y, Ishihara H, Kizuki N, Asano T, Oka Y, Yazaki Y (1997) Massive actin polymerization induced by phosphatidylinositol-4-phosphate 5-kinase in vivo. *J Biol Chem* **272**: 7578-7581
- Shupliakov O, Brodin L (2010) Recent insights into the building and cycling of synaptic vesicles. *Experimental cell research* **316**: 1344-1350
- Sibille N, Bernado P (2012) Structural characterization of intrinsically disordered proteins by the combined use of NMR and SAXS. *Biochemical Society transactions* **40**: 955-962
- Spera S, Bax A (1991) Empirical Correlation between Protein Backbone Conformation and Ca and Cb <sup>13</sup>C Nuclear Magnetic Resonance Chemical Shifts. *J Am Chem Soc* **113**: 5490-5492
- Takaichi R, Odagaki S, Kumanogoh H, Nakamura S, Morita M, Maekawa S (2012) Inhibitory effect of NAP-22 on the phosphatase activity of synaptojanin-1. *J Neurosci Res* **90**: 21-27
- Takasaki A, Hayashi N, Matsubara M, Yamauchi E, Taniguchi H (1999) 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* **274**: 11848-11853
- Tamiola K, Acar B, Mulder FA (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc* **132**: 18000-18003
- Tannus A, Garwood M (1997) Adiabatic pulses. *NMR in biomedicine* **10**: 423-434

- Tantos A, Han KH, Tompa P (2012) Intrinsic disorder in cell signaling and gene transcription. *Molecular and cellular endocrinology* **348**: 457-465
- Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**: 527-533
- Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* **37**: 509-516
- Tompa P, Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* **33**: 2-8
- Toska E, Campbell HA, Shandilya J, Goodfellow SJ, Shore P, Medler KF, Roberts SG (2012) Repression of Transcription by WT1-BASP1 Requires the Myristoylation of BASP1 and the PIP2-Dependent Recruitment of Histone Deacetylase. *Cell reports* **2**: 462-469
- Uversky VN (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *Journal of biomolecular structure & dynamics* **21**: 211-234
- Uversky VN (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *The protein journal* **28**: 305-325
- Uversky VN, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annual review of biophysics* **37**: 215-246
- Uversky VN, Oldfield CJ, Midic U, Xie H, Xue B, Vucetic S, Iakoucheva LM, Obradovic Z, Dunker AK (2009) Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* **10 Suppl 1**: S7
- Welch MD, Mallavarapu A, Rosenblatt J, Mitchison TJ (1997) Actin dynamics in vivo. *Current opinion in cell biology* **9**: 54-61
- Widmer F, Caroni P (1990) Identification, localization, and primary structure of CAP-23, a particle-bound cytosolic protein of early development. *J Cell Biol* **111**: 3035-3047
- Wiederkehr A, Staple J, Caroni P (1997) The motility-associated proteins GAP-43, MARCKS, and CAP-23 share unique targeting and surface activity-inducing properties. *Experimental cell research* **236**: 103-116
- Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* **34**: D668-672
- Wishart DS, Sykes BD (1994) The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J Biomol NMR* **4**: 171-180
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* **19**: 31-38
- Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN (2007a) Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *Journal of proteome research* **6**: 1917-1932
- Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z (2007b) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *Journal of proteome research* **6**: 1882-1898
- Zakharov SD, Hulleman JD, Dutseva EA, Antonenko YN, Rochet JC, Cramer WA (2007) Helical alpha-synuclein forms highly conductive ion channels. *Biochemistry* **46**: 14369-14379
- Zakharov VV, Mosevitsky MI (2010) Oligomeric structure of brain abundant proteins GAP-43 and BASP1. *J Struct Biol* **170**: 470-483
- Zhang H, Neal S, Wishart DS (2003) RefDB: a database of uniformly referenced protein chemical shifts. *J Biomol NMR* **25**: 173-195
- Zhu M, Fink AL (2003) Lipid binding inhibits alpha-synuclein fibril formation. *J Biol Chem* **278**: 16873-16877
- Zhu M, Li J, Fink AL (2003) The association of alpha-synuclein with membranes affects bilayer structure, stability, and fibril formation. *J Biol Chem* **278**: 40186-40197

The author's contributions to the following articles were the preparation of BASP1 samples as well as NMR-spectroscopic measurements and data processing for the elucidation of pH-induced structural and dynamic changes in BASP1. Moreover, the author contributed to the writing of the manuscripts.

# ***Article 1***

# Protonation-dependent conformational variability of intrinsically disordered proteins

Leonhard Geist,<sup>1</sup> Morkos A. Henen,<sup>1,2</sup> Sandra Haiderer,<sup>1</sup> Thomas C. Schwarz,<sup>1</sup> Dennis Kurzbach,<sup>3</sup> Anna Zawadzka-Kazimierczuk,<sup>4</sup> Saurabh Saxena,<sup>4</sup> Szymon Żerko,<sup>4</sup> Wiktor Koźmiński,<sup>4</sup> Dariush Hinderberger,<sup>3</sup> and Robert Konrat<sup>1\*</sup>

<sup>1</sup>Department of Computational and Structural Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, Austria

<sup>2</sup>Faculty of Pharmacy, Mansoura University, Egypt

<sup>3</sup>Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

<sup>4</sup>Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

Received 21 March 2013; Revised 20 June 2013; Accepted 21 June 2013

DOI: 10.1002/pro.2304

Published online 2 July 2013 proteinscience.org

© 2013 The Authors. [Protein Science] published by Wiley-Blackwell on behalf of The Protein Society.

This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

**Abstract:** Intrinsically disordered proteins (IDPs) are characterized by substantial conformational plasticity and undergo rearrangements of the time-averaged conformational ensemble on changes of environmental conditions (e.g., in ionic strength, pH, molecular crowding). In contrast to stably folded proteins, IDPs often form compact conformations at acidic pH. The biological relevance of this process was, for example, demonstrated by nuclear magnetic resonance studies of the aggregation prone (low pH) state of  $\alpha$ -synuclein. In this study, we report a large-scale analysis of the pH dependence of disordered proteins using the recently developed meta-structure approach. The meta-structure analysis of a large set of IDPs revealed a significant tendency of IDPs to form  $\alpha$ -helical secondary structure elements and to preferentially fold into more compact structures under acidic conditions. The predictive validity of this novel approach was demonstrated with applications to the tumor-suppressor BASP1 and the transcription factor Tcf4.

**Keywords:** intrinsically disordered proteins; protein meta-structure; pH dependence; structural biology; biomolecular NMR; EPR spectroscopy

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: FWF; Grant number: P20549-N19 and W-1221-B03 (to R.K.); Grant sponsor: Max Planck Graduate Center with the University of Mainz (MPGC) and Deutsche Forschungsgemeinschaft (DFG); Grant number: HI 1094/2-1; Grant sponsor: Gutenberg Academy of the University of Mainz (to D.K.); Grant sponsor: ÖAD entitled "KU mit Ägypten art 9/2 (to M.A.H.); Grant sponsor: Bio-NMR (Project No. 161863) and EAST-NMR (Project No. 228461) within Transnational Access activity, funded by European Commission's Framework Program 7 (FP7); Grant sponsor: Foundation for Polish science (to S.S., S.Ż., and W.K.; TEAM program).

\*Correspondence to: Robert Konrat; Department of Computational and Structural Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030 Vienna, Austria. E-mail: robert.konrat@univie.ac.at

Intrinsically disordered proteins (IDPs) display conformational plasticity and undergo structural rearrangements upon changes of pH. Here we provide experimental evidence for  $\alpha$ -helix formation and structural compaction of IDPs under acidic conditions. Sequence analysis of a large set of IDPs revealed a general tendency to form  $\alpha$ -helical secondary structure elements and to preferentially fold into more compact structures under acidic conditions suggesting that elements sensitive to protonation-induced stabilization might be crucial in folding-upon-binding events.

## Introduction

Intrinsically disordered proteins (IDPs) display remarkable physicochemical properties that allow

for unique regulatory functionalities in living organisms.<sup>1,2</sup> It is well established that IDPs constitute interaction hubs in the protein interactomes of organisms and have therefore attracted considerable scientific interest.<sup>3</sup> In addition to their innate conformational flexibility (plasticity), IDPs are also sensitive probes for changes of environmental parameters, particularly changes in pH values.<sup>4–6</sup> In a recent thorough analysis, it was shown that the net charge (and proline content) is an important factor determining hydrodynamic radii and governing structural compaction of IDPs.<sup>7</sup> For example, it was shown that although  $\alpha$ -synuclein remains natively unfolded under acidic conditions, local secondary structure propensities in proximity to acidic residues change on variations in pH and the conformational ensemble becomes enriched in compact structures with pronounced local rigidity of the protein backbone.<sup>8</sup> Similar physicochemical behavior was reported for the thymic hormone prothymosin- $\alpha$ .<sup>9</sup> Using a suite of different biophysical techniques (e.g., small-angle X-ray scattering (SAXS), UV, and NMR), it was shown that prothymosin- $\alpha$  cooperatively folds from a denatured state at neutral pH into a partially folded conformation at lower pH. This folding transition is initiated by the formation of  $\alpha$ -helices in glutamic acid-rich regions in which the neutralization of negatively charged side-chains facilitates the hydrophobic collapse.<sup>9</sup>

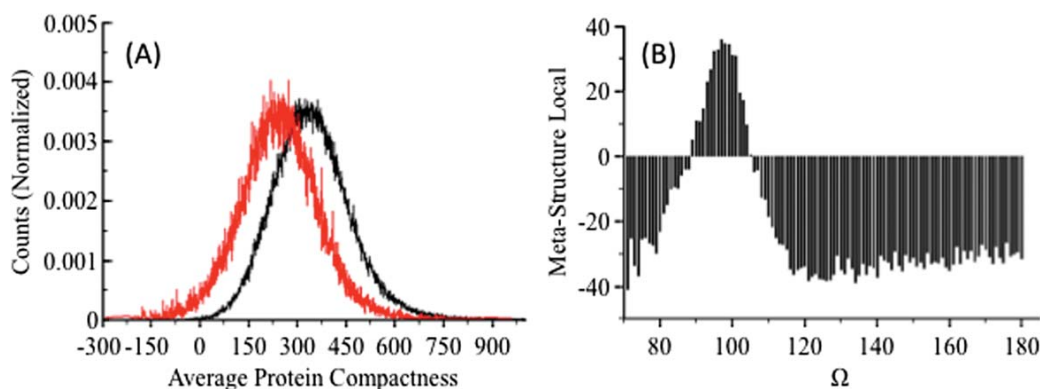
Protonation and deprotonation are fundamental chemical mechanisms to modulate ionization states and physicochemical properties of biomolecules and therefore of great relevance for biological systems. In eukaryotic cells, the pH in different organelles is highly variable ranging from 7.2 to 7.4 in the cytosol and nucleus to 4.0–5.5 in the endosomes and lysosomes.<sup>10</sup> Thus, pH dysregulation frequently leads to cellular dysfunction and ultimately to disease phenotypes. For example, tumor cells have been shown to behave abnormally under acidic cellular pH conditions<sup>11</sup> and pH alterations in the acidic organelles lead to oncological processes and lysosomal storage disorder.<sup>12</sup> In addition to modulating the ionization states of molecules and their interaction patterns with authentic binding partners, alteration of pH can also induce structural changes in proteins. Typically, proteins are only marginally stable and small changes in environmental conditions can induce substantial structural rearrangements. Given the small stability differences between the folded and unfolded state(s), slight changes of structure-disrupting factors (concentration of chemical denaturants, pH, or temperature) may lead to substantial and even cooperative changes in the population of folded and unfolded states, respectively. It is important to note that partial proton transfer is often sufficient for catalytic acceleration of chemical reactions and complete proton transfer might not be necessary. This is

described in the concept of general versus specific acid catalysis in physical organic chemistry and enzymology. Protonation of a given substrate modifies its electronic state and may yield, for example, a strong electrophilic center that may be prone to attacks of nucleophiles. In return, this means that a reduction of electrostatic repulsion among deprotonated, acidic residues of a protein does not necessarily require complete protonation of acidic side chains.

