

---

# **STRUCTURAL STUDIES OF PROTEINS IN INTRACELLULAR ENVIRONMENT: DESIGN AND CHARACTERIZATION OF SUITABLE CELLULAR SYSTEMS**

---

Giovanni Smaldone

Dottorato in Scienze Chimiche – XXV ciclo  
Università di Napoli Federico II







---

# **STRUCTURAL STUDIES OF PROTEINS IN INTRACELLULAR ENVIRONMENT: DESIGN AND CHARACTERIZATION OF SUITABLE CELLULAR SYSTEMS**

---

Dottorando: Smaldone Giovanni

Tutore: Prof. Giancarlo Morelli

Co-Tutore: Dr.ssa Sonia Di Gaetano

Relatore: Prof.ssa Gabriella D'Auria

Coordinatore: Prof. Lucio Previtiera

---

# Index

<b>1. INTRODUCTION</b>	<b>1-20</b>
<b>2. EXPERIMENTAL PROCEDURES</b>	<b>21-28</b>
2.1 <i>Materials</i>	
2.2 <i>Cell culture</i>	
2.3 <i>Cloning strategy of gH625-c-prune and TAT-c-prune</i>	
2.4 <i>Production of recombinant gH625-c-prune and TAT-c-prune</i>	
2.5 <i>Cloning, expression and purification of c-prune and its mutants</i>	
2.6 <i>Circular dichroism measurements</i>	
2.7 <i>Western blot analysis</i>	
2.8 <i>Preparation of Small Unilamellar Vesicles</i>	
2.9 <i>Tryptophan fluorescence experiments</i>	
2.10 <i>Labelling of gH625-c-prune and TAT-c-prune</i>	
2.11 <i>Flow cytometry of cell association of gH625-c-prune</i>	
2.12 <i>Cellular uptake kinetics and endocytosis inhibition of gH625-c-prune and TAT-c-prune by confocal microscopy</i>	
2.13 <i>Transfections</i>	
2.14 <i>NMR sample preparation</i>	
2.15 <i>NMR measurements and analysis</i>	
<b>3. RESULTS</b>	<b>29-49</b>
3.1 <i>Cloning and expression of gH625-c-prune and TAT-c-prune</i>	
3.2 <i>Purification of gH625-c-prune and TAT-c-prune</i>	
3.3 <i>Circular dichroism measurements</i>	
3.4 <i>Analysis of gH625-c-prune biological function</i>	
3.5 <i>Fluorescence measurements</i>	
3.6 <i>Determination of NBD-gH625-c-prune cellular uptake</i>	
3.7 <i>Uptake kinetics and endocytosis inhibition</i>	
3.8 <i>h-prune/Nm23-H1 interaction in unlabeled lysates.</i>	
3.9 <i>h-prune interactions with GSK3beta and gelsolin in unlabeled lysates</i>	
<b>4. DISCUSSION</b>	<b>50-53</b>
<b>5. REFERENCES</b>	<b>54-62</b>

# Introduction

## INTRODUCTION

Investigating proteins ‘at work’ in a living environment at atomic resolution is a major goal of molecular biology, which has not been achieved even though methods for the three-dimensional structure determination of purified proteins in single crystals or in solution are widely used. Various intracellular events such as conformational changes, dynamics and binding events have been investigated by this method. The detailed analysis of the network of interactions between proteins within a cell is a key step to understand the mechanisms that are located at the base of numerous pathologies.

Through the use of high-resolution spectroscopic techniques such as NMR and X-ray crystallography has been possible to obtain important information about protein structures and their function. The problem is that such techniques allow us to study a protein in non-physiological conditions, outside of the intracellular environment in which a protein is naturally found.

Until a short time ago the only way to be able to study an intracellular event consisted of marking the molecules of interest (for example with fluorescent probes such as GFP), and subsequently following the intracellular fate through the use of microscopic techniques such as fluorescence microscopy.

Such techniques have a low resolution and allow to obtain information only on a general intracellular event, preventing from obtaining the molecular details of what happens within a cell.

Using advanced methods of molecular biology, cellular biology and NMR, it was possible to obtain multidimensional heteronuclear NMR spectra of macromolecules *in vivo* at high resolution, both in prokaryotes and in eukaryotes<sup>[1]</sup>. In this way it was possible to analyze

in detail a number of intracellular events such as conformational changes and dynamic binding events between biological molecules<sup>[2]</sup>.

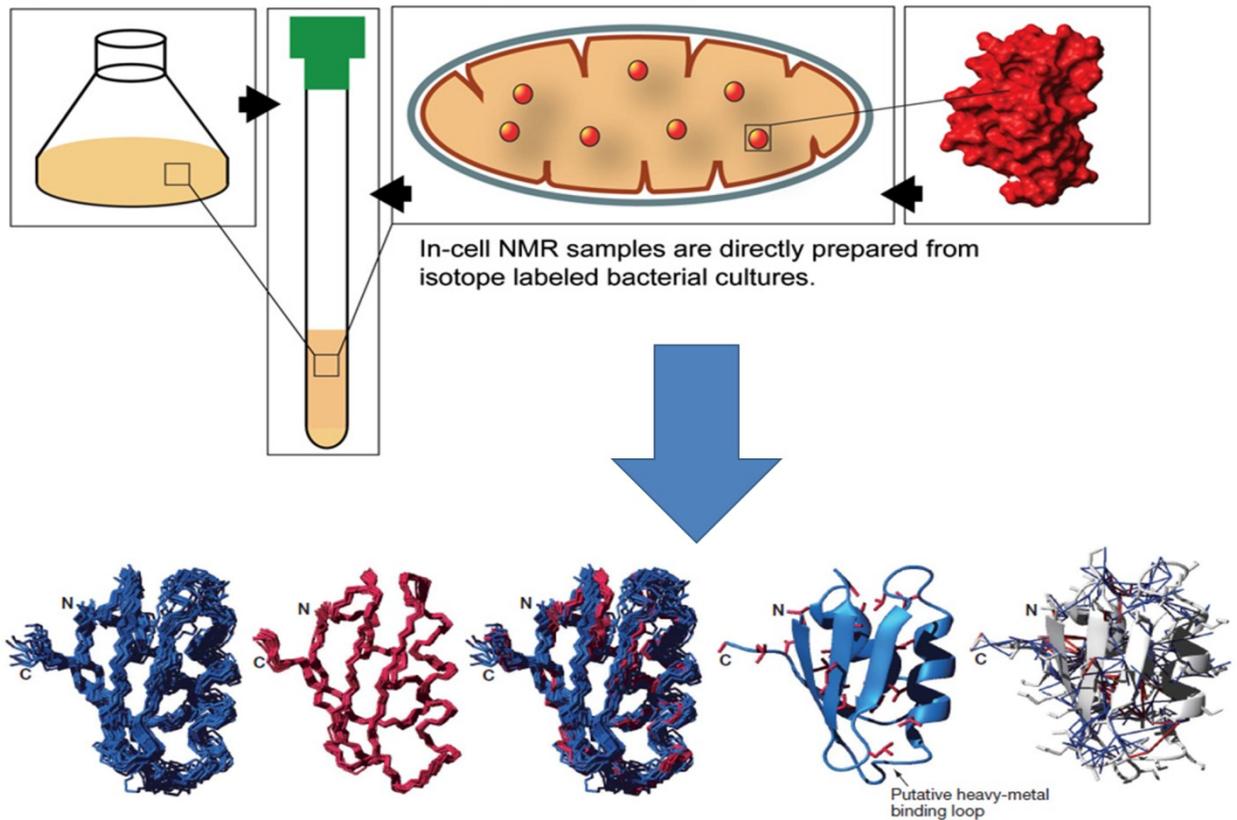
“In-cell NMR” experiments do not aim at determining structures directly in the cellular environment, but use the sensitivity of the chemical shift towards changes in the environment to obtain information about the state of a macromolecule in its natural surroundings. Changes in this environment, caused by post-translational modifications, conformational changes, or binding events, result in changes in the resonance frequencies of the affected nuclei and can thus be detected by “in-cell NMR” experiments. The investigation and characterization of macromolecules in living cells by NMR spectroscopy has to overcome three main difficulties<sup>[3]</sup>. First, the NMR signals of the molecule of interest must be distinguished from the NMR resonances of all other cellular components. Second, the macromolecule must be able to tumble freely, and third the cells have to survive the conditions inside the NMR tube at least for the time period of the experiment without significant changes of their metabolic state. Distinguishing the NMR resonances of a macromolecule from all other resonances of the cell can be achieved by overexpressing the macromolecule and labeling with the NMR-active isotopes <sup>15</sup>N and <sup>13</sup>C.

A major limitation for the application of high-resolution liquid state NMR spectroscopy of biological macromolecules is the requirement that these molecules have to tumble in solution with a sufficiently short correlation time. Long rotational correlation times lead to fast relaxation and, therefore, broad peaks. Since the rotational correlation time is proportional to the surrounding viscosity, the intracellular viscosity is an important parameter for the observation of macromolecules inside living cells<sup>[3]</sup>. Diffusion measurements have indeed shown that the translational diffusion of a macromolecule inside cells can be severely restricted relative to an *in vitro* system with the purified molecule. The intracellular rotational correlation time is only twice as long as the rotational

correlation time of the same molecule in pure water. This increase by a factor of two of the rotational correlation time also increases the apparent molecular weight of the protein by a factor of two. Fortunately, the recent introduction of TROSY<sup>[4]</sup> and similar techniques<sup>[5][6]</sup> has extended the applicability of NMR spectroscopy to large macromolecules with a molecular weight of 100 kDa and more. These technical advances, combined with the relatively low viscosity of the cellular cytoplasm, predict that the cytoplasmic viscosity is not a major limitation for the observation of proteins inside living cells.

The third critical parameter that strongly influences the applicability of in-cell NMR experiments is the survival rate of the cells in the NMR tube. In particular, the high cellular density can cause problems through oxygen starvation and limiting the amount of available nutrients. If the sensitivity of the selected system (mainly the overexpression level) is high enough, the NMR spectra can be measured relatively quickly (less than an hour)<sup>[3]</sup>.

Thanks to the application of more and more faster experiments (FAST-NMR)<sup>[7]</sup> and use of high amounts of labeled proteins of interest, these problems have been exceeded. In fact by “in cell NMR” 3D protein structure was calculated exclusively on the basis of information obtained in living cells<sup>[8]</sup> (Fig. 1).

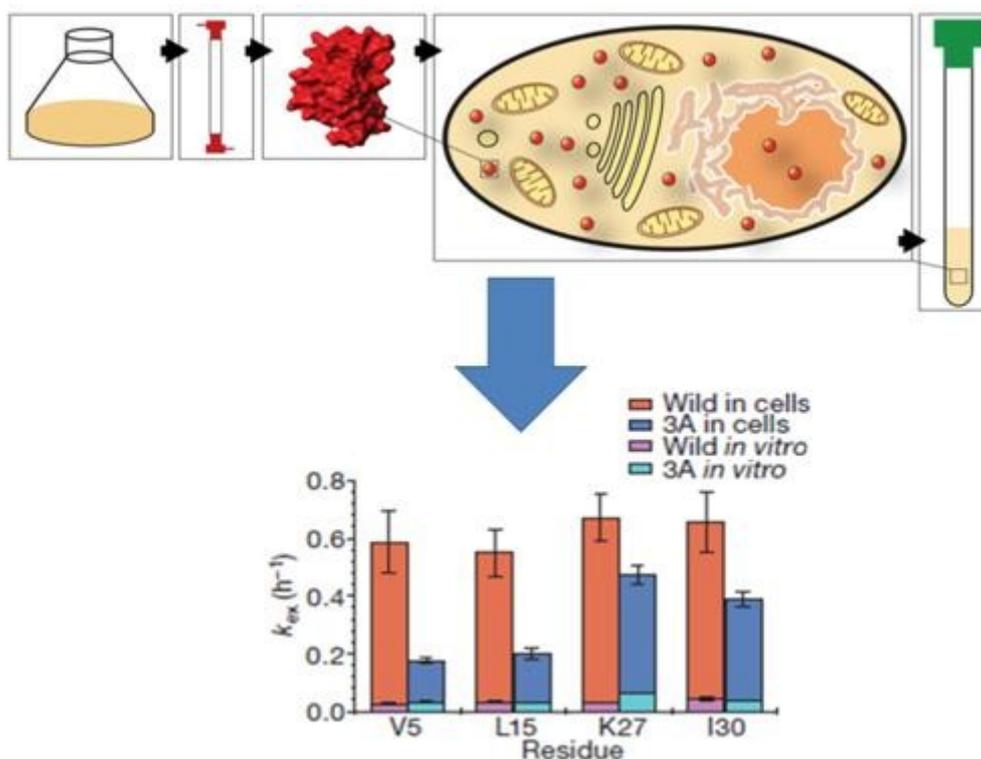


**Fig. 1NMR solution structure of TTHA1718 in living *E. coli* cells.** Prokaryotic in-cell NMR approaches typically employ recombinant protein over-expression, isotope labeling and in-cell NMR analyses within the same cell type. Suspensions of bacterial cells are directly analyzed without purification of the recombinant protein. A superposition of the 20 final structures of TTHA1718 in living *E. coli* cells, showing the backbone (N, C $\alpha$ , C') atoms <sup>[8]</sup>.

The structure of the putative heavy-metal binding protein TTHA1718 from *Thermus thermophilus* HB8 over-expressed in *Escherichia coli* cells was solved by in-cell NMR. Rapid measurement of the 3D NMR spectra by nonlinear sampling of the indirectly acquired dimensions was used to overcome problems caused by the instability and low sensitivity of living *E. coli* samples. Almost all of the expected backbone NMR resonances and most of the side-chain NMR resonances were observed and assigned with high quality. Structures calculated are very similar to the *in vitro* structure of TTHA1718 determined independently. The in-cell NMR approach can thus provide accurate high-resolution structures of proteins in living environments<sup>[8]</sup>.

To demonstrate the broad application of this technique, the in-cell NMR can be used to

studying interactions and protein processing. For example, in-cell NMR spectroscopy demonstrates that both ubiquitin and its mutant 3A have much higher hydrogen exchange rates in the intracellular environment, possibly due to multiple interactions with endogenous proteins<sup>[9]</sup> (Fig. 2). In addition, the same technique can also be applied to studying protein interactions with small compounds, such as screening drugs targeted to specific proteins.

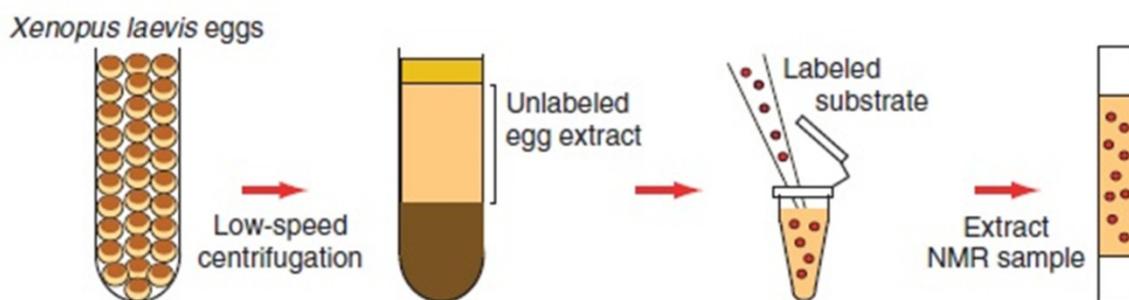


**Fig. 2 The hydrogen exchange rates in eukaryotic cells.** Eukaryotic in-cell NMR applications can involve isotope labeling in bacterial cells and recombinant protein purification prior to in-cell NMR sample preparation. Labeled proteins are then transferred into eukaryotic cells by microinjection, or other vector-based transduction techniques. The hydrogen exchange rates ( $k_{ex}$ ) in cells and in vitro at pH 7.4 of protected amides in wild type ubiquitin and in the mutant 3A<sup>[9]</sup>.

Indeed, the in-cell NMR spectra of FKBP12 (also known as FKBP1A) show the formation of specific complexes between the protein and extracellularly administered immunosuppressants, demonstrating the utility of this technique in drug screening programs<sup>[9]</sup>.

This technique can be also used to study post-translational protein modifications (PTMs) that occur in living environments such as phosphorylation and acetylation<sup>[10]</sup> or methylation<sup>[11]</sup> that regulate a large number of eukaryotic signaling processes.

At the same time comparable results between analyses performed on whole live cells and extracts have been obtained (Fig. 3).



**Fig.3** Overview of the preparation of extract NMR samples.

In a recent paper of Selenko et al.<sup>[12]</sup> it was demonstrated the experimental feasibility of high resolution NMR measurements in the complex environments of crude *Xenopus* egg extracts. These extracts are especially suited to predict the cellular behavior of a protein in a conveniently accessible, homogenous solution that closely resembles the cytoplasmic properties of intact cells. Extract NMR measurements may be used to indicate the likelihood for satisfactory in-cell NMR experiments and/or may serve to independently analyze structural and functional aspects of biological activities of a protein in a cell-free setting. Because extracts are easily prepared in large quantities and because most biological reactions are executed with comparable activities in these *ex vivo* mixtures, we wish to strongly emphasize their supplementary potential for “cellular” NMR analyses. Moreover, these extracts can be selectively depleted of individual cellular components or supplemented with recombinant proteins, small molecule inhibitors, and other components, allowing NMR analyses under defined cellular conditions<sup>[12]</sup>.

Over the years various approaches have been developed for the transfer of macromolecules in animal and plant cells; among these methods, cationic liposomes are able to transfer molecules penetrating into the cell by fusion with the plasma membrane.

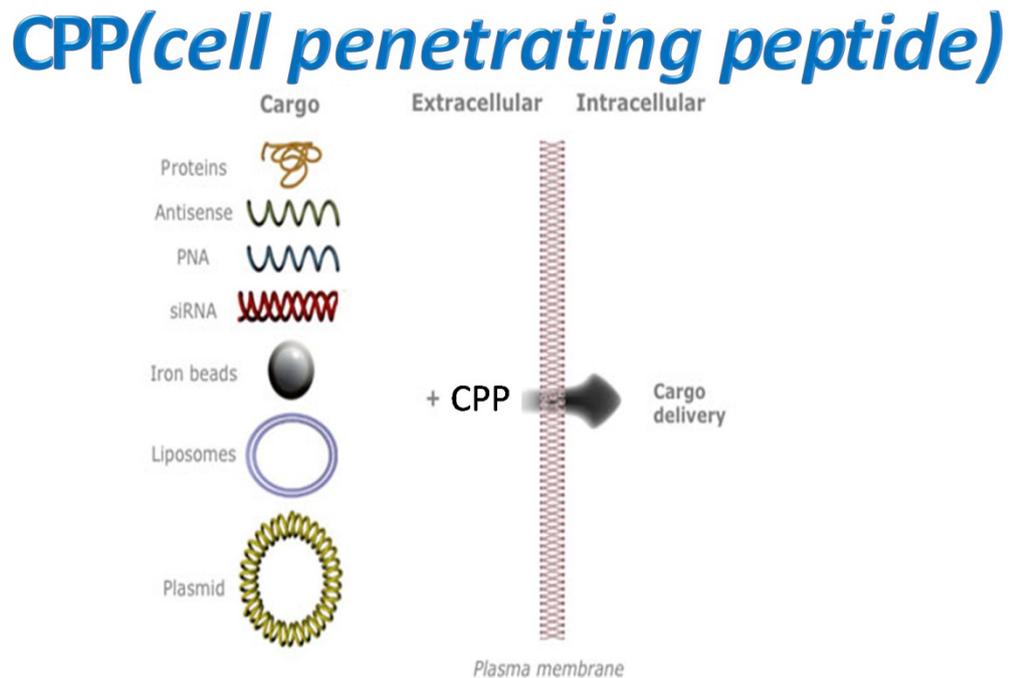
Despite numerous systems have been developed to transfer the macromolecules into living cells, they have a number of limitations that preclude its routine use in *in vivo* studies, in particular for possible therapeutic applications in the clinical field. The main obstacles are summarized in the low efficiency of internalization, in the complexity of handling, in the application of expensive equipment, in the difficulty of release into the cytosol and sometimes in cellular toxicity and immunogenicity.

To facilitate the internalization of proteins of interest into the cells it has been developed a technology that involves the conjugation of the protein with small peptides (10-15 amino acid residues) characterized by the presence of numerous basic amino acids, such as arginine. Using molecular biology techniques it is possible to fuse such peptides to the nucleotide sequence of a specific protein to facilitate the entry into the cell.

Furthermore, there are numerous agents, such as pyrenebutyrate, capable of favoring the passage of the cellular membrane and the entry into the target cell<sup>[13]</sup>.

Cell compartmentalization and impermeability of membranes represent the major obstacle for the delivery of therapeutic molecules; in fact, many bioactive molecules are incapable of overcoming the membrane permeability barrier and this represents the major impediment for gene and drug targeting in theranostics. Although a wide variety of bio-drugs, including peptides and proteins, are now produced on a commercial scale, one challenging task for pharmaceutical researchers is devise ways to deliver these drugs effectively and safely via non-invasive, patient-friendly routes<sup>[14][15]</sup>.

Cell-penetrating peptides (CPPs) are a valuable class of natural short peptide sequences which have been recently considered very promising for cellular uptake and membrane interaction studies<sup>[16]</sup> (Fig. 4).



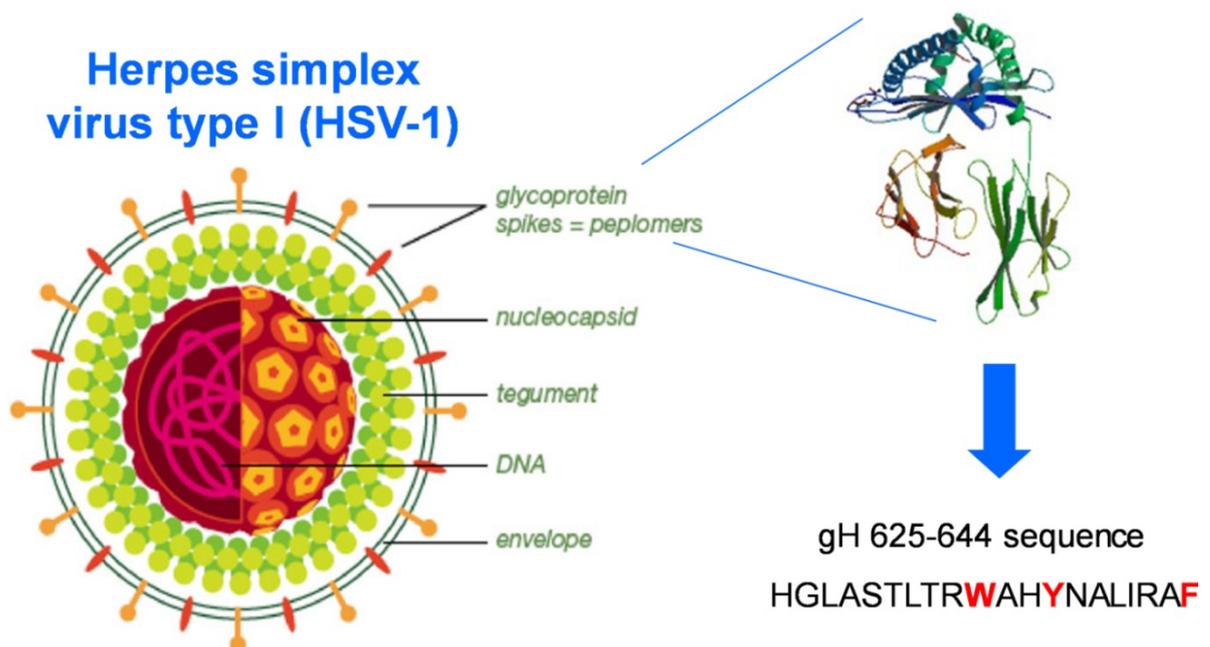
**Fig. 4** Different molecules can be transported into the cell by the CPP.

While individual CPPs differ in length and sequence, they share some common features, which include their amphipathic nature, net positive charge, hydrophobicity and helical moment, the ability to interact with lipidic membranes and to adopt a distinct secondary structure upon association with lipids<sup>[17]</sup>. These short cationic peptides either by covalent or noncovalent binding have been exploited for the delivery of a range of macromolecules in the biologically active form such as proteins, plasmids, siRNA, nanoparticles, PNAs, and liposomes, that are otherwise impermeable to the plasma membrane. Unfortunately, CPPs application in cellular delivery has been hampered by the controversy regarding their uptake mechanism; in fact, no defined picture of the mechanism has yet emerged<sup>[16-20]</sup>.

Among the CPPs, the human immunodeficiency virus type 1 (HIV-1) TAT basic domain has been used extensively for macromolecule delivery<sup>[20-22]</sup>. TAT is a nuclear transcriptional activator protein consisting of 101 amino acids. This protein is required for viral replication by HIV-1. The truncated sequence of HIV-1 TAT 49–57 (RKKRRQRRR) is a highly basic region typically involved in cellular translocation<sup>[20]</sup> and proteins as small as GFPs (30 kDa) and as large as IgG (150 kDa) have been delivered efficiently by TAT peptide<sup>[23]</sup>. Problems in the application of CPPs arise from the unfavorable physicochemical and biological properties of therapeutic peptides and proteins to deliver, which may affect their absorption. In fact, the use of a charged delivery peptide may represent a disadvantage in some cases. Therefore, it is fundamental to exploit novel molecules which present different structural characteristics and use different internalization mechanisms<sup>[24]</sup>. For this reason, great attention has been recently devoted to the study of hydrophobic peptides that efficiently traverse biological membranes, promoting lipid membrane-reorganizing processes, such as fusion or pore formation and thus involving temporary membrane destabilization and subsequent reorganization<sup>[25]</sup>. Viral derived peptides, in particular those derived from viral entry proteins, may be useful as delivery vehicles due to their intrinsic properties of inducing membrane perturbation<sup>[26-28]</sup>. Delivery across cellular membranes involves several mechanisms such as direct transfer through cell surface membrane by lipid membrane fusion or transient permeabilization of the cell membrane or, after endocytosis, transfer across vesicular membranes by lipid disruption, pore formation or fusion.