In this study, we use the recently developed meta-structure concept, a computational approach,<sup>13</sup> to analyze structural changes in IDPs under acidic conditions. The meta-structure analysis provides quantitative (on a per residue basis) information about residue structural complexity (compactness) and local secondary structure. Compactness is inversely related to local residue solvent-exposure, whereas local secondary structure values are defined analogously to the NMR secondary C $\alpha$  chemical shift ( $\alpha$ -helices: positive;  $\beta$ -sheets: negative). Numerous applications have already demonstrated the applicability of this approach.<sup>14–17</sup> In this study, we computationally consider the effect of lowering the pH by altering the primary sequence through replacing the negatively charged carboxylate groups of Asp/Glu residues with neutral (hydrogen-bond donor) amide groups of Asn/Gln residues. The predictive validity of our meta-structure approach is illustrated by its application to the IDPs BASP1, a tumor-suppressor down regulated by the oncogenic transcription factor Myc,<sup>18</sup> and Tcf4, a transcription factor involved in the Wnt signaling pathway.<sup>19</sup>

## Results

The meta-structure derived compactness and local secondary structure parameters were used to assess pH-induced changes in protein structural features. Residue-specific compactness values quantify the spatial embedding of individual residues within the 3D protein structures. Residues in the interior of a structure exhibit large compactness values, whereas residues located on the surface and exposed to the solvent display small (even negative in case of conformationally highly flexible segments) values. The meta-structure derived secondary structure parameter is defined in analogy to the well-established NMR <sup>13</sup>C $\alpha$  chemical shift index, with positive values for  $\alpha$ -helices and negative values indicating the presence of an extended conformation. First, we evaluated the performance of the approach by comparing calculated compactness values of IDPs (taken from the DisProt database) with well-folded proteins deposited in the PDB database. As can be seen from the histograms shown in Figure 1(A), IDPs are characterized by significantly smaller compactness values ( $\sim$ 230) compared to their well-folded counterparts ( $\sim$ 330), indicating that compactness



**Figure 1.** Large-scale performance evaluation of meta-structure protein parameters. (A) Histogram of calculated residue compactness values for proteins taken from the PDB (black) or DisProt database (red). On average, significantly smaller compactness values are found for intrinsically disordered proteins. (B) Comparison between calculated secondary structure parameters and local secondary structure (expressed by the backbone parameter  $\Omega$ ). Protein backbone dihedral angles  $\varphi$  and  $\psi$  are transformed into a single geometric parameter using the well-established relationship:  $3\cos\Omega = 1 - 4\cos^2[(\varphi + \psi)/2]$ .<sup>21</sup> Note that  $\alpha$ -helical regions have  $\Omega$  values between 90 and 110, whereas  $\beta$ -sheets are characterized by  $\Omega \approx 140$ –180. The protein dataset was taken from Bax and coworkers.<sup>20</sup>

values are valuable quantitative probes for structural compaction of proteins. Then, we analyzed whether the calculated local secondary structure parameters are indicative of  $\alpha$ -helices and  $\beta$ -strands. To facilitate the analysis, protein backbone dihedral angles  $\varphi$  and  $\psi$  were extracted for a set of 857 proteins taken from the PDB database (protein structure database designed by Kontaxis *et al.*<sup>20</sup>). Dihedral angle combinations were transformed into a single geometrical parameter using the well-established relationship:  $3\cos\Omega = 1 - 4\cos^2[(\varphi + \psi)/2]$ .<sup>21</sup> Figure 1(B) shows that meta-structure derived secondary structure values are consistently positive for residues located in  $\alpha$ -helical segments, whereas residues populating extended structural elements ( $\beta$ -strands or polyproline II helices) display negative values. It has to be noted that, although  $\alpha$ -helices can be directly inferred from the meta-structure analysis, negative secondary structure values are found for both  $\beta$ -strands and other extended segments (polyproline II helices). As a summary, we conclude that changes in local structure preferences (e.g.,  $\alpha$ -helix to  $\beta$ -strand transition) can be correctly identified by this approach.

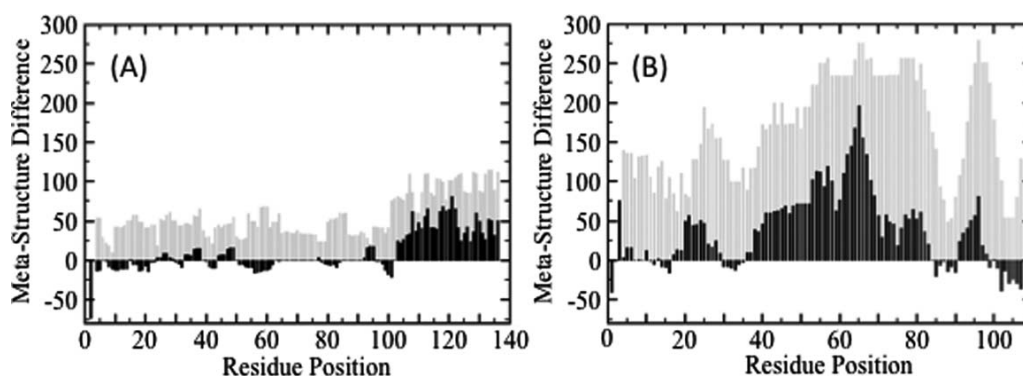
We have applied our meta-structure approach to two well-studied examples of acid-induced structural changes of IDPs. A recent NMR analysis of the pH-dependence of  $\alpha$ -synuclein revealed significant structural adaptations in the C-terminal (Asp/Glu-rich) region. Although at near neutral pH residues in the C-terminal part populate extended conformations, propensities for  $\alpha$ -helical elements are increased on lowering pH.<sup>8</sup> Surprisingly, independent FTIR data suggested an increase in populations of  $\beta$ -sheet conformations.<sup>22</sup> The secondary structure change was accompanied by an overall compaction of the polypeptide chain at low pH as probed by pulsed-field

gradient (PFG) measurements of the hydrodynamic radius,  $R_H$ , and long-range paramagnetic relaxation enhancement data.<sup>8</sup> Guinier analysis of SAXS data also showed that at neutral pH the radius of gyration of  $\alpha$ -synuclein amounts to  $R_g = 40$  Å, whereas it drops to  $\sim 30$  Å at low pH.<sup>22</sup> The meta-structure analysis shown in Figure 2(A) convincingly supports these findings: (i) increase in  $\alpha$ -helical propensities in the C-terminal (Asp- and Glu-rich) region; (ii) increased population of  $\beta$ -sheet conformations in the N-terminal region that houses the aggregation relevant NAC region of  $\alpha$ -synuclein; and (iii) significant local compaction of the C-terminal region.

A second example is given by prothymosin  $\alpha$ . It was shown that prothymosin  $\alpha$  exhibits an increase in  $\alpha$ -helical content from nearly 0% at high pH to about 15% at low pH, coincident with a considerable reduction of the radius of gyration (high pH: 37.8 Å; low pH: 27.6 Å).<sup>9</sup> Inspection of Figure 2(B) shows that the significant increase in compactness and  $\alpha$ -helical content compared to neutral conditions is correctly identified by the meta-structure analysis. Summing up, these two applications clearly demonstrated that the novel meta-structure approach provides reliable residue-specific information about the impact of protonation on local second structure and structural compaction.

In a second step, we applied our approach to two IDPs: BASP1, a tumor-suppressor, and Tcf4, a transcription factor involved in Wnt signaling. BASP1 is an N-myristoylated protein that is involved in neurite outgrowth and plasma membrane organization<sup>23</sup> and has been shown to be involved in transcription regulation by acting as a co-suppressor of WT1 function (Wilms' tumor suppressor protein 1).<sup>24,25</sup> Figure 3(A) shows the results of meta-structure analysis of BASP1. From the



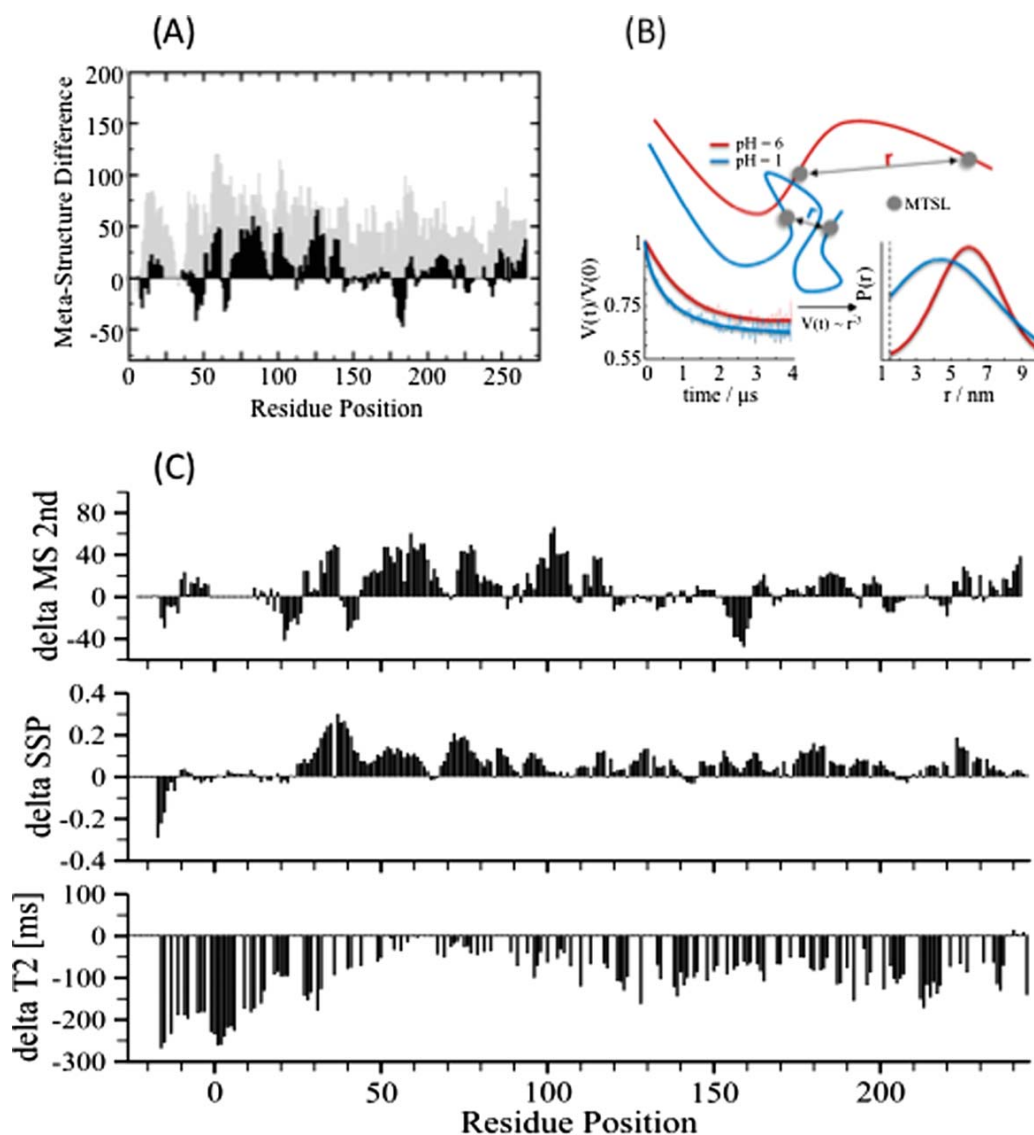


**Figure 2.** Meta-structure derived pH dependence of intrinsically disordered proteins. The effect of protonation (low pH condition) is considered by replacing Asp and Glu with Asn and Gln residues in the primary sequence. The meta-structure parameters are calculated for both wild-type and “mutated” protein forms. Differences in meta-structure derived local secondary structure (black) and compactness values (gray) are related to pH-induced alterations of local backbone geometry and compaction of the polypeptide chain. The differences shown in the figure are defined as low pH–high pH. Data are given for (A)  $\alpha$ -synuclein and (B) prothymosin  $\alpha$ .

residue plot it can be deduced that under acidic conditions (pH 2) BASP1 displays larger fractions of  $\alpha$ -helical regions as compared to pH 6. Furthermore, it samples more compact conformations, as indicated by the significantly larger compactness values. The meta-structure-predicted overall compaction of BASP1 was experimentally verified using PFG NMR diffusion measurements (PFG-DOSY), although significant scatter in the extracted diffusion constants impaired a detailed quantitative analysis. Additional support for the structural compaction of BASP1 was obtained from well-established NMR saturation experiments.<sup>26</sup> It was shown that the SOFAST-HMQC experiment can efficiently be used to probe  $^1\text{H}$ - $^1\text{H}$  spin diffusion or NOE effects, when a selective inversion pulse ( $\text{H}^{\text{sat}}$ ) is applied on aliphatic protons before the start of the pulse sequence. Two data sets are recorded with ( $\text{I}^{\text{sat}}$ ) and without ( $\text{I}^{\text{ref}}$ ) the inversion pulse  $\text{H}^{\text{sat}}$ . The intensity ratio ( $\lambda_{\text{NOE}} = \text{I}^{\text{sat}}/\text{I}^{\text{ref}}$ ) measured in the two SOFAST- $^1\text{H}$ - $^{15}\text{N}$ -HMQC spectra depends on spin diffusion effects and thus probes the structure and dynamics of proton spin networks (density and mobility) in proteins.<sup>26</sup> In well-structured, globular proteins spin diffusion is highly efficient leading to  $\lambda_{\text{NOE}} \ll 1$ , whereas in loosely folded proteins (random coils, molten globules)  $\lambda_{\text{NOE}} \approx 1$ . Thus, we have applied HET-SOFAST-HMQC to BASP1 to probe structural compaction under high and low pH conditions. The significant decrease of  $\lambda_{\text{NOE}}$  on lowering pH (0.75–0.60) again indicates structural compaction of BASP1 under low pH conditions. The overall compaction of BASP1 at low pH was independently corroborated by electron paramagnetic resonance (EPR)-based double electron-electron resonance (DEER) measurements.<sup>27</sup> Figure 3(B) shows DEER time traces obtained at different pH values (pH 1 and pH 6). Changes in solution pH clearly influence effective modulation depths,  $\Delta_{\text{eff}}$ , and the pair-wise distance distributions,  $P(r)$ . At low pH, the effective

modulation depths,  $\Delta_{\text{eff}}$ , is larger and  $P(r)$  is shifted toward shorter distances. More specific information about local structural dynamics changes was obtained using residue-specific NMR data. Inspection of Figure 3(C) displays a clear correlation between meta-structure derived [Fig. 3(C), top] and experimental NMR-derived differential secondary structural propensity (SSP) changes [SSP (pH2) - SSP (pH6); Fig. 3(C), middle]. Consistently, most pronounced secondary structural changes are found in the glutamic acid-rich region 30–100, whereas smaller secondary structural changes were observed for segments around residues 180 and 230. The regions of increased  $\alpha$ -helical content at low pH coincide with regions displaying reduced local mobility under acidic conditions as indicated by differential  $^{15}\text{N}$   $T_2$  values ( $^{15}\text{N}$   $T_2$  (pH6) -  $^{15}\text{N}$   $T_2$  (pH2)). These regions appear as local maxima in the differential  $^{15}\text{N}$   $T_2$  versus residue plot [Fig. 3(C), bottom]. Interestingly, the global compaction of BASP1 at low pH is further indicated by larger  $^{15}\text{N}$   $T_2$  values (smaller apparent correlation time  $\tau_c$ ) [Fig. 3(C), bottom]. It should be noted that on structural compaction BASP1 retains significant internal mobility that is significantly different to the behavior of well-folded (ordered) proteins. These local changes in internal dynamics lead to residue-specific modulation of the experimental  $^{15}\text{N}$   $T_2$  values, for example, local secondary structure formation.

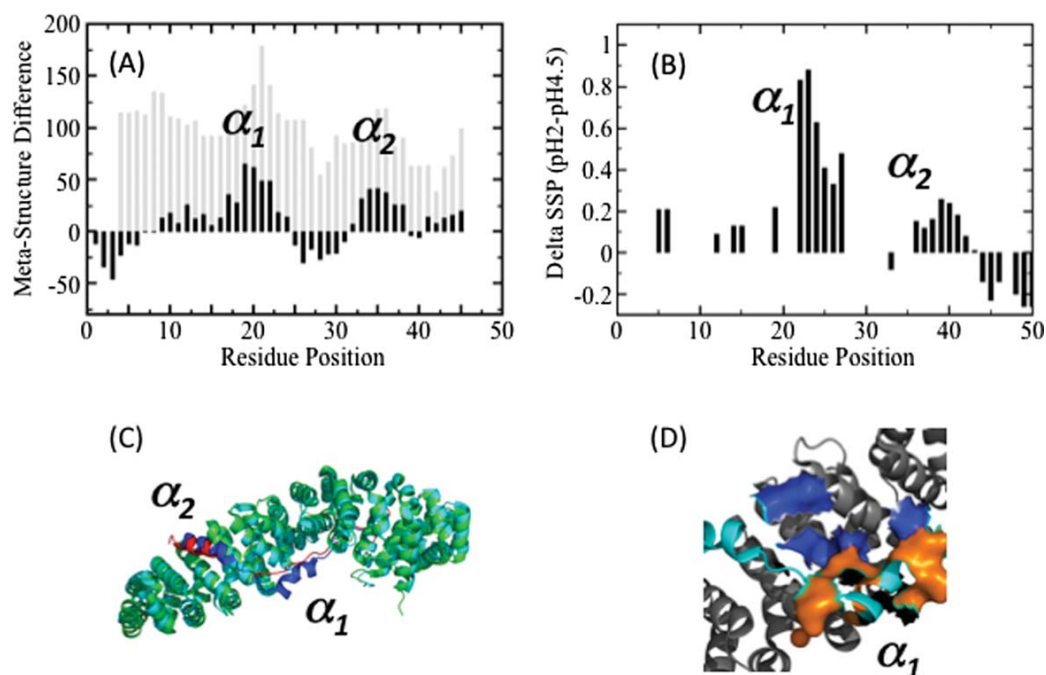
The second IDP analyzed was the DNA-binding T-cell factor/lymphoid enhancer factor (Tcf4/LEF), a prominent binding partner of  $\beta$ -catenin.  $\beta$ -Catenin is a central unit in the Wnt signaling pathway, and constitutive activation of the Wnt signaling pathway is involved in the development of various human malignancies, like colorectal carcinomas, melanomas, and ovarian carcinomas.<sup>28</sup> Although detailed structural information of unbound Tcf4 is still lacking, it was shown by CD spectroscopy that Tcf4 in



**Figure 3.** Experimental verification of meta-structure derived structural changes of BASP1 on lowering pH. (A) Differences in meta-structure derived local secondary structure (black) and compactness values (gray) between acidic and neutral conditions. (B) Schematic illustration of coarse-grained structural changes of BASP1 on changes in solution pH, as detected by DEER spectroscopy. On lowering pH, the two labelling sites (C110 and C229) spatially approach, indicating compaction of the IDP (top). This is evident from time traces (lower left) at pH 6 (red) and at pH 1 (blue) and the corresponding distance distributions (lower right). (C) Comparison between meta-structure derived secondary structure changes (upper trace), NMR derived secondary structure propensities (SSP) changes (middle trace) and differential local conformational dynamics probed by  $^{15}\text{N}$   $T_2$  relaxation on lowering pH (lower trace).

its apo-state is devoid of significant secondary structure elements and is reminiscent of a random coil.<sup>29</sup> This lack of global structure is also reflected in small meta-structure derived compactness. The average residue compactness of Tcf4 was found to be about 54, which is considerably smaller than the average value for a stably structured, globular protein (about 300).<sup>13</sup> Meta-structure calculations performed under acidic conditions revealed that at low pH Tcf4 is more compact and significantly more  $\alpha$ -

helical [Fig. 4(A)]. To verify this prediction, NMR signal assignment was performed at neutral and low pH (2.0). The increase in  $\alpha$ -helical content at low pH was clearly observed in the SSP analysis [Fig. 4(B)]. Interestingly, the two  $\alpha$ -helical regions ( $\alpha_1$  and  $\alpha_2$ ) displaying more  $\alpha$ -helical propensities under acidic than under neutral conditions constitute major parts of the Tcf4 interaction interface and provide most of the binding energy for the  $\beta$ -Catenin/Tcf4 complex [Fig. 4(C)]. It is particularly noteworthy, that the



**Figure 4.** NMR data and meta-structure analysis of Tcf4. (A) Differences in meta-structure derived local secondary structure (black) and compactness values (gray) between low and high pH conditions. (B) Differences of NMR-derived secondary structure propensities (Delta SSP) derived under acidic conditions and neutral pH as a function of residue position. Overall, only small  $^{13}\text{C}\alpha$  secondary chemical shift deviations from random coil values are observed for Tcf4 under neutral conditions indicative of marginal residual local secondary structure in Tcf4. Significant increases in  $^{13}\text{C}\alpha$  secondary chemical shifts on lowering the pH were identified in the recognition helices  $\alpha_1$  and  $\alpha_2$ . Missing peaks are due to low signal-to-noise ratios resulting from considerable chemical exchange of amide protons with bulk water. (C) Conformational plasticity of the Tcf4 interaction region. Overlay of 3D structures of  $\beta$ -catenin/Tcf4 protein complexes (1JPW, 1JDH).  $\beta$ -Catenin and Tcf4 molecules are shown in green and red (1JPW)<sup>39</sup> or cyan and blue (1JDH).<sup>40</sup> Interestingly, conformational plasticity (polymorphism, see text) was found for the recognition element  $\alpha_1$ . (D) The polymorphic recognition element  $\alpha_1$  of Tcf4 binds into a basic  $\beta$ -catenin-binding cleft (1JDH).<sup>40</sup> Negatively charged Tcf4 (cyan) residues located in the recognition element  $\alpha_1$  are shown in orange, whereas  $\beta$ -catenin (gray) lysine and arginine residues located in the basic binding cleft are shown as blue surfaces.

region  $\alpha_1$  exhibits the largest SSP changes [Fig. 4(B,D)] and thus points to a significantly larger conformational plasticity of this part of Tcf4.