The nineteen residues peptide gH625 was previously identified as a membrane-perturbing domain in the gH protein of *Herpes simplex* virus type I (Fig. 5); gH625 interacts with model membranes, contributing to their merging; it is able to traverse the membrane bilayer and to transport a cargo into the cytoplasm and across the blood brain barrier<sup>[25-29]</sup>.

In particular, it's reported the ability of gH625 peptide to transport quantum dots inside the cytoplasm in an efficient way and only partially involving endocytic pathways<sup>[30]</sup>, moreover it's reported its ability to enhance intracellular penetration of liposomes functionalized with the peptide on the external leaflet<sup>[31]</sup>, nanoparticles<sup>[15]</sup> and dendrimers<sup>[14]</sup>. Compared to the TAT peptide which mainly exploits the endocytic pathway, the viral membranotropic peptide gH625 crosses membrane bilayers mainly through a translocation mechanism.

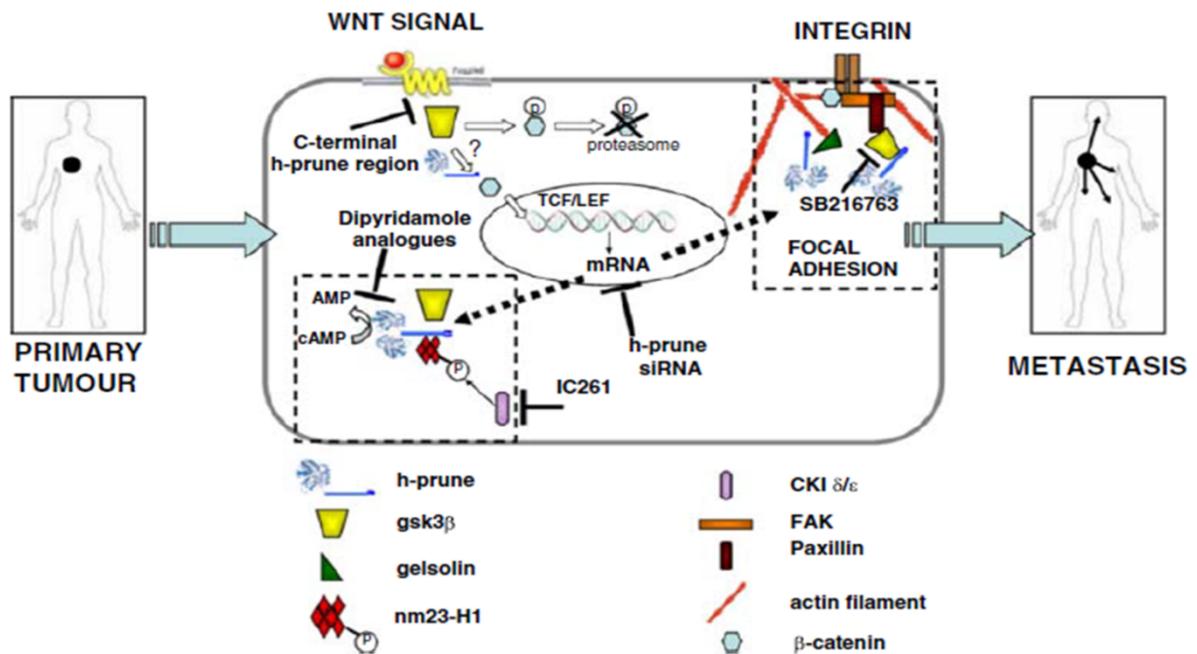


**Fig.5** Representation of the region of Herpes Simplex virus type 1 in which the peptide gh625-644 has been identified. In red the aromatic residues of peptide.

The hydrophobicity of gH625 compared to the positively charged TAT can represent an advantage for the delivery of an acidic and natively unfolded protein such as the C-terminal domain of the protein h-prune.

H-prune is the best characterized interactor of Nm23-H1. Its mRNA and protein expression is high in advanced neuroblastoma (NBL), combined with high protein levels in serum<sup>[32]</sup>, which indicates that it has a function as a metastasis-promoter gene in NBL.

Indeed, amplification and over-expression of Nm23-H1, and also of the S120G mutation of Nm23-H1 (Nm23-H1-S120G) have been detected in 14% to 30% of patients with advanced NBL stages. The Nm23-H1 mechanism in the mediation of NBL aggressiveness remains to be understood and it is still subject to debate. In this mechanism the complex Nm23-H1/h-prune is involved, so it is very important to understand the molecular basis of this interaction. The h-prune protein is a member of the phosphoesterases (DHH) protein superfamily, and its overexpression in breast, colorectal and gastric cancers correlates with the level of invasion and the degree of lymph-node and distant metastases<sup>[33-36]</sup> (Fig.6).

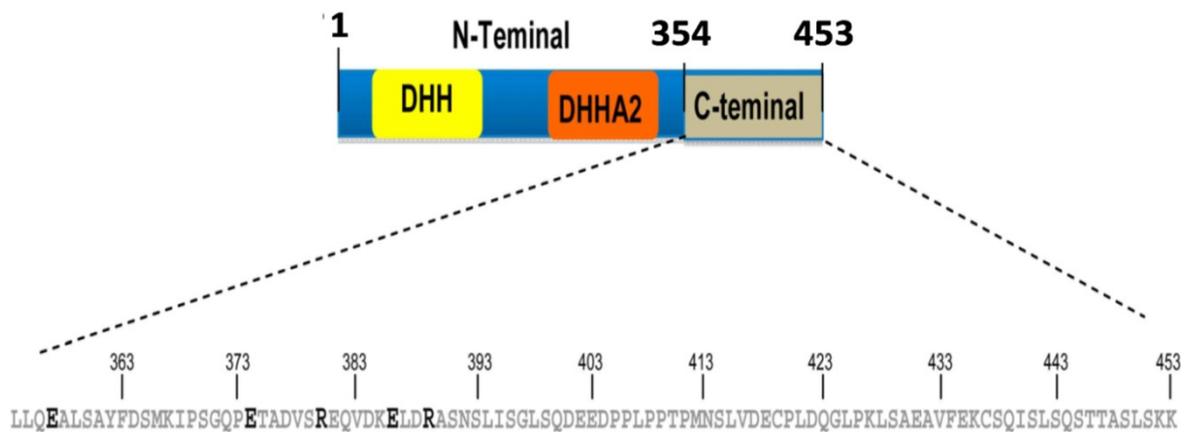


**Fig.6** Model of the activation of h-prune and its protein-interactor functions that result in a high motility phenotype, thus leading to substantial changes in cell status: from a primary to a metastatic cancer cell environment.

H-prune also has a phosphodiesterase (cAMP-PDE) activity, and inhibition of this PDE activity with dipyridamole suppresses cell motility in breast cancer cell lines<sup>[37-39]</sup>. The N-terminus of the h-prune sequence contains the DHH (amino acids 10-180) and DHHA2 (amino acids 215-360) domains that are involved in its enzymatic function. Of note, there

is an additional function of h-prune within this N-terminus that includes the DHH domains: an exopolyphosphatase (PPASE) activity that was described by Lahti and colleagues through its homology similarities with the yeast protein<sup>[40]</sup>. However, this additional activity does not affect either the cell motility of breast cancer cell lines or their metastasis formation *in vivo*<sup>[41]</sup>.

These domains are then followed by a cortexillin-homology region (CDH, amino acids 360-453) that contains putative coiled-coil and proline-rich regions. The C-terminal region of h-prune (amino acids 353-454, c-prune) (Fig.7) is responsible for its interaction with GSK-3 $\beta$ <sup>[42]</sup> and with Nm23-H1<sup>[43]</sup>.

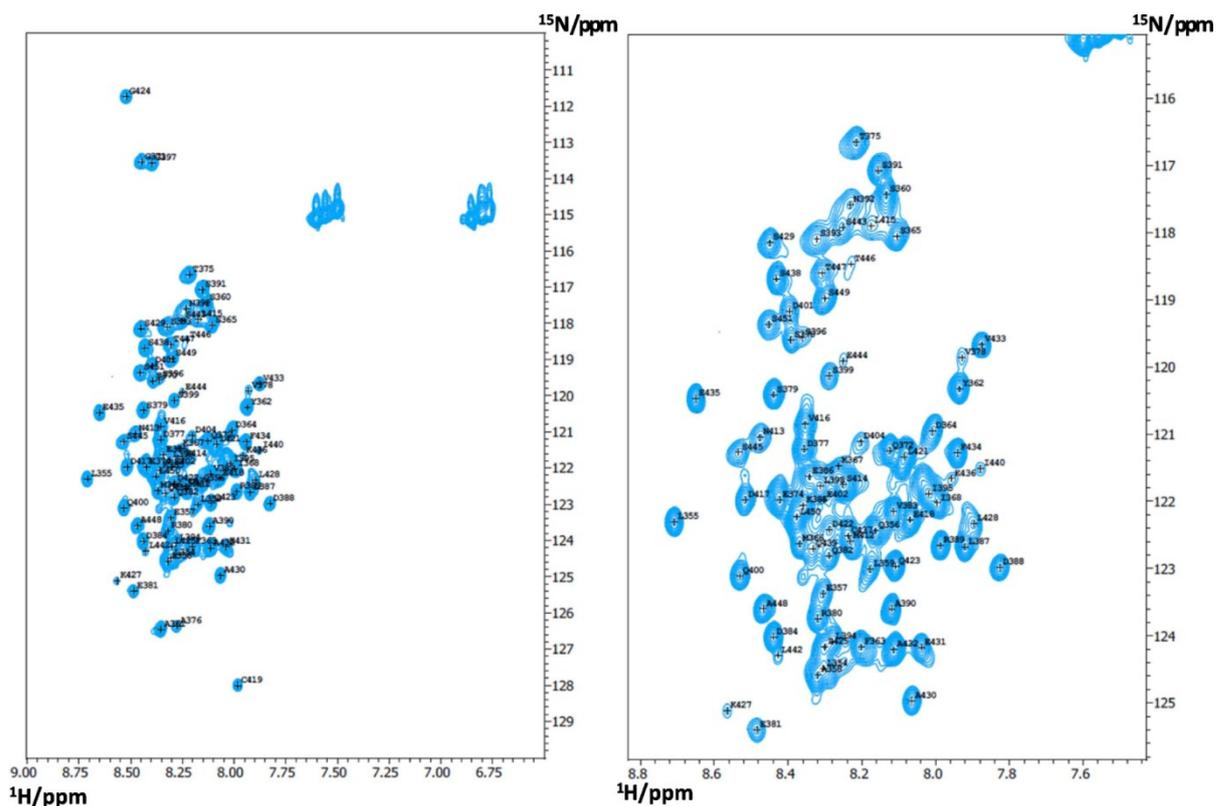


**Fig. 7** The h-prune C-terminal sequence. The amino acids given in bold represent those that are more exposed according to limited proteolysis experiments.

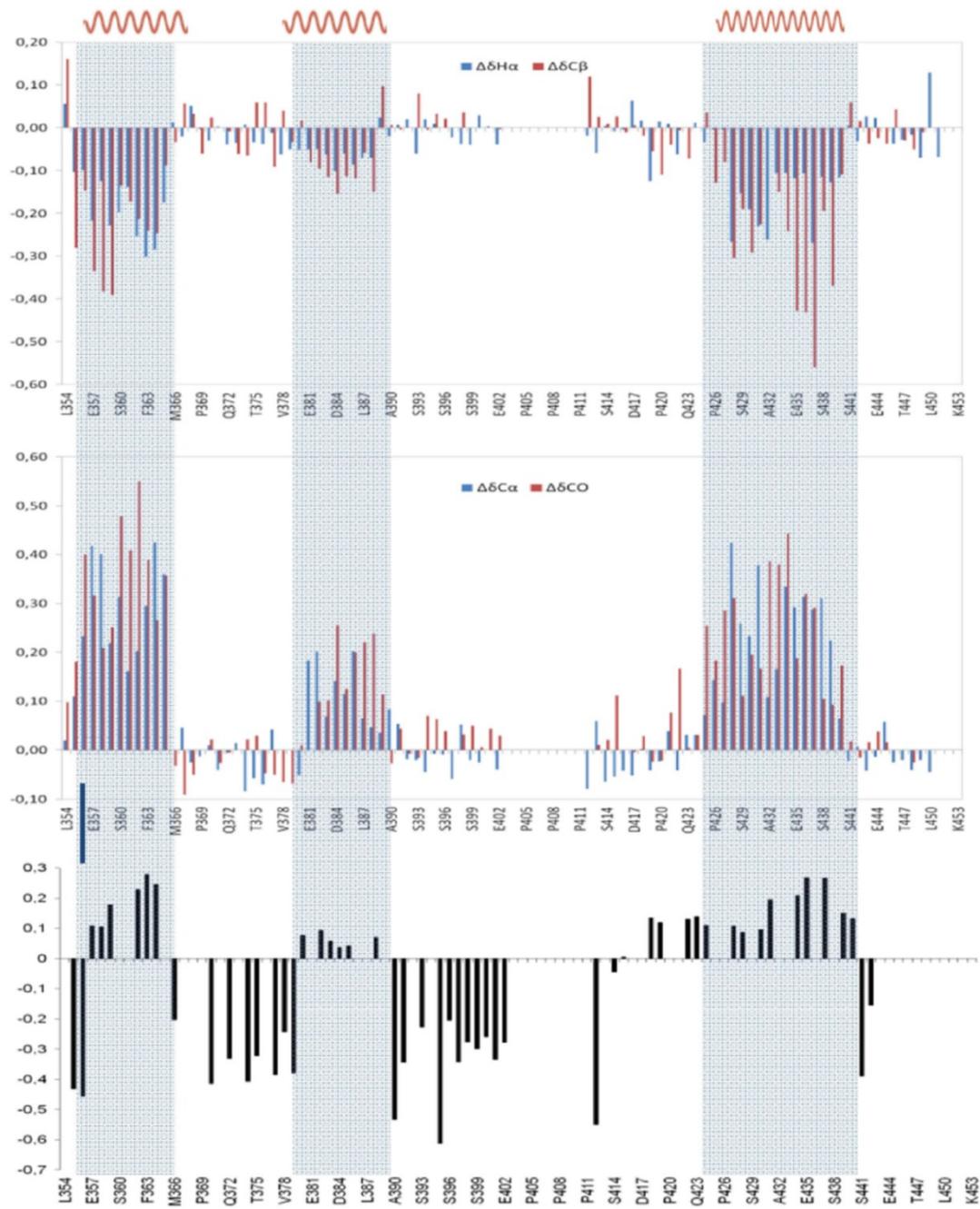
The Nm23-H1/h-prune interaction is mediated through casein kinase (CKI, CKII) phosphorylation of Ser120, Ser122 and Ser125 of Nm23-H1<sup>[44]</sup>, phosphorylation sites that have been highly conserved through evolution, and that promote enhanced cell motility<sup>[45][46]</sup>. The conformational analysis of the c-prune performed by NMR showed that it assumes a random coil conformation with the exception of some protein regions with helical structural propensity and a disulphuric bridge as unique rigid linkage defining it an intrinsically disordered domain (IDD). In addition the recombinant c-prune was sufficient

to bind the characterized interacting Nm23-H1 endogenous protein as verified by western blotting analysis and NMR spectroscopy<sup>[41]</sup>. C-prune is stable and soluble. His sequence has a low overall hydrophobicity and a high net negative charge, and analyses according to several algorithms have suggested that it is mostly unfolded in the native protein. The  $^1\text{H}$  NMR spectra of c-prune have been very difficult to assign due to poor dispersion of the resonances. To increase the dispersion of the NMR signals and to allow analysis of the dynamics using multinuclear NMR spectroscopy, uniformly  $^{15}\text{N}$ -labeled and  $^{15}\text{N}/^{13}\text{C}$ -labeled c-prune samples were prepared. Sequential assignments of the  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}$ ,  $^{13}\text{CO}$ ,  $^{13}\text{C}_\alpha$ ,  $^1\text{H}_\alpha$  and  $^{13}\text{C}_\beta$  resonances were obtained in full following a standard triple-resonance procedure. A  $^1\text{H}$ - $^{15}\text{N}$  hetero-nuclear single quantum coherence (HSQC) spectrum of c-prune is shown in Fig.8. The chemical shift assignments allowed an analysis of the secondary structure of the protein by comparison with random coil values corrected for local sequence effects. As shown in Fig.9, multinuclear ( $\text{H}_\alpha$ ,  $\text{C}_\beta$ ,  $\text{C}_\alpha$ ,  $\text{C}'$ ) chemical shift indexing (CSI) plots suggest that the majority of the resonances of c-prune are within the random coil range. Nevertheless, the CSI values of the amino-acid sequences from Leu355 to Ser365 ( $\alpha$ 1 helical region), from Glu381 to Asp388 ( $\alpha$ 2 helical region), and from Leu428 to Gln439 ( $\alpha$ 3 helical region) are consistent with an  $\alpha$ -helical secondary structure propensity. The CSI data are also supported by the  $^3\text{J}_{\text{HNHA}}$  coupling-constant measurements, which are characteristic of unfolded molecules, with the exception of the three helical regions, which have means of 5.3 Hz, as typical of helical structures<sup>[41]</sup>. A three-dimensional model of the full-length h-prune protein was then built using an N-terminal (amino acids 1-352) h-prune structure derived by homology modelling and attached to a representative NMR c-prune structure. The full-length h-prune model was then refined through molecular dynamics simulation *in vacuo*. The analysis of the whole

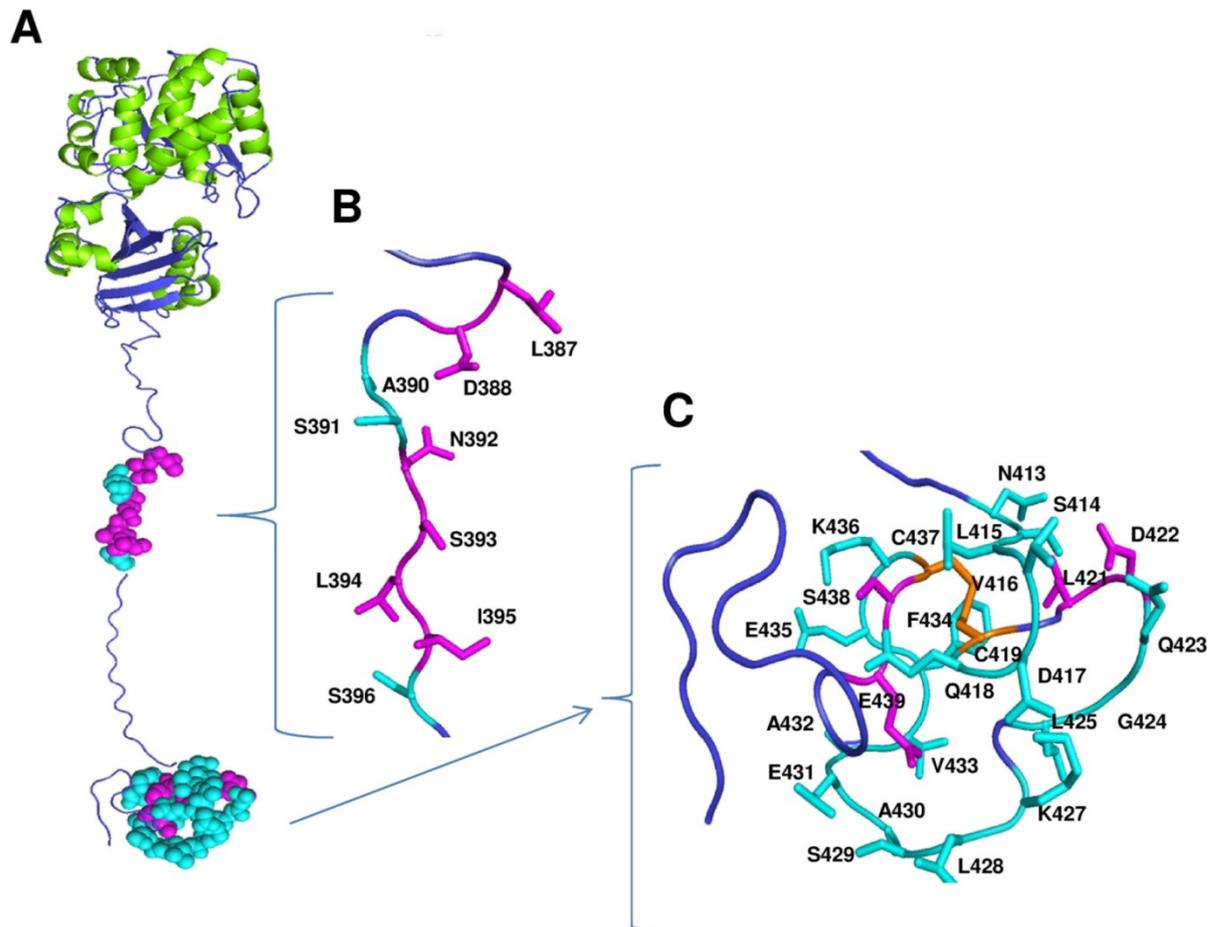
structure indicates that the N terminus (amino acids 353-370) of c-prune (Fig.10A) is indeed part of the h-prune DHH2 domain, and in particular constitutes the second part of the last helix and a turned region that interacts with the preceding helix; accordingly, this region in c-prune has a clear helical propensity (Fig. 9). Therefore, the IDD c-prune that does not have specific interactions with the globular portions of the whole protein begins at residue 371 and retains the secondary structure propensities ( $\alpha 2$  and  $\alpha 3$ ) indicated by the NMR analysis, with a more compact C-terminal region (amino acids 410-440)<sup>[41]</sup>. To better understand the Nm23-H1/c-prune interaction, a series of two-dimensional [ $^1\text{H}$ ,  $^{15}\text{N}$ ]-HSQC experiments were performed.



**Fig.8** Two dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of the h-prune C-terminal. (Left) Backbone amide proton resonances are narrowly dispersed over the range of 7.8 ppm to 8.8 ppm, which is characteristic of unstructured proteins. (Right) Expanded region to allow labeling of the crowded central region of the spectrum.



**Fig.9** The secondary chemical shift index, corrected for sequence-dependent contributions, based on  $^1\text{H}_\alpha$ ,  $^{13}\text{C}_\beta$ ,  $^{13}\text{C}_\alpha$  and  $^{13}\text{CO}$  chemical shifts of the h-prune C-terminal. Protein regions with a propensity to a helical structure (a1, a2 and a3) are highlighted. The  $^1\text{H}$ - $^{15}\text{N}$  steady-state heteronuclear NOE values are also reported, according to the h-prune C-terminal amino acids.



**Fig. 10. Three-dimensional model of full-length h-prune.** *A.* The amino acids showing large variations upon complex formation with Nm23-H1 and the CPP peptide are mapped in cyan. The regions colored in magenta correspond to the amino acids recognized exclusively by Nm23-H1, as shown in detail in panels *B* and *C*.

Progressive additions of Nm23-H1 induced line broadening and a strong reduction in the peak intensities of some of the amino acids located in the 387-395 region and in the more compact region constituted by the disulfide-bridged cycle and the  $\alpha$ 3 helical region (amino acids 410-440). This should therefore represent the Nm23-H1 binding epitope (Fig. 10 A,B,C)<sup>[41]</sup>.

These studies have been performed exclusively *in vitro* and allowed to identify the binding site of Nm23-H1 on c-prune surface. In addition further recent studies using a mimetic peptide, derived from the minimal interaction region of Nm23-H1 with h-prune, showed the impairment of NBL *in vivo* and elected it as possible therapeutic strategy<sup>[41]</sup>.

Gene expression studies and protein–protein pull-down analyses coupled to mass spectrometry have led to the identification, to date, of three h-prune interacting partners that themselves have important features: glycogen synthase kinase-beta (GSK-3 $\beta$ ), gelsolin<sup>[43]</sup> and ASAP1<sup>[47]</sup>. GSK-3 $\beta$  is a serine–threonine kinase<sup>[48]</sup> that was initially shown to phosphorylate and inactivate glycogen synthase<sup>[49]</sup>. GSK-3 $\beta$  is known to be involved in energy metabolism, neuronal cell development and body pattern formation<sup>[48]</sup>. Together with the strong co-localization of h-prune with the focal adhesions proteins F-actin, paxillin and vinculin in several cell types (including breast cancer cell lines), this h-prune–GSK-3 $\beta$  interaction has led to the hypothesis of a specific association of h-prune with microtubule architecture and dynamics<sup>[42-44]</sup>. This would, in turn, results in h-prune modulation of the cell-shape processes involved in cell development, proliferation, migration and differentiation; all of which are activities that are linked to cancer<sup>[50]</sup>. Gelsolin is a member of a family of proteins that can modulate the structural dynamics of the actin cytoskeleton, through the sequestration and capping of monomeric G-actin, and then severing and capping of the barbed ends of polymerized F-actin filaments<sup>[51]</sup>. Thus

gelsolin is involved in the different physiological processes that enable cells to control their shape and motility. The formation of the h-prune-gelsolin complex might also include the Nm23-H1 protein<sup>[50]</sup>.

ASAP1 is a multi-domain adaptor protein with ADP-ribosylation factor-GAP activity, as being potentially involved in tumor progression. ASAP1 has been implicated in regulating cell motility and invasion<sup>[52]</sup>. It is amplified and overexpressed in uveal melanomas and in colorectal, prostate and breast carcinomas<sup>[53-56]</sup>. ASAP1 has been functionally linked with breast cancer metastasis<sup>[55]</sup>. It was demonstrated that ASAP1 interacts with h-prune<sup>[47]</sup>. Furthermore, ASAP1 stimulates the phosphodiesterase activity of h-prune, an activity required for its motility-promoting properties<sup>[37]</sup>. To be able to understand the molecular details of these interactions would shed light on the complex mechanisms at the base of many cancers.