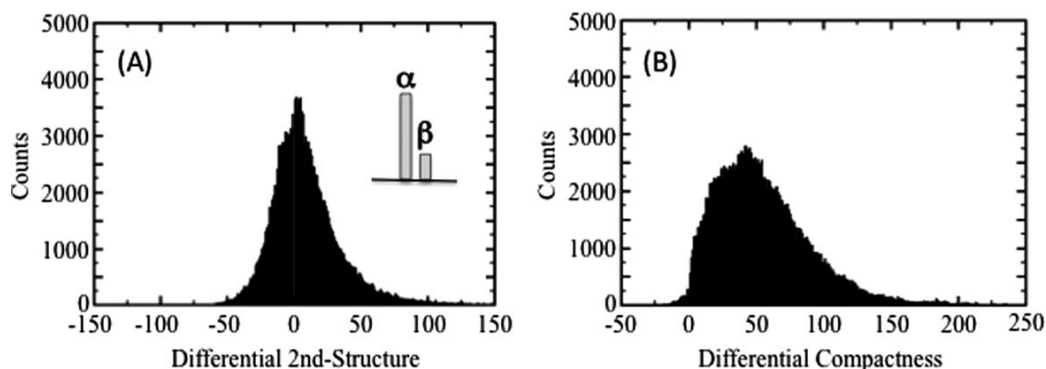
Finally, to analyze whether the observed pH-dependent conformational changes might constitute a conserved feature within the IDP domain, we performed a large-scale analysis of human IDPs. Because the average mean compactness value for a protein in the PDB is about 300,<sup>13</sup> we used a threshold of 200 for the selection of human IDPs. Only proteins with mean compactness values smaller than this threshold were considered (in total 1013 proteins were analyzed). We thus only consider globally disordered proteins and exclude globular proteins with (long) unstructured linker segments.

For this set of human IDPs, we calculated local secondary structure and compactness values for high and low pH forms. The histogram of pair-wise secondary structure and compactness differences is shown in Figure 5. Overall, we found a slight but significant preponderance of  $\alpha$ -helical content at low pH and increased compactness values on decreasing

pH seem to be a conserved feature for IDPs. Moreover, considering only longer residue stretches (larger than 7 residues) with consistent local secondary structure changes, the preponderance of  $\alpha$ -helical structures at low pH becomes even more pronounced [Fig. 5(A), insert].

## Discussion

IDPs are intriguingly sensitive to changing environmental conditions. Given that local changes of pH are of high relevance to *in vivo* functionalities of these protein species, local pH levels or proton concentrations are tightly controlled in living organisms. Although detailed studies of the effect of pH on IDP conformational space are rare, recent experimental findings indicate that a decreasing pH is likely to be accompanied by the formation of  $\alpha$ -helices in glutamic acid-rich regions and a subsequent folding transition to a non-molten globule state for which the hydrophobic collapse is facilitated by neutralization of negatively charged side-chains. In this study, we have introduced a novel



**Figure 5.** Large-scale analysis of pH dependencies of (1013) human IDPs. Histograms of meta-structural changes for individual residues are shown. (A) Histogram of individual local secondary structure changes on lowering pH. Positive and negative deviations are indicative for larger  $\alpha$ -helical or  $\beta$ -strand propensities. The slightly skewed distribution suggests a more pronounced tendency for IDPs to form  $\alpha$ -helical structures at low pH conditions. This effect becomes more visible when considering meta-structure deviations for a continuous stretch of seven residues (corresponding to two  $\alpha$ -helical turns). Relative counts for longer (larger than seven residues) continuous  $\alpha$ -helical and  $\beta$ -sheet regions are shown in the inset (see text). The formation of  $\alpha$ -helical structures is approximately three times more favourable compared to  $\beta$ -strand elements. (B) Histogram of individual compactness changes on lowering pH. Positive deviations indicate structural compaction of the polypeptide chain.

meta-structure approach that can be used to predict the pH dependence of IDPs using only primary sequence information. The pH-dependent differences in meta-structure derived local structure and compactness values were found to be in good agreement with experimental findings. Data obtained on a large set of intrinsically disordered human proteins show that IDPs fold under acidic conditions into more compact structures with higher  $\alpha$ -helical content largely due to reduced electrostatic repulsion of negatively charged side chains. It was shown recently that the inherent context-independent  $\alpha$ -helical propensities are significantly higher for Asn and Gln than for Asp and Glu, respectively.<sup>30</sup> The meta-structure derived increases in  $\alpha$ -helical propensities are not only related to individual (context-independent) amino acid specific properties but also take into account context-dependent (primary sequence) influences.

This observation is also particularly interesting in view of a recent finding showing that the evolution of IDP sequences is constrained and involves a significant bias in the amino acid composition.<sup>31</sup> In contrast to globular, stably folded proteins structural compaction in IDPs is accompanied by the formation of  $\alpha$ -helices and characterized by a low  $\beta$ -structure propensity.<sup>31</sup> Therefore, avoiding aggregation despite preformed molecular recognition elements is an important driving force in protein sequence evolution of IDPs. We showed that similar mechanisms are responsible for maintaining soluble, functionally active and nonaggregating IDPs on protonation or under acidic conditions.

Furthermore, as illustrated with the  $\beta$ -catenin/Tcf4 protein complex, stabilization of preformed  $\alpha$ -helices via (full or partial) protonation might also be

of relevance for protein complex formation involving IDPs. Although globular proteins typically involve autonomously structured domains of about 60–300 amino acid lengths, IDPs mediate protein interaction using short sequence motifs. Depending on the context, they are called eukaryotic linear motifs,<sup>32</sup> short linear motifs,<sup>33</sup> molecular recognition features,<sup>34</sup> or preformed structural elements (PSEs),<sup>35</sup> although they are not entirely unrelated and share analogous features. Although protein recognition by IDPs often proceeds via folding-on-binding events (e.g., disorder-to-order transitions),<sup>36</sup> there is growing evidence that even in the bound state IDPs (can) retain substantial conformational flexibilities, a conceptual view which is sometimes referred to as “fuzziness”.<sup>37</sup> The Tcf4/ $\beta$ -catenin protein complex is an example for static disorder (“polymorphic”) as Tcf4 can bind in several distinct conformations to  $\beta$ -catenin.<sup>38</sup> Specifically, the Tcf4/ $\beta$ -Catenin binding interface comprises three epitopes centered at Tcf4 residues D16, E29 and D43 (PDB: 1JPW, 1JDH).<sup>39,40</sup> The central epitope (E29),  $\alpha_1$ , shows considerable plasticity and docks to  $\beta$ -Catenin in several different conformations (extended or  $\alpha$ -helical). Inspection of the Tcf4/ $\beta$ -catenin structure (PDB: 1JPW, 1JDH)<sup>39,40</sup> reveals that in this complex epitope  $\alpha_1$  is surrounded by a basic patch comprising the positively charged  $\beta$ -Catenin residues K312, R342, K345, R376, and R386. The finding that protonation of negatively charged side-chains stabilizes  $\alpha$ -helical structures provides an additional chemical mechanism for stabilizing recognition elements in IDP protein complexes. Favorable electrostatic interactions between the basic proton acceptor located at the surface of the IDP recognition element and the acidic proton donor of the interaction partner lead to stabilization



of the partially PSE. However, it should be noted that to stabilize the basic proton acceptor complete proton transfer is not required, as even the realization of hydrogen bonds (e.g., partial proton transfer) can already lead to stabilization of the system. Based on our results obtained for the Tcf4/ $\beta$ -catenin protein complex we conclude that tuning of protonation efficiency might constitute a relevant chemical mechanism to modulate interaction affinities and associated recognition selectivity.

As a conclusion, the availability of a fast and reliable prediction tool for preformed  $\alpha$ -helical elements sensitive to protonation-induced stabilization may be of great relevance for the (sequence-based) identification of folding-on-binding systems in different organisms and the analysis of recognition patterns with potential implications for large-scale predictions of IDP interaction partners.

## Materials and Methods

### Meta-structure description of proteins under acidic conditions

The pH dependence of IDPs was analyzed applying the recently developed meta-structure approach.<sup>13</sup> In this study, we focus on the assessment of protonation events on the meta-structural parameters of IDPs. In brief, protein 3D structural information is converted into a network structure in which a node refers to a particular residue and edges indicate the existence of neighbourhood relationships (provided that the C $\alpha$ -C $\alpha$  distances are below a distance cutoff, typically 8 Å). Subsequently, the topological relationship between two residues is quantified by the shortest path length  $\theta$  connecting these two residues in the network. The topological relationship between two residues (of type A and B) characteristically depends on the amino acid types and their primary sequence distance,  $l_{AB}$ . The statistical distribution functions  $\rho(\theta, A, B, l_{AB})$  were extracted from a selected subset of PDB structures.<sup>13</sup> The statistical distribution functions  $\rho(\theta, A, B, l_{AB})$  can subsequently be used to predict topological information solely based on the primary sequence. The input primary sequence is used to predict for each possible amino acid pair (of residue type A and B and separated by  $l_{ab}$  in the sequence) in the protein and based on  $\rho(\theta, A, B, l_{AB})$  an average topology parameter,  $e_{ij}$ . Typically primary sequence separations  $l_{AB} > 5$  are not further differentiated and considered long-range. In brief, the topological parameter  $e_{ij}$  is related to the probability of finding a shortest path length  $\theta_{ij}=1$  (direct neighbour contact). Summation of all pairwise contributions  $e_{ij}$  leads to the compactness value  $C_i$  of residue  $i$ . Large compactness  $C_i$  values are found for residues embedded in the interior of the structure and surrounded by many neighbouring residues whereas small (even negative) values are found for

flexible residues located on the surface and exposed to the solvent. The local secondary structure of proteins is assessed by using statistical distribution function  $\rho(\theta, A, B, l_{AB})$  with sequence separations  $l_{AB} \leq 4$ . The local secondary structure parameter  $S_i$  is defined as  $S_i = \alpha P_i - \beta P_i$ , where  $\alpha P_i$  and  $\beta P_i$  are defined as:  $\alpha P_i = N_\alpha * \gamma_2 * \gamma_3 * \gamma_4$ ,  $\beta P_i = N_\beta * \delta_3 * \delta_4$ .  $\gamma_i = \rho(1, i) * (1.0 - \rho(2, i) / [\rho(2, i) * (1.0 - \rho(1, i))])$ ;  $\delta_i = Y_i^{-1}$ .  $\rho(1, i)$  and  $\rho(2, i)$  denote the probabilities of finding shortest path length values of 1 and 2 between residues  $i$  and  $j$ .  $N_\alpha$  and  $N_\beta$  are empirical constants ensuring comparable compactness and local secondary structure values.<sup>13</sup> Typically, residues located in  $\alpha$ -helices display positive  $S_i$  values whereas for residues located in extended regions ( $\beta$ -strands or polyproline II helices) significantly smaller (negative) values were observed. Of particular relevance is the fact, that this sequence analysis provides quantitative information about the local secondary structure and residue compactness for each residue position.

The statistical distribution functions  $\rho(\theta, A, B, l_{AB})$  used for the calculation of meta-structural parameters were extracted from proteins predominantly crystallized under near neutral pH conditions. Under these conditions the carboxylic groups of Asp and Glu side-chains are predominantly deprotonated (negatively charged). The residue types in  $\rho(\theta, A, B, l_{AB})$  thus refer to negatively charged Asp and Glu side-chains. Because distribution functions for protonated Asp and Glu side-chains are not directly accessible through crystal structure analysis, we mimic protonation events by replacing Asp and Glu side-chains by Asn and Gln in the primary sequence. The rationale for our approach is as follows: Firstly, this amino acid replacement ensures electroneutrality (on protonation of carboxylate) of the side-chain. Secondly, the Asn/Gln amide groups are chemically comparable to protonated Asp/Glu carboxyl groups. This is due to the fact that both, amide groups in Asn/Gln and carboxyl groups of protonated (neutral) Asp and Glu, are planar and can act as hydrogen bond donors. The respective meta-structure parameters are calculated for both wild-type (negatively charged Asp/Glu side chains) and “mutated” protein (neutral Asn/Gln) forms. Subsequent differences in meta-structure derived local structure and compactness values are correlated with pH-induced alterations of local backbone geometry and compaction of the polypeptide chain.

### NMR spectroscopy and signal assignment

All spectra were acquired at 298 K on an Agilent Direct Drive 700 MHz (BASP1) or Varian Unity 500 and 600 MHz (Tcf4, 1 mM, pH 4.5 and 2) spectrometer using standard 5 mm  $^1\text{H}$ - $^{13}\text{C}$ - $^{15}\text{N}$  triple-resonance probe heads. NMR data for Tcf4 were obtained on spectrometers at 298 K. Signal assignment was achieved as described elsewhere.<sup>41</sup> PFG

NMR diffusion measurements were performed and analyzed as described elsewhere.<sup>42</sup>

### EPR spectroscopy

Double electron-electron (DEER)<sup>27,43,44</sup> was applied to glassy solids obtained by freeze-quenching aqueous solutions of BASP1 double mutants (MTSL labeling at Cys-mutated residues 110 and 229 of the chicken analogue of human BASP1) after addition of 30 v/v % glycerol. More details of the experimental procedure are given in the Supporting Information.

### Acknowledgments

M.A.H. acknowledges the support of Professor Dr. Hassan Eisa, Dr. Alaa El-Din Barghash, and Dr. Laila A. Abouzeid; Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University, Egypt.

### References

1. Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27:527–533.
2. Fink AL (2005) Natively unfolded proteins. *Curr Opin Struct Biol* 15:35–41.
3. Uversky VN, Oldfield CJ, Midic U, Xie H, Xue B, Vucetic S, Iakoucheva LM, Obradovic Z, Dunker AK (2009) Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* 10 Suppl 1:S7.
4. Uversky VN (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J* 28:305–325.
5. Uversky VN, Dunker AK (2010) Understanding protein non-folding. *Biochim Biophys Acta* 1804:1231–1264.
6. Uversky VN (2013) Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* 1834:932–951.
7. Marsh JA, Forman-Kay JD (2010) Sequence determinants of compaction in intrinsically disordered proteins. *Biophys J* 98:2383–2390.
8. Cho MK, Nodet G, Kim HY, Jensen MR, Bernado P, Fernandez CO, Becker S, Blackledge M, Zweckstetter M (2009) Structural characterization of alpha-synuclein in an aggregation prone state. *Protein Sci* 18:1840–1846.
9. Uversky VN, Gillespie JR, Millett IS, Khodyakova AV, Vasiliev AM, Chernovskaya TV, Vasilenko RN, Kozovskaya GD, Dolgikh DA, Fink AL, Doniach S, Abramov VM (1999) Natively unfolded human prothymosin alpha adopts partially folded collapsed conformation at acidic pH. *Biochemistry* 38:15009–15016.
10. Paroutis P, Touret N, Grinstein S (2004) The pH of the secretory pathway: measurement, determinants, and regulation. *Physiology* 19:207–215.
11. Warburg O (1956) Origin of cancer cells. *Science* 123:309–314.
12. Schultz ML, Tecedor L, Chang M, Davidson BL (2011) Clarifying lysosomal storage diseases. *Trends Neurosci* 34:401–410.
13. Konrat R (2009) The protein meta-structure: a novel concept for chemical and molecular biology. *Cell Mol Life Sci* 66:3625–3639.
14. Henen MA, Coudeville N, Geist L, Konrat R (2012) Toward rational fragment-based lead design without 3D structures. *J Med Chem* 55:7909–7919.
15. Mayer C, Slater L, Erat MC, Konrat R, Vakoniakis I (2012) Structural analysis of the Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J Biol Chem* 287:7182–7189.
16. Schedlbauer A, Gandini R, Kontaxis G, Paulmichl M, Furst J, Konrat R (2011) The C-terminus of ICLn is natively disordered but displays local structural preformation. *Cell Physiol Biochem* 28:1203–1210.
17. Hasenohrl D, Konrat R, Blasi U (2011) Identification of an RNase J ortholog in Sulfolobus solfataricus: implications for 5'-to-3' directional decay and 5'-end protection of mRNA in Crenarchaeota. *RNA* 17:99–107.
18. Hartl M, Nist A, Khan MI, Valovka T, Bister K (2009) Inhibition of Myc-induced cell transformation by brain acid-soluble protein 1 (BASP1). *Proc Natl Acad Sci U S A* 106:5604–5609.
19. Clevers H, Nusse R (2012) Wnt/beta-catenin signaling and disease. *Cell* 149:1192–1205.
20. Kontaxis G, Delaglio F, Bax A (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol* 394:42–78.
21. Dickerson RE, Geis I (1969) The structure and action of proteins. New York: Harper & Row.
22. Uversky VN, Li J, Fink AL (2001) Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J Biol Chem* 276:10737–10744.
23. Korshunova I, Caroni P, Kolkova K, Berezin V, Bock E, Walmod PS (2008) Characterization of BASP1-mediated neurite outgrowth. *J Neurosci Res* 86:2201–2213.
24. Carpenter B, Hill KJ, Charalambous M, Wagner KJ, Lahiri D, James DI, Andersen JS, Schumacher V, Royer-Pokora B, Mann M, Ward A, Roberts SG (2004) BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1. *Mol Cell Biol* 24:537–549.
25. Green LM, Wagner KJ, Campbell HA, Addison K, Roberts SG (2009) Dynamic interaction between WT1 and BASP1 in transcriptional regulation during differentiation. *Nucleic Acids Res* 37:431–440.
26. Schanda P, Forge V, Brutscher B (2006) HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn Reson Chem* 44 Spec No:S177–S 184.
27. Jeschke G, Chechik V, Ionita P, Godt A, Zimmermann H, Banham J, Timmel CR, Hilger D, Jung H (2006) Deera-analysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Appl Magn Reson* 30:473–498.
28. Kobayashi M, Honma T, Matsuda Y, Suzuki Y, Narisawa R, Ajioka Y, Asakura H (2000) Nuclear translocation of beta-catenin in colorectal cancer. *Br J Cancer* 82:1689–1693.
29. Knapp S, Zamai M, Volpi D, Nardese V, Avanzi N, Breton J, Plyte S, Flocco M, Marconi M, Isacchi A, Caiola VR (2001) Thermodynamics of the high-affinity interaction of TCF4 with beta-catenin. *J Mol Biol* 306:1179–1189.
30. Moreau RJ, Schubert CR, Nasr KA, Torok M, Miller JS, Kennedy RJ, Kemp DS (2009) Context-independent, temperature-dependent helical propensities for amino acid residues. *J Am Chem Soc* 131:13107–13116.
31. Naranjo Y, Pons M, Konrat R (2012) Meta-structure correlation in protein space unveils different selection rules for folded and intrinsically disordered proteins. *Mol Biosyst* 8:411–416.
32. Neduva V, Russell RB (2005) Linear motifs: evolutionary interaction switches. *FEBS Lett* 579:3342–3345.

33. Davey NE, Shields DC, Edwards RJ (2006) SLIMDisc: short, linear motif discovery, correcting for common evolutionary descent. *Nucleic Acids Res* 34:3546–3554.
34. Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK (2007) Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res* 6:2351–2366.
35. Fuxreiter M, Simon I, Friedrich P, Tompa P (2004) Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* 338:1015–1026.
36. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *J Mol Biol* 293:321–331.
37. Tompa P, Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* 33:2–8.
38. Hazy E, Tompa P (2009) Limitations of induced folding in molecular recognition by intrinsically disordered proteins. *ChemPhysChem* 10:1415–1419.
39. Poy F, Lepourcelet M, Shivdasani RA, Eck MJ (2001) Structure of a human Tcf4-beta-catenin complex. *Nat Struct Biol* 8:1053–1057.
40. Graham TA, Ferkey DM, Mao F, Kimelman D, Xu W (2001) Tcf4 can specifically recognize beta-catenin using alternative conformations. *Nat Struct Biol* 8:1048–1052.
41. Geist L, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Kozminski W, Konrat R (2012) (1)H, (13)C and (15)N resonance assignments of human BASP1. *Biomol NMR Assign*; doi: 10.1007/s12104-012-9436-4.
42. Platzer G, Schedlbauer A, Chemelli A, Ozdowy P, Coudeville N, Auer R, Kontaxis G, Hartl M, Miles AJ, Wallace BA, Glatter O, Bister K, Konrat R (2011) The metastasis-associated extracellular matrix protein osteopontin forms transient structure in ligand interaction sites. *Biochemistry* 50:6113–6124.
43. Jeschke G, Koch A, Jonas U, Godt A (2002) Direct conversion of EPR dipolar time evolution data to distance distributions. *J Magn Reson* 155:72–82.
44. Martin RE, Pannier M, Diederich F, Gramlich V, Hubrich M, Spiess HW (1998) Determination of end-to-end distances in a series of TEMPO diradicals of up to 2.8 nm length with a new four-pulse double electron resonance experiment. *Angew Chem Int Ed* 37:2834–2837.

## Supplementary Material

### Protonation-Dependent Conformational Variability of Intrinsically Disordered Proteins

Leonhard Geist<sup>1</sup>, Morkos A Henen<sup>1,4</sup>, Sandra Haiderer<sup>1</sup>, Thomas Schwarz<sup>1</sup>,  
Dennis Kurzbach<sup>2</sup>, Anna Zawadzka-Kazimierczuk<sup>3</sup>, Saurabh Saxena<sup>3</sup>,  
Szymon Żerko<sup>3</sup>, Wiktor Koźmiński<sup>3</sup>, Dariusz Hinderberger<sup>2</sup> and Robert  
Konrat<sup>1\*</sup>

*<sup>1</sup>Department of Computational and Structural Biology, Max F. Perutz  
Laboratories, University of Vienna, Campus Vienna Biocenter 5, A-1030  
Vienna, Austria*

*<sup>2</sup>Max Planck Institute for Polymer Research, Ackermannweg 10, 55128  
Mainz, Germany*

*<sup>3</sup>Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw,  
Poland*

*<sup>4</sup>Faculty of Pharmacy, Mansoura University, Egypt*

\*To whom correspondence should be addressed. E-mail:  
[robert.konrat@univie.ac.at](mailto:robert.konrat@univie.ac.at); Phone : ++43-(0)1-4277-52202 ; FAX : ++43-(0)1-  
4277-9522