This PhD project has been focused on the study of the IDD c-prune belonging to a class of proteins/domains, whose involvement in many cellular processes is emerging. Indeed, many gene sequences in eukaryotic genomes encode entire proteins or large segments of proteins that lack a well-structured three-dimensional fold. Disordered regions can be highly conserved between species in both composition and sequence and, contrary to the traditional view that protein function equates with a stable three-dimensional structure, disordered regions are often functional, in ways that we are only beginning to discover. Many disordered segments fold on binding to their biological targets (coupled folding and binding), whereas others constitute flexible linkers that have a role in the assembly of macromolecular arrays<sup>[57]</sup>.

Many eukaryotic proteins are modular that is, they contain independently folded globular domains that are separated by flexible linker regions. Linker sequences vary greatly in length and composition, but many are rich in polar, uncharged amino acids (such as Ser,

Thr, Gln and Asn), in the small residues Ala and Gly, and in Pro residues. In the absence of their targets, modular proteins often behave as “beads on a flexible string”, where the function of the linker is, primarily, to enable a relatively unhindered spatial search by the attached domains<sup>[58]</sup>. However, binding can induce structure formation in linkers, which can have significant functional consequences. The intrinsically disordered proteins (IDPs) have a lot of functions that include the regulation of transcription and translation, cellular signal transduction, protein phosphorylation, the storage of small molecules, and the regulation of the self-assembly of large multiprotein complexes such as the bacterial flagellum and the ribosome. A recent review<sup>[58]</sup> highlights the occurrence of unfolded regions in proteins that function as chaperones for other proteins and for RNA molecules, and proposes that the unfolded regions work by binding to misfolded proteins and RNA molecules<sup>[58]</sup>.

The aim of this PhD project was to design a suitable system to study a protein in its physiological conditions. This methodology becomes more intriguing when the selected molecule is an IDP, whose folding may be induced by its interaction with the partners. Our target has been the IDD c-prune. Firstly we tried to obtain a fusion protein able to introduce c-prune in the cell to perform in cell-NMR spectra. To achieve this goal we have studied the mechanism of internalization of this domain in eukaryotic HeLa cell line using different CPPs. Successively the behavior of c-prune in cell lysates of eukaryotic cells has been evaluated.

In particular we focused on the use of gH625 for delivery of c-prune. The peculiar properties of gH625 render it an optimal candidate to act as a carrier for net acidic charged molecules by comparison with the positively charged TAT. Furthermore this study is focused also on the interaction studies between c-prune and its different partners over-expressed in human cells via NMR spectroscopy. Besides Nm23-H1, the interaction with

GSK-3 $\beta$ , ASAP1 and gelsolin has also been studied. All these proteins are involved in different physiological processes that enable cells to control their shape and motility and that are strictly linked to tumor progression.

## Experimental Procedures

## EXPERIMENTAL PROCEDURES

### *Materials*

*Escherichia coli* strains DH5 $\alpha$  (Invitrogen, California, USA) and BL21 (DE3) STAR (Invitrogen, California, USA) were used for cloning and expression of the recombinant proteins, respectively. pET28b+ (Novagen, Wisconsin, USA) and pETM11 (EMBL) were used as expression vectors. The oligonucleotides were synthesized by Primm (Milano, IT). All the other reagents used were from Delchimica (Naples, Italy). The anti-Nm23-H1 antibody and anti-his antibody were from Sigma Aldrich.

### *Cell culture*

HeLa cells (ATCC, USA) SH-SY5Y cells (ATCC, USA) and Human Embryonic Kidney cells (HEK-293) were grown in DMEM supplemented with 10% fetal bovine serum (FBS), 1% glutamine, 100 U/mL penicillin and 100  $\mu$ g/mL streptomycin (Invitrogen, Carlsbad, CA) at 37° C in humidified air with 5% CO<sub>2</sub>.

### *Cloning strategy of gH625-c-prune and TAT-c-prune*

The coding sequence of gH-625-644 (60 bp) was amplified by polymerase chain reaction using specific primers from genomic DNA of Herpes simplex virus 1 (HSV-1). To ensure correct orientation into the multiple cloning sites of the vector, the forward primer (5'-GGAATTCCATATGCACGGCCTCGCCTCG-3') incorporated a *Nde*I restriction enzyme site, whereas the reverse primer (5'-CATGCCATGGGGAAGGCGCGGATCAG-3') incorporated a *Nco*I site. The digested fragment was ligated to the gene encoding c-prune (aminoacid sequence 353-454, pETM11 c-prune) previously digested with the restriction enzymes *Nco*I and *Xho*I<sup>[41]</sup>. The fused segment was finally ligated to the expression vector pET28b+ which provided a cleavable N-terminal His<sub>6</sub> tag. The coding sequence of *TAT-c-*

*prune* was obtained through two consecutive PCR reactions. The first PCR reaction was carried out using pETM11-c-*prune* as a template and the forward primer (5'-CCGCCAGCGCCGCCGCCCTGCTCCAGGAAGC-3') and the reverse primer (oligo rev454) incorporated *XhoI* restriction enzyme site (5'-CCGCTCGAGTTACTTCTTGGACAGGGAGGCTG-3'). The product of this PCR was used as template for a second PCR reaction using the forward primer incorporated *NcoI* restriction enzyme site (5'-CATGCCATGGGCTACGGCCGCAAGAAACGCCGCCAGCGCCGCC-3') and the oligo rev454.

The coding sequence of *TAT-c-prune* was ligated to the expression vector pETM11 which provided a cleavable N-terminal His<sub>6</sub> tag. The resulting positive plasmids were used to transform *E. coli* BL21 (DE3) STAR strain. The expression of the recombinant proteins (gH625-c-*prune* and TAT-c-*prune*) was carried out using the transformed cells grown at 37°C in 0.5L of LB containing 50 µg/mL kanamycin and then inducing them overnight with 0.5 mM IPTG at 22 °C.

#### ***Production of recombinant gH625-c-prune and TAT-c-prune***

C-*prune* was expressed and purified as elsewhere described <sup>[41]</sup>. The gH625-c-*prune* and TAT-c-*prune* proteins were purified from inclusion bodies. Briefly, the cells were resuspended in 30 mL of 50 mM Tris-HCl (pH 7.5), 8 M urea, 300 mM NaCl, and 5% glycerol (buffer A) and incubated for 16h at RT and then sonicated to promote cell lysis. The soluble and insoluble fractions of the lysate were then separated by centrifugation (17000 rpm for 20 min at 4 °C). The soluble fraction of the lysate was mixed with 4 mL of Ni-NTA resin (Qiagen) equilibrated with the same buffer and then incubated for 1 h at RT. The column was washed with ten volumes of denaturing buffer (buffer A) plus 50 mM imidazole. The denatured proteins were then eluted by increasing imidazole concentration

(300 mM). The purity of the proteins were analyzed by 15% SDS–PAGE whereas their concentration was determined by the Bradford assay using BSA as protein standard. The refolding procedure was carried out by the dilution of the denatured proteins concentrated at 0.2-0.4 mg/mL in 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 5% glycerol, 0.5M L-Arginine and 5mM dithiothreitol (DTT). After an overnight mild stirring at 4 °C, the insoluble proteins were removed by centrifugation at 12000 rpm for 30 min at 4 °C. To remove the His-tags, the refolded proteins were subjected to proteolytic digestion with thrombin, for gH625-c-prune and TEV protease for TAT-c-prune, respectively. The concentration of the refolded proteins were estimated by Bradford assay. Then, the proteins were loaded onto a Superdex 75 gel filtration column (HiLoad 10/30, GE Healthcare) pre-equilibrated with 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 5% glycerol, 1.5% D-glucose and 5 mM DTT.

### ***Cloning, expression and purification of c-prune and its mutants***

The construct, pETM-11-c-prune, has been already described elsewhere<sup>[41]</sup>. The single mutants, namely L394A, D422A and D388A, were produced by following the protocol outlined in the Quick Change II site directed mutagenesis kit (Stratagene). Two overlapping complementary primers containing the desired nucleotide changes were designed for each mutation using appropriate oligonucleotides. Plasmids carrying the above mentioned mutations were sequenced at the Primm DNA sequencing service (Naples Facility). The plasmids containing the mutations inserted were used to transform BL21 (DE3) STAR competent cells. Protein expression and purification protocols for <sup>15</sup>N labeled c-prune and its mutants have been performed as previously reported with slight modifications <sup>[41]</sup>. Briefly, the transformed BL21 (DE3) STAR cells were grown at 37°C on a LB agar plate containing 100 µg/mL kanamycin for 16 h. A single colony was picked

and grown in 80 mL of a LB medium containing 100  $\mu\text{g/mL}$  kanamycin at 37° C overnight. To switch from LB medium to a minimal medium, the cells were pelleted at 3000 rpm for 10 min, then washed by using 10 mL of minimal medium containing  $^{15}\text{NH}_4\text{Cl}$  1 g/L, glucose 2 g/L,  $\text{Na}_2\text{HPO}_4$  6 g/L,  $\text{KH}_2\text{PO}_4$  3 g/L, NaCl 0,5 g/L,  $\text{MgSO}_4$  120 mg/L, thiamine 300 mg/L,  $\text{CaCl}_2$  11 mg/L and 100  $\mu\text{g/mL}$  kanamycin. Bacteria were diluted (1:100) into 0,5 L of minimal medium and grown at 37 °C until  $\text{OD}_{600} \approx 0.8$ . Protein expression was induced with 0.5 mM IPTG at 22 °C and cells were harvested after 16 h of incubation. The harvested cells were suspended in 30 ml of a cold buffer containing 50 mM Tris-HCl, 300 mM NaCl, pH 7.5 (Buffer A) supplemented with a protease inhibitor cocktail (Complete EDTA-FREE, Roche). The cells were sonicated 10 minutes (10s on / 10s off). The lysate was centrifuged at 12000 rpm for 30 min to remove cell debris and other particles. The supernatant was loaded on Hi-Trap column (GE Healthcare) equilibrated with Buffer A and eluted with a step of imidazole (500 mM). The His-tag was removed by TEV cleavage followed by an additional step of Ni-NTA affinity. The purity of c-prune and its mutants was evaluated by using 15% SDS-PAGE. The molecular mass was analyzed by ESI-Quadrupole mass spectroscopy. For NMR experiments proteins were extensively dialyzed against PBS buffer (pH 7.4).

### ***Circular dichroism measurements***

CD spectra were recorded with a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system (Model PTC-423-S). Far-UV measurements (195–250 nm) were carried out at 20° C in a buffer containing 10 mM Tris-HCl (pH 7.5), NaCl 100mM, 1.25% glycerol and dithiothreitol (DTT) 1mM using a 0.1 cm optical path length cell and a concentration of 40 $\mu\text{M}$  for gH-625-644 peptide and 8  $\mu\text{M}$  for all proteins analysed. CD spectra, recorded with a time constant of 4 s, a 2 nm band width, and a scan rate of 10 nm

$\text{min}^{-1}$ , were signal-averaged over at least three scans. The baseline was corrected by subtracting the buffer spectrum. Thermal denaturation curves were recorded over the 20–95 °C temperature interval, following the CD signal at 222 nm. The curve was registered using a 0.1 cm path length cell and a scan rate of  $1.0\text{ }^{\circ}\text{C min}^{-1}$ .

### ***Western blot analysis***

His-tagged gH625-c-prune (2 mg) was immobilized on Ni-NTA resin. To obtain cellular lysates from HeLa cells, they were re-suspended in PBS supplemented with 1% Triton X-100 and protease inhibitors. After 30 minutes on ice, cells were centrifuged at 12000 rpm for 30 minutes. The amount of total proteins present in the extract was determined by Bradford assay. Total extract of HeLa cells (4 mg of protein) was incubated with the resin at 4°C overnight. After extensive washing the elution was carried out with increasing imidazole concentration (300mM). The collected fractions were analysed by western blotting using an anti-Nm23-H1 antibody and an anti-his antibody as control of immobilized protein.

### ***Preparation of Small Unilamellar Vesicles***

Small Unilamellar Vesicles (SUVs) consisting of DOPC/Chol (1:1) and 25% of Br-PC, were prepared according to the extrusion method of Hope et al. <sup>[59]</sup> in 5 mM Hepes, 100 mM NaCl, pH 7.4. Lipids were dried from a chloroform solution under a stream of nitrogen-gas and lyophilized overnight. For fluorescence experiments, dry lipid films were suspended in buffer by vortexing for 1h and sonicated for 30 min.