## EPR spectroscopy

Freeze-quenching was achieved by immersing the sample tubes in supercooled iso-pentane. Such, a snapshot representative for the solution at the glass transition temperature is detected. The sample volume was always large enough to fill the complete resonator. The four pulse DEER sequence  $\pi/2(\nu_{\text{obs}}) - \tau_1 - \pi(\nu_{\text{obs}}) - (\tau_1 + t) - (\nu_{\text{pump}}) - (\tau_2 - t) - \pi(\nu_{\text{obs}}) - \tau_2 - \text{echo}$  was used to obtain dipolar time evolution data at X-band frequencies (9.2 to 9.4 GHz) with a Bruker Elexsys 580 spectrometer equipped with a Bruker Flexline split-ring resonator ER4118X\_MS3. The dipolar evolution time  $t$  was varied, whereas  $\tau_2 = 4 \mu\text{s}$  and  $\tau_1$  were kept constant. Proton modulation was averaged by the addition of eight time traces of variable  $\tau_1$ , starting with  $\tau_{1,0} = 200 \text{ ns}$  and incrementing by  $\Delta\tau_1 = 8 \text{ ns}$ . The resonator was overcoupled to  $Q \approx 100$ . The pump frequency,  $\nu_{\text{pump}}$ , was set to the maximum of the EPR spectrum. The observer frequency,  $\nu_{\text{obs}}$ , was set to  $\nu_{\text{pump}} + 61.6 \text{ MHz}$ , coinciding with the low field local maximum of the nitroxide spectrum. The observer pulse lengths were 32 ns for both  $\pi/2$  and  $\pi$  pulses, and the pump pulse length was 12 ns. The temperature was set to 50 K by cooling with a closed cycle cryostat (ARS AF204, customized for pulse EPR, ARS, Macungie, PA). The total measurement time for each sample was around 12 h. Intermolecular contributions to the DEER time-traces were removed through dividing by an experimentally (DEER on singly labeled BASP1) determined background function. The raw time domain DEER data were processed with the program package DeerAnalysis2010 (Jeschke et al., 2006). The resulting time traces were normalized to  $t = 0$ . Intermolecular contributions to the DEER time-traces were removed through dividing by an experimentally (DEER on singly

labeled BASP1) determined background function. Distance distributions,  $P(r)$ , were determined by fits based on Gaussian distance distributions of the background-corrected time traces.

# ***Article 2***

# $^1\text{H}$ , $^{13}\text{C}$ and $^{15}\text{N}$ resonance assignments of human BASP1

Leonhard Geist · Anna Zawadzka-Kazimierczuk ·  
Saurabh Saxena · Szymon Żerko · Wiktor Koźmiński ·  
Robert Konrat

Received: 27 June 2012 / Accepted: 31 October 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** Brain acid-soluble protein 1 (BASP1, CAP-23, NAP-22) appears to be implicated in diverse cellular processes. An N-terminally myristoylated form of BASP1 has been discovered to participate in the regulation of actin cytoskeleton dynamics in neurons, whereas non-myristoylated nuclear BASP1 acts as co-suppressor of the potent transcription regulator WT1 (Wilms' Tumor suppressor protein 1). Here we report NMR chemical shift assignment of recombinant human BASP1 fused to an N-terminal cleavable His6-tag.

**Keywords** BASP1 · NMR signal assignment · Intrinsically disordered protein · Myc oncogene · WT1

## Biological context

The homologues of human BASP1 were first identified as the brain-specific proteins CAP-23 (cortical cytoskeleton-associated protein) in chicken brain (Widmer and Caroni 1990) and its rat homologue NAP-22 (neuron-specific acidic protein; Maekawa et al. 1993). Human BASP1 was originally isolated from neuronal cells (Mosevitsky 2005). Interestingly it appears to fulfill quite diverse tasks in the cell. N-myristoylated BASP1 has been described to be

involved in neurite outgrowth and plasma membrane organization (Korshunova et al. 2008). It is able to interact with the inner leaflet of the plasma membrane via its myristoyl-anchor and sequesters Phosphatidyl-inositol-4,5-diphosphate (PIP2) into lipid rafts (Epand et al. 2004; Shaw et al. 2006). Recently it has been shown that liposomes containing anionic phospholipids induce oligomerization of BASP1. Interaction with calmodulin is followed by dissociation of BASP1 from the membrane and disruption of the oligomers (Zakharov and Mosevitsky 2010). Additionally, BASP1 is under the control of protein kinase C (PKC), which phosphorylates BASP1 at Ser5. It is hypothesized that phosphorylation leads to the disruption of the interaction of the N-terminal positive effector domain of BASP1 with anionic phospholipids (Laux et al. 2000).

Furthermore, non-myristoylated BASP1 appears to influence transcription regulation in the nucleus, greatly affecting the differentiation pathway of a cell. It has been discovered as a co-suppressor of WT1 function (Wilms' Tumor suppressor protein 1) exerting its function by interacting with an N-terminal suppression domain of WT1 (Carpenter et al. 2004; Green et al. 2009). WT1 itself is a potent transcriptional regulator that activates or represses target genes including those for growth factors and regulators of cell division (Wagner and Roberts 2004). Aberrant expression of WT1 is associated with several childhood and adult cancers (Rivera and Haber 2005; Yang et al. 2007). Additionally, a recent study discovered BASP1 to be downregulated in v-Myc-transformed chicken fibroblasts. Strikingly, ectopic expression of BASP1 renders fibroblasts resistant to subsequent cell transformation by v-Myc and it has been shown that the inhibition of v-Myc-induced cell transformation by BASP1 affects the transcriptional regulation of Myc target genes (Hartl et al. 2009). Other findings, reporting the frequent down-regulation of BASP1

L. Geist · R. Konrat (✉)  
Department of Computational and Structural Biology, Max F.  
Perutz Laboratories, University of Vienna, Campus Vienna  
Biocenter 5, 1030 Vienna, Austria  
e-mail: robert.konrat@univie.ac.at

A. Zawadzka-Kazimierczuk · S. Saxena · S. Żerko ·  
W. Koźmiński  
Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093  
Warsaw, Poland

**Table 1** Maximum evolution times (tmax, ms) and spectral width (sw, kHz) used for acquisition of spectra for H6-hBASP1

	3D HNCO	4D HabCabCONH	5D HNCOCACB	5D HN(CA) CONH	5D H(CC-tocsy) CONH	5D (H)NCO(NCA) CONH	5D (HACA)CON(CA) CONH
Number of points	700	550	550	550	650	610	610
Experiment duration (hours)	4	13	13	13	17	14	14
sw <sub>1</sub>	2	4	14	6	8	2.5	3.8
sw <sub>2</sub>	2.5	14	14	2.5	18	2	2
sw <sub>3</sub>		2	2	2	2	2	2
sw <sub>4</sub>		2.5	2.5	2.5	2.5	2.5	3.8
t <sub>1</sub> <sup>max</sup>	100	20	10	25	10	50	50
t <sub>2</sub> <sup>max</sup>	150	7.1	10	50	10	45	45
t <sub>3</sub> <sup>max</sup>		45	45	50	45	45	45
t <sub>4</sub> <sup>max</sup>		75	75	75	75	75	75
Sampling density versus conventional	$9.3 \times 10^{-3}$	$4.1 \times 10^{-6}$	$1.7 \times 10^{-6}$	$1.6 \times 10^{-6}$	$2.7 \times 10^{-6}$	$3.2 \times 10^{-8}$	$1.4 \times 10^{-7}$

expression in ALL (acute lymphocytic leukaemia) and CLL (chronic lymphocytic leukaemia) (Yeoh et al. 2002; Wang et al. 2004), as well as apoptosis-induced cleavage of BASP1 and its subsequent translocation to the cytoplasm (Ohsawa et al. 2008), again highlight the importance of BASP1 in transcription regulation.

To provide molecular information about this potential tumour suppressor protein we have started the NMR structure determination of recombinant human BASP1. The near complete chemical shift assignment reveals that BASP1 belongs to the class of intrinsically disordered proteins.

## Methods and results

### Protein expression and purification

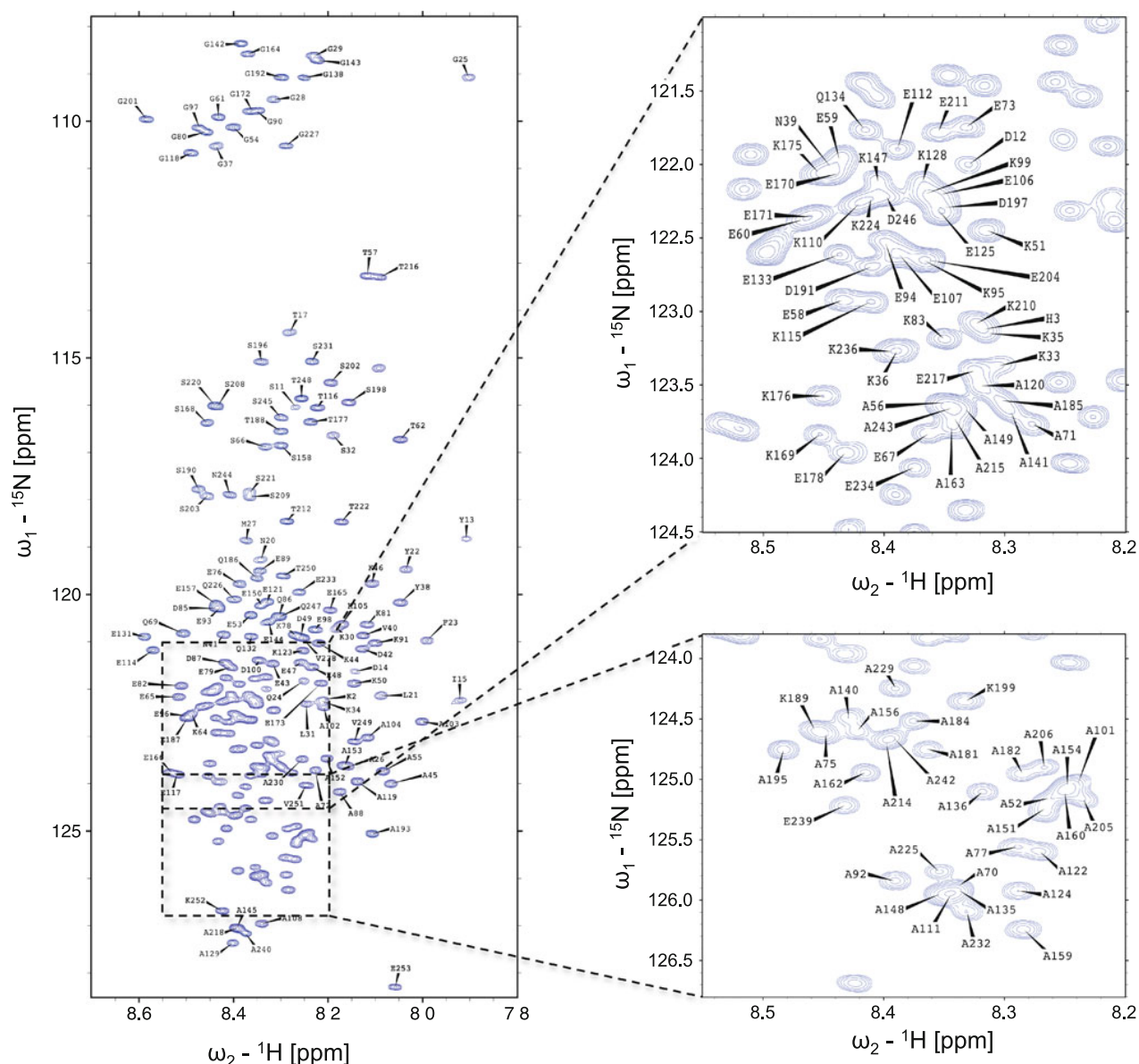
The coding region for hBASP1 (human BASP1) was amplified by PCR from the mammalian expression vector Flag-hBASP1-pTKX3 (Ohsawa et al. 2008) introducing a 5' NcoI and 3' NotI site. Subsequently the fragment was inserted in-frame into the NcoI and NotI sites of the bacterial expression vector pET-M11 (Pinotsis et al. 2006), yielding pET-M11-hBASP1, encoding hBASP1 fused to an N-terminal His6-tag plus the TEV-cleavage site (H6-hBASP1). <sup>15</sup>N/<sup>13</sup>C labeled H6-hBASP1 was expressed in the *E. coli* strain Rosetta(DE3)pLysS following a new expression protocol for efficient isotopic labeling of recombinant proteins using a fourfold cell concentration in isotopically labeled minimal medium (Marley et al. 2001). The cells were collected after 4 h of expression at 37 °C by centrifugation at 5,000 rpm for 15 min and resuspended in 40 ml of ice-cold lysis buffer (20 mM Na<sub>x</sub>H<sub>(3-x)</sub>PO<sub>4</sub>, 50 mM NaCl, 10 mM imidazole, pH

7.2) 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<sup>2+</sup> loaded HiTrap 5 ml affinity column (GE Healthcare), washed with 2 column volumes of high salt buffer (20 mM Na<sub>x</sub>H<sub>(3-x)</sub>PO<sub>4</sub>, 1.5 M NaCl, 10 mM imidazole, pH 7.2) and eluted with high imidazole buffer (20 mM Na<sub>x</sub>H<sub>(3-x)</sub>PO<sub>4</sub>, 50 mM NaCl, 0.5 M imidazole, pH 7.2) using a linear gradient of 15 column volumes. The H6-hBASP1 containing fractions were collected and the buffer was exchanged by 4 steps of concentration in an Amicon Ultra-15 centrifugal filter device 10 K NMWL (Amicon) and subsequent dilution in target buffer (20 mM Na<sub>x</sub>H<sub>(3-x)</sub>PO<sub>4</sub>, 50 mM NaCl, pH 6.0). NMR samples contain 1.5 mM uniformly <sup>15</sup>N/<sup>13</sup>C labeled protein in 20 mM sodium phosphate (pH 6.0, in 90 % H<sub>2</sub>O and 10 % D<sub>2</sub>O), 50 mM NaCl and 0.2 % sodium azide.

### NMR experiments

All spectra were acquired at 298 K on an Agilent Direct Drive 700 MHz spectrometer using the standard 5 mm <sup>1</sup>H-<sup>13</sup>C-<sup>15</sup>N triple-resonance probe head.

The backbone <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N resonances were assigned using sparse random sampling of indirectly detected time domains, in order to increase resolution. A 3D HNCO experiment was used as a base spectrum for SMFT (Sparse Multidimensional Fourier Transform) processing of higher dimensionality experiments (Kazimierczuk et al. 2010). Backbone assignment was achieved using 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) and HNCOCACB (Zawadzka-Kazimierczuk et al. 2012b)



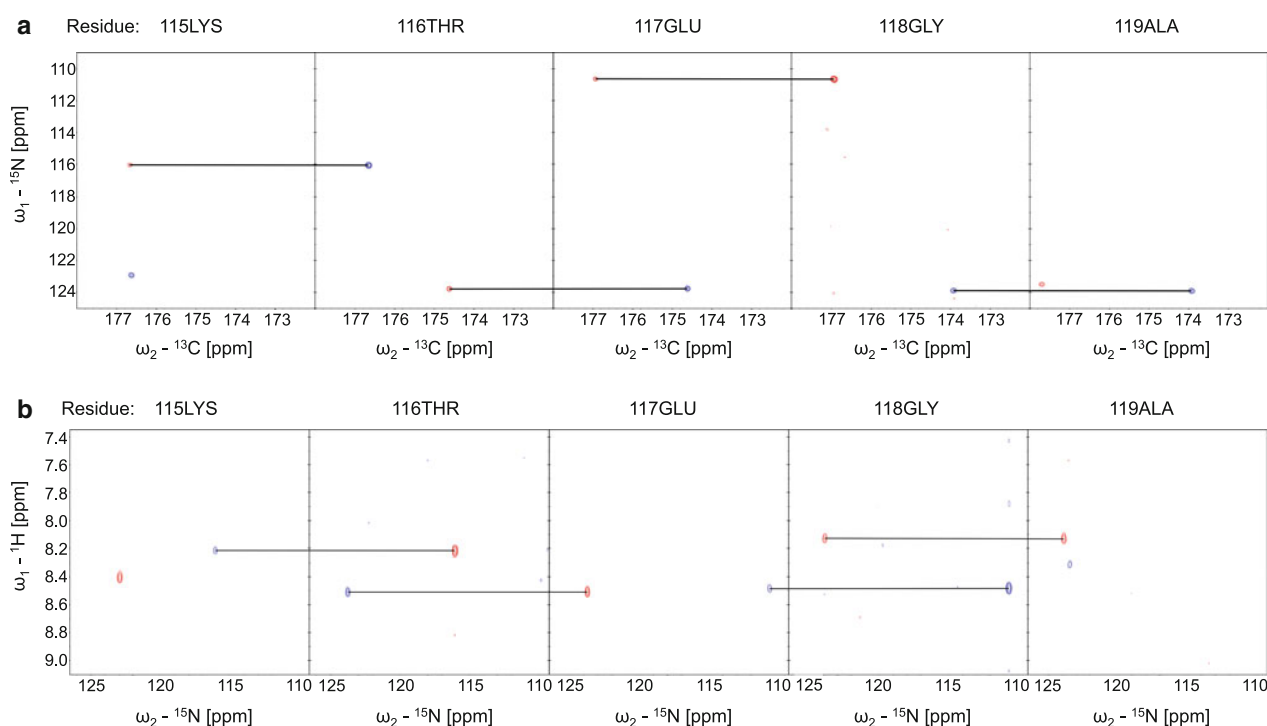
**Fig. 1**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of H6-hBASP1 at pH 6 and 298 K. Assignments of backbone amides are labeled in *single letter* amino acid code and residue number (His6-tag: 1–26; hBASP1: 27–253)

experiments. Side-chain assignments were obtained using 5D HabCabCONH (Kazimierczuk et al. 2010), and H (CC-tocsy)CONH (Kazimierczuk et al. 2009) experiments.

All NMR data sets were processed by multidimensional Fourier transformation using the home written software package (<http://nmr700.chem.uw.edu.pl/formularz.html>). The resonance assignment was performed using the TSAR program (Zawadzka-Kazimierczuk et al. 2012a). The input data for TSAR was prepared using Sparky software (Goddard and Kneller 2002). Table 1 shows the maximum evolution times and spectral width used for the acquisition of the spectra.

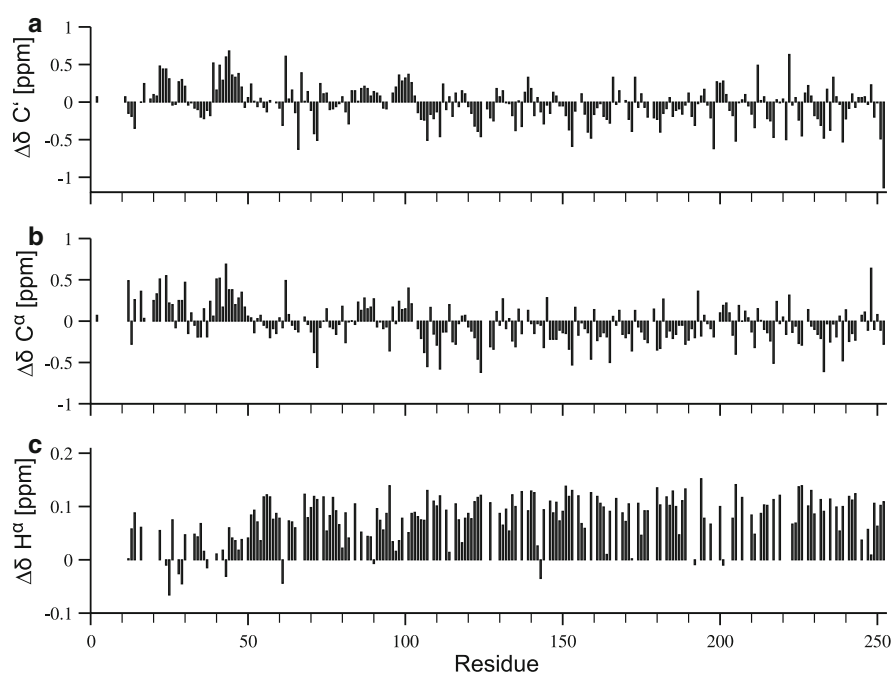
#### Extent of assignment and data deposition

The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of H6-hBASP1 shows a very narrow peak dispersion in the  $^1\text{H}$  dimension typical for intrinsically disordered proteins (Fig. 1). Extensive signal overlap in conventional 2D & 3D spectra could be overcome by using the aforementioned 5D experiments. 99 % of backbone  $^{15}\text{N}$ , 99.5 % of  $^1\text{H}$ , 96.5 % of  $^{13}\text{C}^\alpha$ , 74 % of  $^1\text{H}^\alpha$ , 86.2 % of  $^{13}\text{C}^\beta$ , 81.4 % of  $^1\text{H}^\beta$  and 98.7 % of  $^{13}\text{C}'$  resonances have been assigned. Additionally, H(CC-tocsy) CONH spectra allowed the assignment of several side-chain atoms. Figure 2 shows sample strips of sequential



**Fig. 2** 2D spectral planes for consecutive amino acids in H6-hBASP1 obtained by SMFT processing of the 5D randomly sampled signal. 2D cross-sections of **a** 5D (HACA)CON(CA)CONH ( $N_i$ - $CO_{i-1}$  &  $N_{i-1}$ - $CO_{i-2}$ ) and **b** 5D HN(CA)CONH ( $HN_i$ - $N_i$  &  $HN_{i+1}$ - $N_{i+1}$ )

**Fig. 3** Secondary chemical shifts for **a**  $^{13}C'$ , **b**  $^{13}C^\alpha$ , and **c**  $^1H^\alpha$  using sequence-specific random coil chemical shifts of intrinsically disordered proteins (Tamiola et al. 2010)



resonance assignment in a 5D (HACA)CON(CA)CONH and HN(CA)CONH experiment. Secondary chemical shifts for  $^{13}C'$ ,  $^{13}C^\alpha$ ,  $^1H^\alpha$  (Fig. 3) show only minor deviations

from random coil chemical shift values. Interestingly, the N-terminus appears to harbour stretches with slight  $\alpha$ -helical structure propensities, whereas the rest of the

protein seems to adopt a rather extended conformation indicated by positive <sup>1</sup>H $\alpha$  chemical shift differences.

The <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shifts have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>) under the BMRB accession number 18417.