### ***Tryptophan fluorescence experiments***

Br-PC employed as quencher of tryptophan fluorescence is suitable for probing membrane insertion of gH625-c-prune, since it acts over a short distance and does not drastically

perturb the membrane<sup>[60][61]</sup>. gH625-c-prune, containing the tryptophan residue in the peptide sequence, was added (final concentration of 4 $\mu$ M) to 2 ml of buffer (5mM HEPES, 100mMNaCl pH 7.4) containing 100 $\mu$ M of Br-PC/Chol SUV and a total lipid concentration of 400  $\mu$ M, thus establishing a lipid:peptide molar ratio of 100:1. Emission spectrum of the tryptophan was recorded with excitation set at 295nm. SUV composed of DOPC/Chol (1/1) and which contained 25% of either 6,7 Br-PC, or 9,10 Br-PC or 11,12 Br-PC were used. Three separate experiments were conducted for each condition. In control experiments, gH625-c-prune in DOPC/Chol (1/1) SUVs without Br-PC was used.

#### ***Labelling of gH625-c-prune and TAT-c-prune***

The gH625-c-prune, TAT-c-prune and c-prune were labeled with N-((2-(iodoacetoxy)ethyl)-N-Methyl)amino-7-Nitrobenz-2-Oxa-1,3-Diazole (IANBD, Invitrogen) a fluorophore that interacts specifically with reduced cysteines. The reaction was conducted in a buffer containing 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 5% glycerol, 1.5% D-glucose and 5mM DTT with a 10-fold molar excess of IANBD added as a stock solution (10 mg/ml in dimethylformamide) for 4 h at 4°C in slow agitation. The excess of IANBD was removed by extensive dialysis at 4°C in the reaction buffer.

#### ***Flow cytometry of cell association of gH625-c-prune***

Cell association was measured by flow cytometry using a FACScan (Becton Dickinson). NBD-labelled-gH625-c-prune (5  $\mu$ M) was added to HeLa cells (5 x 10<sup>5</sup> cells in 0.5 mL) in OptiMEM medium at 37°C in slow agitation for 10 min. Thereafter, the cells were washed in PBS, 0.1% BSA and then re-suspended in 0.5 mL of PBS, 0.1% BSA for flow cytometric analysis. The NBD-labelled-gH625-c-prune was excited at 465 nm and the fluorescence was measured at 530 nm. A stock solution of dithionite (1M) was freshly

prepared in 1 M Tris-HCl (pH 10). Following flow cytometric analysis, 5  $\mu$ L of dithionite stock solution was added to cells maintained at 4°C and the fluorescence of internalized gH625-c-prune was measured after 5 min.

### ***Cellular uptake kinetics and endocytosis inhibition of gH625-c-prune and TAT-c-prune by confocal microscopy***

For protein uptake-kinetics experiments on non-fixed living cells, 10<sup>5</sup> HeLa cells were incubated with 30  $\mu$ M of NBD-protein solutions at 37°C for 10 and 30 min. After incubation, cells were rinsed twice with PBS and fresh cell culture medium without phenol red was added.

To inhibit the endocytic pathways, HeLa cells were treated with sodium azide, cytochalasin D and low temperature. In particular, a 40  $\mu$ M sodium azide solution was added to the cell culture medium for 30 minutes followed by 30 min of incubation with proteins. Moreover, to disassemble actin microfilaments, HeLa cells were treated with 30  $\mu$ M cytochalasin D in cell culture medium for 30 min at 37 °C before protein incubation. After cytochalasin D treatment, cells were incubated with protein solution for 30 min. Furthermore, 4° C blocking experiments were carried out by incubating cells with protein solutions at 4° C for 30 min.

All the samples were observed by a confocal laser scanning microscope (CLSM) (LSM510, Zeiss) equipped with a 488 nm Argon laser line and with a 63 $\times$  water-immersion objective.

### ***Transfections***

Before transfection, HEK-293 cells were plated and incubated overnight at 37°C. The following day, immediately prior to transfection, medium was changed. Transfections were

achieved using calcium phosphate according to the protocol. 60% confluent HEK-293 were transfected with 15  $\mu\text{g}$  of DNA per 10-cm dish, using vectors encoding HA-NM23H1, HA-NM23H1-P96S and HA-NM23H1-S120G, each mutant was created as described by Reymond A. et al. [46]. The cells were collected 48 h later the transfection assays and the lysates were obtained as described above.

### ***NMR sample preparation***

NMR samples were prepared mixing  $^{15}\text{N}$ -labeled stock solution of c-prune and its mutants (100  $\mu\text{M}$  final concentration) with cell extracts (2 mg of total protein) in a final volume of 500  $\mu\text{L}$  supplemented with 10%  $\text{D}_2\text{O}$ . By this procedure, the overall extract dilution factor was kept constant and uniformly small (20%). To prevent c-prune degradation, protease inhibitors were added

### ***NMR measurements and analysis***

All experiments were recorded on a Varian INOVA 600-MHz NMR spectrometer equipped with a cold-probe and at room temperature. 2D [ $^1\text{H}$ - $^{15}\text{N}$ ] heteronuclear single quantum coherence (HSQC) correlation spectra were acquired by using a fast version of sensitivity enhanced HSQC pulse-sequence. Lysate data were recorded with 16 transients and 1024 ( $^1\text{H}$ ) x 64 ( $^{15}\text{N}$ ) complex points. The reference experiments of pure samples in NMR buffer were obtained with identical spectrometer settings. NMR peaks parameters, like signal intensities, background noise and chemical shift values, were independently extracted and analyzed with Neasy, a tool of CARRA software. Intensity ratios were computed as  $\text{Int}(\text{buffer})/\text{Int}(\text{lysates})$  and  $\text{Int}(\text{lysates})/\text{Int}(\text{lysates overexpressing c-prune partners})$ .

## Results

## RESULTS

### *Cloning and expression of gH625-c-prune and TAT-c-prune*

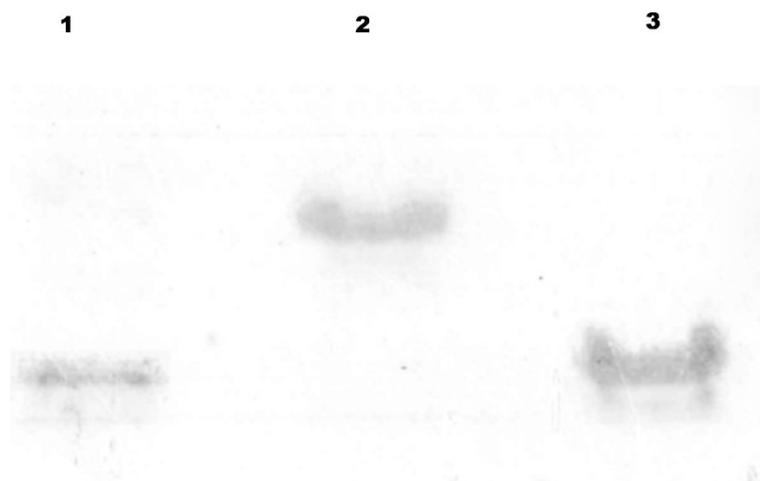
In order to obtain a protein able to enter into cells and to study it in its physiological conditions, we designed two constructs containing c-prune fused with two different peptides for cellular internalization, gH625-644 and TAT. The coding sequence of gH625-644 (60 bp) was amplified by polymerase chain reaction using specific primers for genomic DNA of Herpes simplex virus 1 (HSV-1). The digested fragment was ligated to the gene encoding c-prune. The fused segment was finally inserted into the expression vector pET28b+ which provided a cleavable N-terminal His<sub>6</sub> tag.

The coding sequence of TAT was fused to c-prune cDNA sequence performing two consecutive PCR reactions as describe in the experimental section. The coding sequence of TAT-c-prune was inserted into the expression vector pETM11 which provided a cleavable N-terminal His<sub>6</sub> tag. The recombinant clones of both constructs obtained were sequenced by Primm sequencing service. Different cell strains and environmental conditions (temperature and IPTG concentrations) were tested to optimize the expression of the recombinant proteins. The best level of expression were achieved using *E. coli* BL21 STAR (DE3) cells grown at 37 °C and induced overnight with 0.5 mM IPTG at 22 °C.

### *Purification of gH625-c-prune and TAT-c-prune*

Recombinant gH625-c-prune and TAT-c-prune were purified from inclusion bodies using a nickel affinity chromatography in denaturing conditions. The collected fractions were subjected to a refolding process via direct dilution to prevent aggregation and precipitation of the protein. The refolded proteins were subjected to proteolytic digestion with thrombin (gH625-c-prune) and TEV protease (TAT-c-prune), in the optimized conditions, to remove the His-tag with an efficiency of about 60%. The digested sample was loaded onto a

Superdex 75 gel filtration column pre-equilibrated with an appropriate buffer (see Experimental Procedures) to separate the protein aggregates from the properly folded forms. The purity of the proteins were assessed by SDS-PAGE (Fig. 11). Since gH625-c-prune and TAT-c-prune have shown the tendency to aggregate during their concentration it was not possible to concentrate them over 30  $\mu$ M.

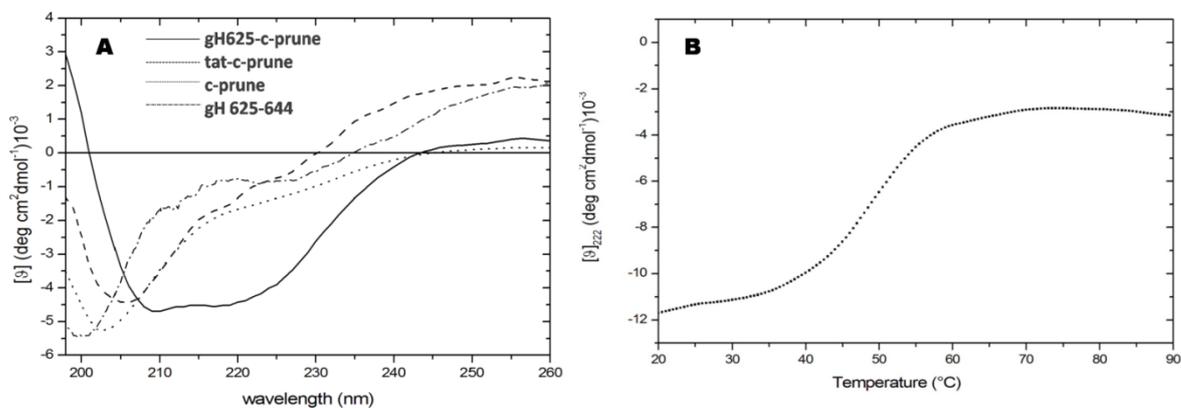


**Fig. 11** Electrophoretic analysis carried out on 15% SDS-PAGE. Tat-c-prune (1), gh625-c-prune (2), and c-prune (3) purified to homogeneity.

### ***Circular dichroism measurements***

The secondary structure of gH625-c-prune, TAT-c-prune, c-prune and gH625-644 peptide was determined by CD spectroscopy in the far UV spectral region (195–260 nm). CD spectra show that gH-625-c-prune protein has a degree of structure in solution which was absent in the protein alone and in the peptide alone. In fact, circular dichroism studies performed on c-prune<sup>[41]</sup>, TAT-c-prune and on gH625-644 peptide highlight the lack of any structure (**Fig 12A**). Instead, the CD spectrum of the recombinant protein gH625-c-prune, in the same conditions, is characterized by the presence of two minima (at 208 and 222 nm) and one maximum (at 195 nm) which are typical fingerprints of  $\alpha/\beta$  proteins (**Fig**

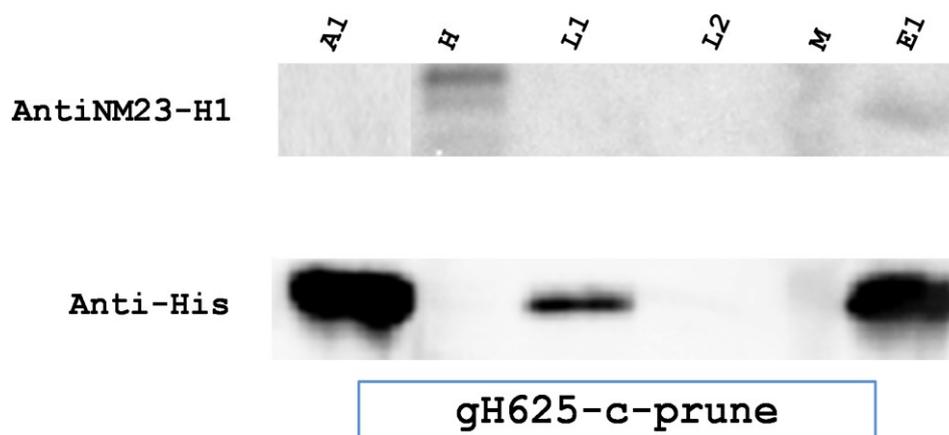
**12A).** The thermal denaturation curve of gH625-prune presents a sigmoidal shape with a single inflection point, corresponding to Td values of about 52 °C (**Fig. 12B**).



**Fig. 12** Far-UV CD spectra. **A)** Far-UV CD spectra of gh625-c-prune (solid line), tat-c-prune (dashed line), c-prune (dotted line) and gH 625-644 (dashed-dotted line) in 10mM Tris-HCl (pH 7.5), NaCl 100mM, 5% glycerol and dithiothreitol (DTT) 1mM. The spectra were recorded using 0.1 cm path length cells. The protein and peptide concentration were 8  $\mu$ M and 40  $\mu$ M, respectively. **B)** Temperature-induced denaturation curves of gh625-c-prune obtained by recording the molar ellipticity at 222 nm in 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5% (v/v) glycerol and DTT 1mM.

### ***Analysis of gH625-c-prune biological function***

To verify if c-prune retained its biological activity after conjugation with the peptide gH625, we verified the complex formation with Nm23-H1 by western blotting<sup>[41]</sup>. His tagged gH-625-c-prune was immobilized on the resin and total extract of HeLa cells was loaded onto it (see experimental procedures). The western blotting analysis showed that gH-625-c-prune is still able to interact with Nm23-H1 in HeLa extract (**Fig. 13**), suggesting that the folding of the protein does not affect its biological function.

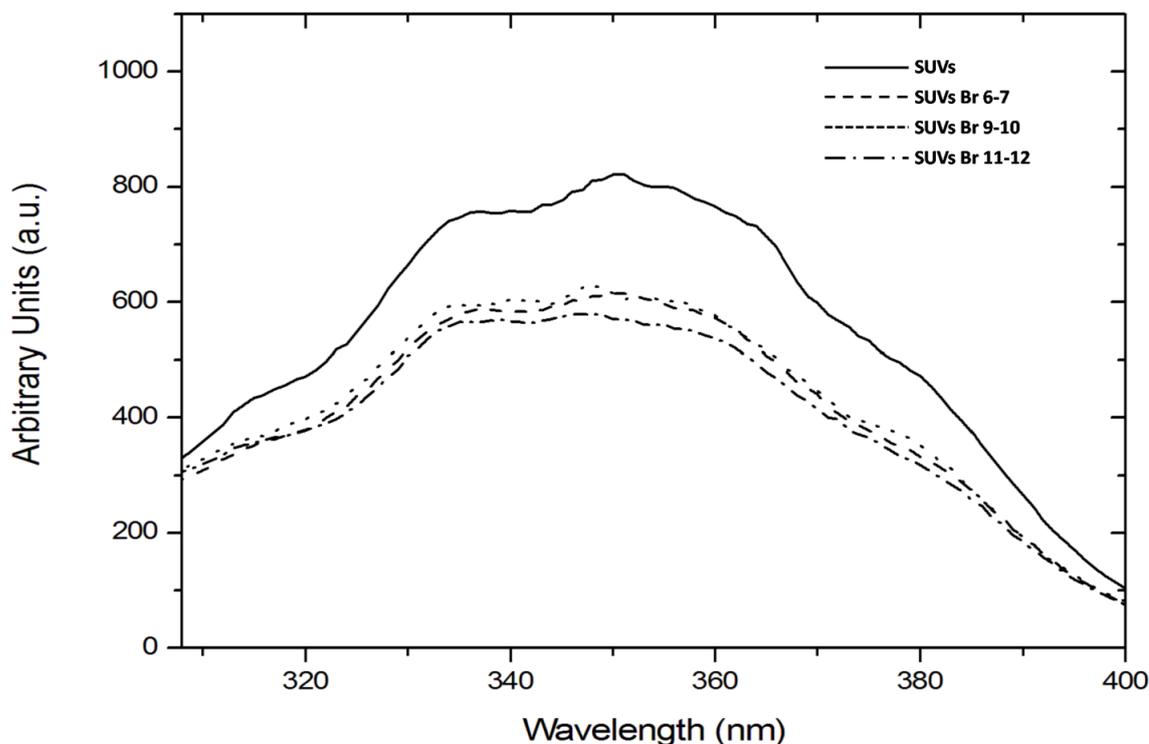


**Fig. 13** Western blotting analysis of gH-625-c-prune biological function. Lane A1: His tagged gH-625-c-prune. Lane H: HeLa total extract. Lane L1 wash with 50mM imidazole. Lane L2: wash with 50mM imidazole of resin without gh625-c-prune immobilized. Lane M: molecular weight markers. Lane E1: elution with 300mM imidazole.

### ***Fluorescence measurements***

A tryptophan residue naturally present in the sequence of a protein or a peptide can serve as an intrinsic probe for the localization of the peptide within a membrane; in fact, the fluorescence emission of a tryptophan residue increases when the amino acid enters a more hydrophobic environment, and together with an increase in quantum yield, the maximal spectral position will be shifted toward shorter wavelengths (blue shift). The peptide gH625 contains a tryptophan residue in the middle of the sequence; Fig. 14 shows the fluorescence emission spectra of the peptides gH625 upon interaction with Br-DOPC/Chol vesicles. In this experiment the gH625-prune location inside the bilayer was investigated by measuring the relative quenching of the gH625 Trp fluorescence by the probes 11,12-Br-PC, 9,10-Br-PC and 6,7-Br-PC, which differ in the position of the quencher moiety along the hydrocarbon chain and allow to establish the depth of the peptide insertion into the membrane. For all the probes used, it is evident a quench of the Trp emission. These

results indicate that, upon binding to vesicles, the peptide is deeply inserted into the membrane bilayer.

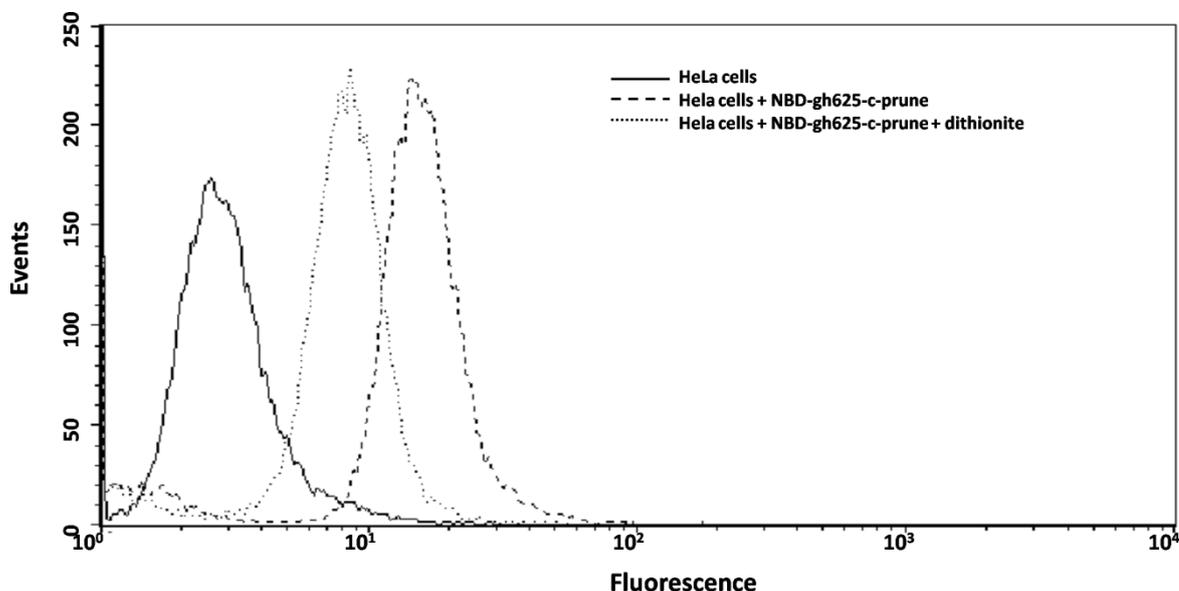


**Fig. 14 Tryptophan fluorescence emission spectra.** Tryptophan fluorescence emission spectra of the gH625-c-prune upon interaction with Br-DOPC/Chol vesicles. gH625-c-prune incubated with SUV (solid line), with SUV which contained 6,7 Br-PC (dashed line), with SUV which contained 9,10 Br-PC (dotted line) and with SUV which contained 11,12 Br-PC (dashed-dotted line).

#### ***Determination of NBD-gH625-c-prune cellular uptake***

We have determined the fraction of NBD-labeled gH625-c-prune taken up into HeLa cells by flow cytometry. This fraction is determined by comparing the fluorescence intensity before and after addition of sodium dithionite, an essentially membrane-impermeant molecule, which suppresses irreversibly the fluorescence of the accessible NBD-moiety localized on the external cell surface. We incubated NBD-gH625-c-prune (5 $\mu$ M) with HeLa cells at 37°C. After 10 min of incubation, we measured the quenching of the NBD-gH625-c-prune by dithionite treatment. The dithionite reaction was performed at low temperature (4 °C) because the dithionite diffusion in the biological membranes is strongly

reduced in comparison with high temperatures<sup>[62]</sup>. After the addition of dithionite there is an increase of fluorescence compared to the intrinsic fluorescence of the cells (Fig.15).



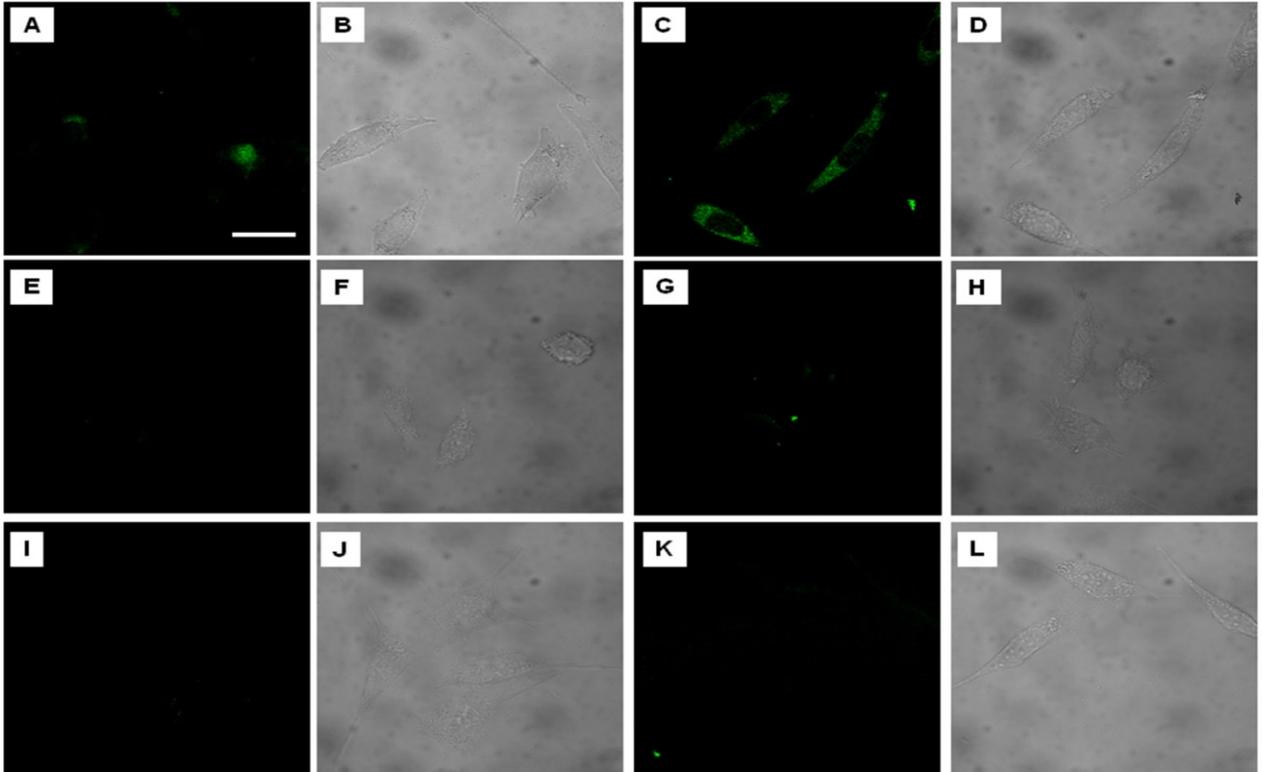
**Fig. 15** Flow cytometric analysis of HeLa cells treated with NBD-gh625-c-prune. Cells are treated for 10 min with 5  $\mu$ M NBD-gh625-c-prune in the presence or absence of dithionite. Not treated cells (solid line), treated with NBD-gh625-c-prune (dashed line), treated with NBD-gh625-c-prune and dithionite (dotted line). Single result representative of three similar experiments.

This shows that a percentage of protein (ca. 50%) has been internalized in 10 min. Increasing protein concentration, time of incubation or dithionite concentration did not result in an increase of yield's internalization. NBD was used in the same experiment to exclude that the fluorescence signal was due to internalization of the fluorophore alone and no cellular fluorescence was observed at concentration up to 15  $\mu$ M (data not shown).