**Acknowledgments** The authors thank Professor Masayuki Miura for providing the Flag-hBASP1-pTKX3 plasmid (Ohsawa et al. 2008). 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 EAST-NMR project (contract no. 228461) inside of a transnational access program (proposal acronym: SEQUASSIGNIUPBASP1) and by the Grant P 20549-N19 from the Austrian Science Foundation FWF. A.Z.-K. thanks the Foundation for Polish Science for supporting her with the MPD Programme, S.S., S.Ż., and W.K. thank the Foundation for Polish Science for support with the TEAM Programme. MPD and TEAM programs were co-financed by the EU European Regional Development.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Carpenter B, Hill KJ, Charalambous M, Wagner KJ, Lahiri D, James DI, Andersen JS, Schumacher V, Royer-Pokora B, Mann M, Ward A, Roberts SG (2004) BASP1 is a transcriptional cosuppressor for the Wilms' tumor suppressor protein WT1. *Mol Cell Biol* 24(2): 537–549
- Epand RM, Vuong P, Yip CM, Maekawa S, Epand RF (2004) Cholesterol-dependent partitioning of PtdIns(4,5)P<sub>2</sub> into membrane domains by the N-terminal fragment of NAP-22 (neuronal axonal myristoylated membrane protein of 22 kDa). *Biochem J* 379(Pt 3):527–532. doi:10.1042/BJ20040204
- Green LM, Wagner KJ, Campbell HA, Addison K, Roberts SG (2009) Dynamic interaction between WT1 and BASP1 in transcriptional regulation during differentiation. *Nucleic Acids Res* 37(2): 431–440. doi:10.1093/nar/gkn955
- Hartl M, Nist A, Khan MI, Valovka T, Bister K (2009) Inhibition of Myc-induced cell transformation by brain acid-soluble protein 1 (BASP1). *Proc Natl Acad Sci USA* 106(14):5604–5609. doi:10.1073/pnas.0812101106
- Kazimierczuk K, Zawadzka A, Kozminski W (2009) Narrow peaks and high dimensionalities: exploiting the advantages of random sampling. *J Magn Reson* 197(2):219–228. doi:10.1016/j.jmr.2009.01.003
- Kazimierczuk K, Zawadzka-Kazimierczuk A, Kozminski W (2010) Non-uniform frequency domain for optimal exploitation of non-uniform sampling. *J Magn Reson* 205(2):286–292. doi:10.1016/j.jmr.2010.05.012
- Korshunova I, Caroni P, Kolkova K, Berezin V, Bock E, Walmod PS (2008) Characterization of BASP1-mediated neurite outgrowth. *J Neurosci Res* 86(10):2201–2213. doi:10.1002/jnr.21678
- Laux T, Fukami K, Thelen M, Golub T, Frey D, Caroni P (2000) GAP43, MARCKS, and CAP23 modulate PI(4,5)P<sub>2</sub> at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. *J Cell Biol* 149(7):1455–1472
- Maekawa S, Maekawa M, Hattori S, Nakamura S (1993) Purification and molecular cloning of a novel acidic calmodulin binding protein from rat brain. *J Biol Chem* 268(18):13703–13709
- Marley J, Lu M, Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20(1):71–75
- Mosevitsky MI (2005) Nerve ending “signal” proteins GAP-43, MARCKS, and BASP1. *Int Rev Cytol* 245:245–325. doi:10.1016/S0074-7696(05)45007-X
- Ohsawa S, Watanabe T, Katada T, Nishina H, Miura M (2008) Novel antibody to human BASP1 labels apoptotic cells post-caspase activation. *Biochem Biophys Res Commun* 371(4):639–643. doi:10.1016/j.bbrc.2008.04.056
- Pinotsis N, Petoukhov M, Lange S, Svergun D, Zou P, Gautel M, Wilmanns M (2006) Evidence for a dimeric assembly of two titin/telethonin complexes induced by the telethonin C-terminus. *J Struct Biol* 155(2):239–250. doi:10.1016/j.jsb.2006.03.028
- Rivera MN, Haber DA (2005) Wilms' tumour: connecting tumorigenesis and organ development in the kidney. *Nat Rev Cancer* 5(9):699–712. doi:10.1038/nrc1696
- Shaw JE, Epand RF, Sinnathamby K, Li Z, Bittman R, Epand RM, Yip CM (2006) Tracking peptide-membrane interactions: insights from in situ coupled confocal-atomic force microscopy imaging of NAP-22 peptide insertion and assembly. *J Struct Biol* 155(3):458–469. doi:10.1016/j.jsb.2006.04.015
- Tamiola K, Acar B, Mulder FA (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc* 132(51):18000–18003. doi:10.1021/ja105656t
- Wagner KJ, Roberts SG (2004) Transcriptional regulation by the Wilms' tumour suppressor protein WT1. *Biochem Soc Trans* 32(Pt 6):932–935. doi:10.1042/BST0320932
- Wang J, Coombes KR, Highsmith WE, Keating MJ, Abruzzo LV (2004) Differences in gene expression between B-cell chronic lymphocytic leukemia and normal B cells: a meta-analysis of three microarray studies. *Bioinformatics* 20(17):3166–3178. doi:10.1093/bioinformatics/bth381
- Widmer F, Caroni P (1990) Identification, localization, and primary structure of CAP-23, a particle-bound cytosolic protein of early development. *J Cell Biol* 111(6 Pt 2):3035–3047
- Goddard TD, Kneller DG (2002) SPARKY 3. University of California, San Francisco
- Yang L, Han Y, Suarez Saiz F, Minden MD (2007) A tumor suppressor and oncogene: the WT1 story. *Leukemia* 21(5): 868–876. doi:10.1038/sj.leu.2404624
- Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR (2002) Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 1(2):133–143
- Zakharov VV, Mosevitsky MI (2010) Oligomeric structure of brain abundant proteins GAP-43 and BASP1. *J Struct Biol* 170(3): 470–483. doi:10.1016/j.jsb.2010.01.010
- Zawadzka-Kazimierczuk A, Kozminski W, Billeter M (2012a) TSAR a program for automatic resonance assignment using cross-sections of high dimensionality, high-resolution spectra. *J Biomol NMR* 54:81–95. doi:10.1007/s10858-012-9652-3
- Zawadzka-Kazimierczuk A, Kozminski W, Sanderova H, Krasny L (2012b) High dimensional and high resolution pulse sequences for backbone resonance assignment of intrinsically disordered proteins. *J Biomol NMR* 52:329–337. doi:10.1007/s10858-012-9613-x



# C U R R I C U L U M   V I T A E

**Mag. Leonhard Geist**

[leonhard.geist@univie.ac.at](mailto:leonhard.geist@univie.ac.at)

## E D U C A T I O N

- since Jan 2009:      Doctoral thesis in NMR at the Department of Structural and Computational Biology, MFPL, University of Vienna, Austria.  
Supervisor: Prof. Robert Konrat  
“Structural Dynamics of the Intrinsically Disordered Protein BASP1.”
- 13<sup>th</sup> Nov 2008:      Diploma degree (with honors)
- Oct 2007–Nov 2008: Diploma thesis at the Department of Structural and Computational Biology, MFPL, University of Vienna, Austria.  
Supervisor: Prof. Robert Konrat  
“Investigation of direct interaction between Myc and the Myc-target BASP1.”
- Oct 2004–June 2005: Internship at the Centre de Biologie du Développement at the Université Paul Sabatier in Toulouse (France).  
Supervisor: Prof. Marc Haenlin.
- 2001–2008:      Studies of Molecular Biology, University of Vienna (Austria) with emphases in Structural Biology, Genetics and Molecular Medicine.
- 2000–2001:      Military Service: Paramedic
- 1992–2000:      High School Bundesrealgymnasium Wieselburg/Erlauf
- 1988–1992:      Primary School Volksschule Wieselburg/Erlauf

## S K I L L S

- Languages:      German (mother tongue), English (fluent), French (intermediate).  
Language course in France (2004).
- Computational skills: Windows, basic skills in Matlab and Python programming.

## CONFERENCE CONTRIBUTIONS

- International School of Biological Magnetic Resonance: 9<sup>th</sup> Course: Biophysics and Structure 2009 (Erice, Sicily, Italy):  
Presentation: **“NMR Characterization of the Ligand Binding Properties of the Model Siderocalin Q83”**
- ICMRBS 2010 (Cairns, Australia):  
Poster: **“Application of the Protein Meta-Structure to Weak Binder Screening by NMR for Fragment Based Drug Discovery”**
- PhD-PostDoc-Retreat 2010 (Krems, Austria):  
Presentation: **“Application of the Protein Meta-Structure to Weak Binder Screening by NMR for Fragment Based Drug Discovery”**
- MMCE 2011 (Tatranská Lomnica, Slovakia):  
Poster: **“NMR Probing of Transient Structure in the ‘Intrinsically Unstructured’ Protein BASP1”**
- Careers in Science 2012 (Wien, Austria):  
Poster: **“Structural Dynamics of the Intrinsically Disordered Protein BASP1”**
- MMCE 2013 (Semmering, Austria):  
Poster: **“Structural Dynamics of the Intrinsically Disordered Protein BASP1”**

## PUBLICATIONS

- **Geist L**, Henen MA, Haiderer S, Schwarz T, Kurzbach D, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Koźmiński W, Hinderberger D, Konrat R (2013)  
*Protein Sci.* 2013 Jul 2. doi: 10.1002/pro.2304. PubMed PMID: 23821606.  
Protonation-dependent Conformational Variability of Intrinsically Disordered Proteins.
- **Geist L**, Zawadzka-Kazimierczuk A, Saxena S, Zerko S, Koźmiński W, Konrat R (2012)  
*Biomol NMR Assign.*  
(1)H, (13)C and (15)N resonance assignments of human BASP1.
- Stanek J, Saxena S, **Geist L**, Konrat R, Koźmiński W (2013)  
*Angew Chem Int Ed Engl*, 52(17):4604–6  
Probing local backbone geometries in intrinsically disordered proteins by cross-correlated NMR relaxation.
- Solyom Z, Schwarten M, **Geist L**, Konrat R, Willbold D, Brutscher B (2013)  
*J Biomol NMR*, 55(4):311–21  
BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins.

- Orbán-Németh Z, Henen MA, **Geist L**, Zerko S, Saxena S, Stanek J, Koźmiński W, Propst F, Konrat R (2013)  
*Biomol NMR Assign.*  
Backbone and partial side chain assignment of the microtubule binding domain of the MAP1B light chain.
- Henen MA, Coudevylle N, **Geist L**, Konrat R (2012)  
*J Med Chem*, 55(17):7909-19  
Toward rational fragment-based lead design without 3D structures.
- Coudevylle N, Hoetzing M, **Geist L**, Kontaxis G, Hartl M, Bister K, Konrat R (2011)  
*Biochemistry*, 50(43):9192-9  
Lipocalin Q83 reveals a dual ligand binding mode with potential implications for the functions of siderocalins.
- Coudevylle N, **Geist L**, Hoetzing M, Tollinger M, Konrat R (2011)  
*J Biomol NMR*, 51(1-2):83-8  
Siderocalin Q83 exhibits differential slow dynamics upon ligand binding.
- Coudevylle N, **Geist L**, Hötzinger M, Hartl M, Kontaxis G, Bister K, Konrat R (2010)  
*J Biol Chem*, 285(53):41646-52  
The v-myc-induced Q83 lipocalin is a siderocalin.
- Auer R, Kloiber K, Vavrinska A, **Geist L**, Coudevylle N, Konrat R (2010)  
*J Am Chem Soc*, 132(5):1480-1  
Pharmacophore mapping via cross-relaxation during adiabatic fast passage.

## REFERENCES

- Prof. Dr. Robert Konrat: Department of Structural & Computational Biology, Vienna Biocenter 5, 1030, Vienna, Austria. [robert.konrat@univie.ac.at](mailto:robert.konrat@univie.ac.at)
- Prof. Dr. Georg Kontaxis: Department of Structural & Computational Biology, Vienna Biocenter 5, 1030, Vienna, Austria. [georg.kontaxis@univie.ac.at](mailto:georg.kontaxis@univie.ac.at)