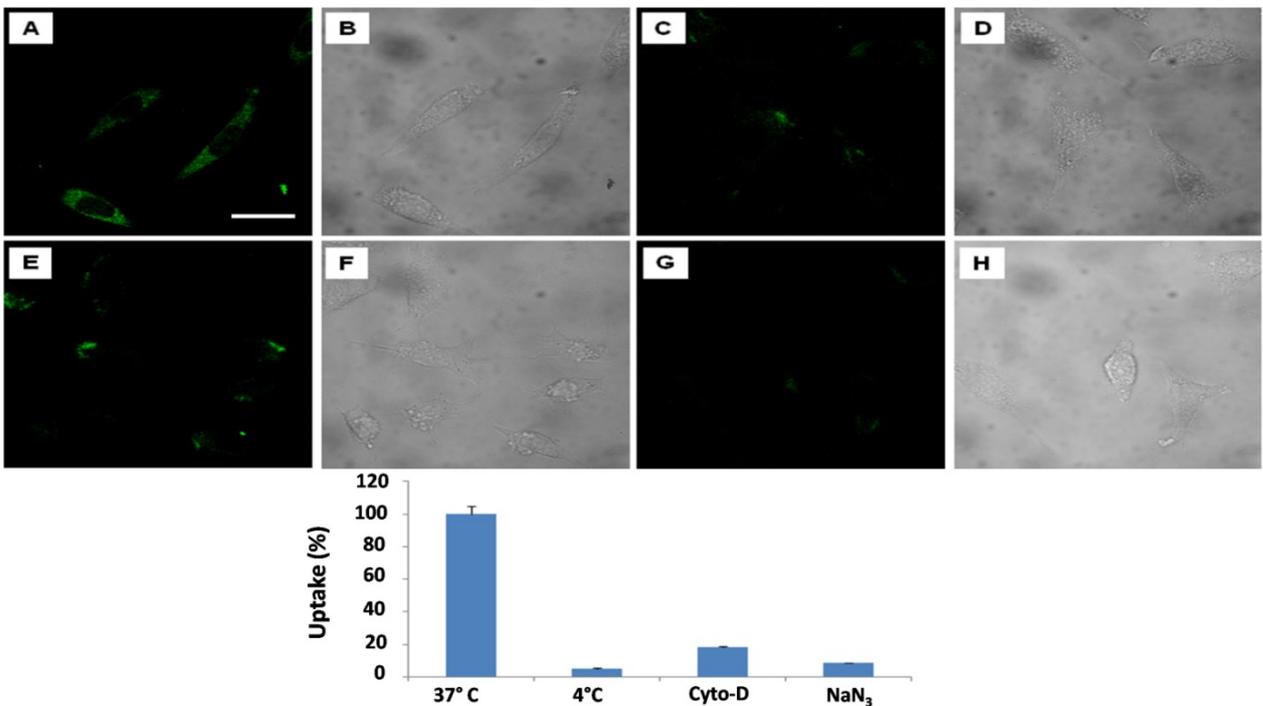
#### ***Uptake kinetics and endocytosis inhibition***

To verify if gH625 peptide could promote c-prune internalization, we examined uptake in HeLa cells by confocal microscopy. We incubated 30  $\mu$ M gH625-c-prune solution in cell culture medium for 10 and 30 min at 37°C. Images show a fluorescence signal in HeLa cells already after 10 min incubation for gH625-c-prune (**Fig. 16A**). Moreover, by increasing incubation time, the intracellular fluorescence intensity increased (**Fig. 16C**).

Conversely, no fluorescence was observed in case of c-prune (**Fig. 16 I and K**) and only a slight intracellular fluorescence was detectable after incubation with TAT-c-prune (**Fig. 16 E and G**). gH625-c-prune uptake was partially due to endocytic mechanisms as demonstrated by endocytosis inhibition experiments. In particular, gH625-c-prune uptake was significantly inhibited by  $\text{NaN}_3$  and 4 °C treatments (**Fig.17 C and G**). In addition, 80% uptake inhibition was observed after incubation with cytochalasin D (**Fig. 17 E**). These data demonstrate that gH625-c-prune internalization significantly depends on active, ATP dependent mechanisms and suggest that there is a small amount of gH625-c-prune that use a different, probably passive, mechanism to cross cell membrane.



**Fig. 16 Uptake kinetics.** Uptake kinetics of HeLa cells incubated at 37°C with gh625-c-prune (A-D), TAT-c-prune (E-H) and c-prune (I-L) for 10 (left panels) and 30 (right panels) min. Fluorescence (A, C, E, G, I and K) and transmitted light (B, D, F, H, J and L) images. Magnification bar: 20  $\mu$ m.



**Fig. 17 Uptake inhibition experiments.** Confocal laser scanning microscope images of HeLa cells incubated with gh625-c-prune protein at 37°C (A and B), NaN<sub>3</sub> (C and D), cytochalasin D (E and F) at 4°C (G and H). Fluorescence (A, C, E and G) and transmitted light (B, D, F and H) images. I. Percentage of uptake as a function of inhibition treatments. Magnification bar: 20  $\mu$ m.

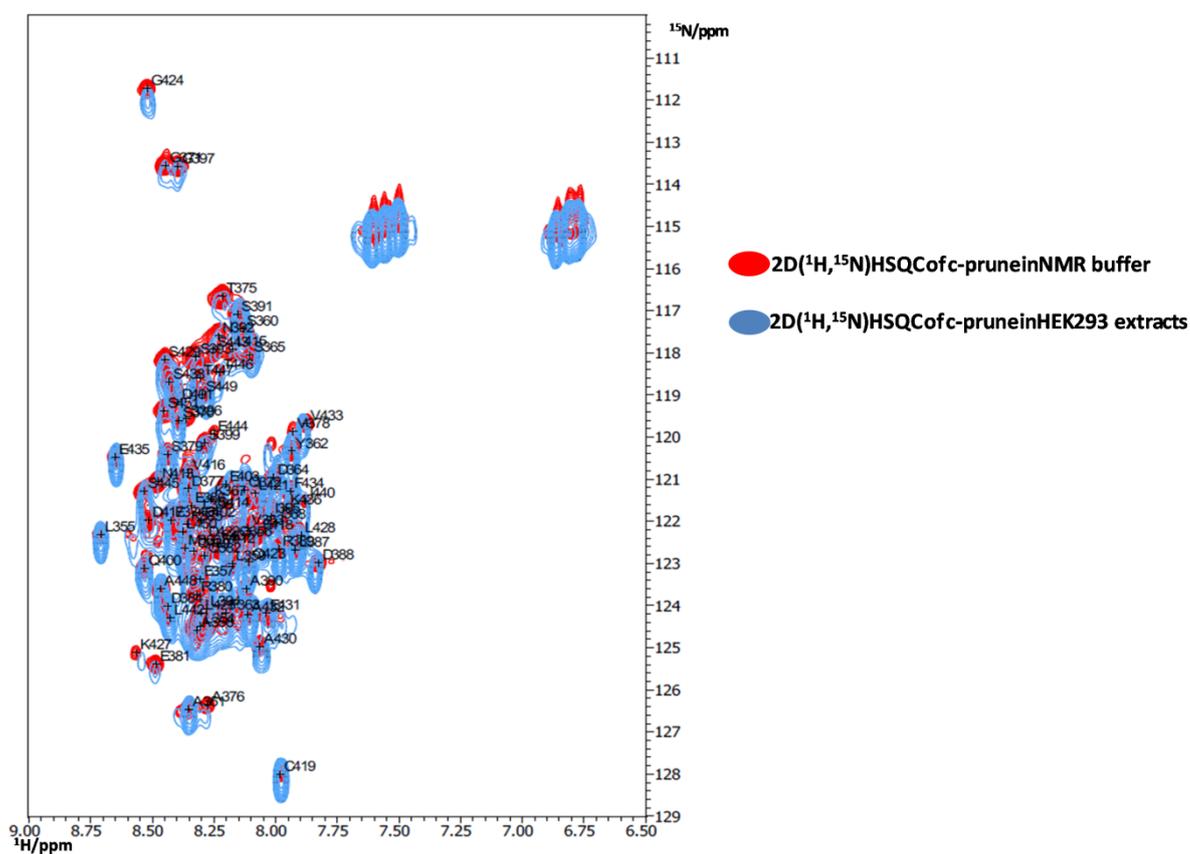
***h-prune/Nm23-H1 interaction in unlabeled lysates.***

Due to the low concentration of recombinant gH625-c-prune samples obtained a new route has been undertaken. We performed some NMR experiments by using eukaryotic cellular extracts which potentially provide further information about the binding mode of these proteins close to physiological conditions. Therefore, we began our cellular analyses by adding labeled c-prune to HEK-293 lysates. Despite the high abundance of endogenous, unlabeled proteins, the resulting 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectra of c-prune in HEK-293 lysates were superimposable with the reference spectra of the protein in buffer solution (Fig. 18), confirming that c-prune retains the same conformation in cellular environment and does not interact with any of the endogenous components in the lysate. However the average c-prune amide resonance signal intensity in the HEK-293 lysate was notably reduced with an overall increase in proton line-widths correlating well with the expected increase in the protein total correlation time ( $\tau_c$ ) due to cellular viscosity<sup>[12]</sup>. To investigate the specific association between c-prune and Nm23-H1 as much as possible closest to physiological conditions, a starting solution of purified  $^{15}\text{N}$ -labeled c-prune was added to HEK-293 lysate overexpressing Nm23-H1 (Fig. 19).

The appearance of some dispersed peaks, together with the disappearance or the decrease in intensity of several signals in the c-prune [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectrum, indicate that the protein in the presence of HEK-293 lysates is able to form a specific complex with Nm23-H1, in agreement with the results observed in previous experiments *in vitro*<sup>[41]</sup>.

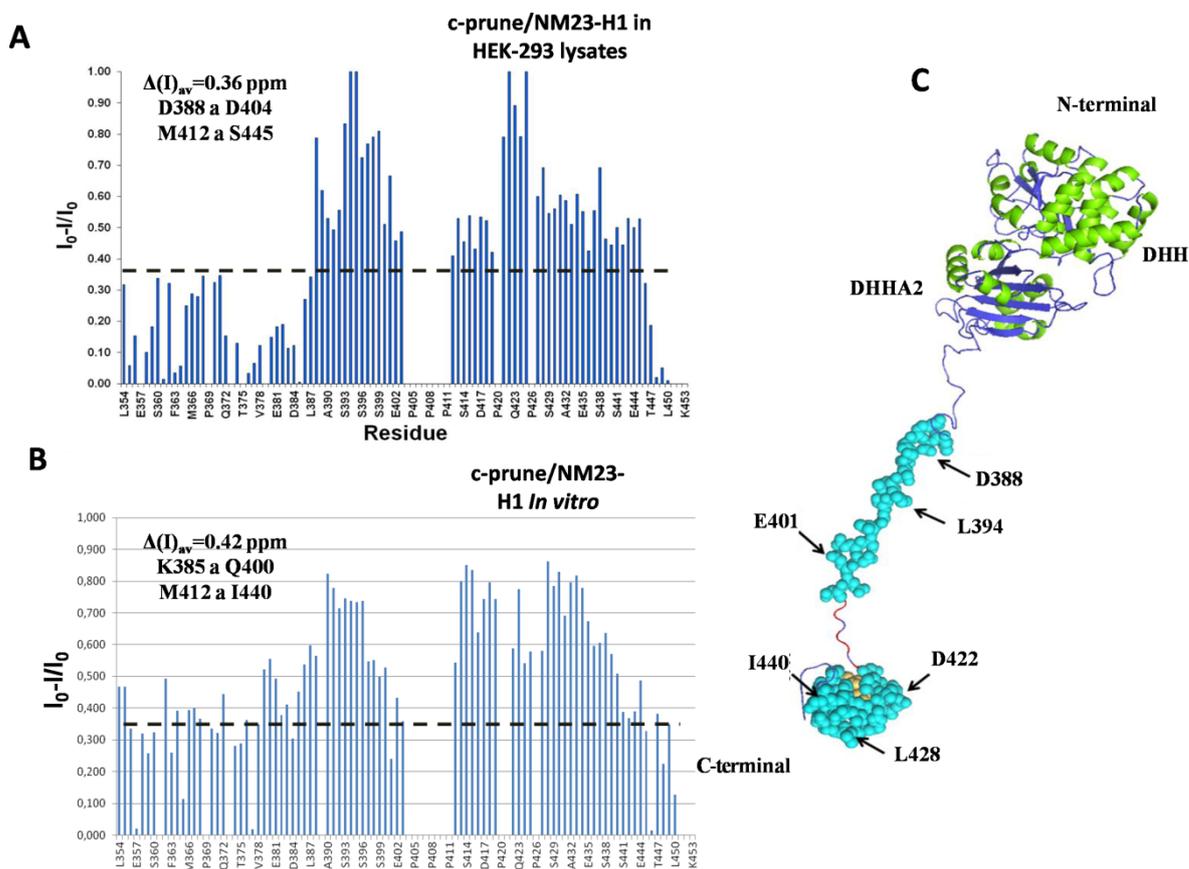
In effect, we have demonstrated that the Nm23-H1 binding epitope on c-prune surface is located in the region encompassing amino acids from 388 to 404 and in the more compact region constituted by the disulfide-bridged cycle and the  $\alpha 3$  helical region (amino acids 412-445), as showed in Fig. 20 A-B-C.

To further confirm the specific interaction of c-prune/Nm23-H1, the NMR binding experiments were repeated by using human NBL SH-SY5Y lysates expressing high levels of Nm23-H1. As showed in Fig. 21, SH-SY5Y extracts confirmed that, upon the complex formation, Nm23-H1 retains the binding to the two regions (from L387 to D401 and from N413 to E444) on the surface of c-prune.

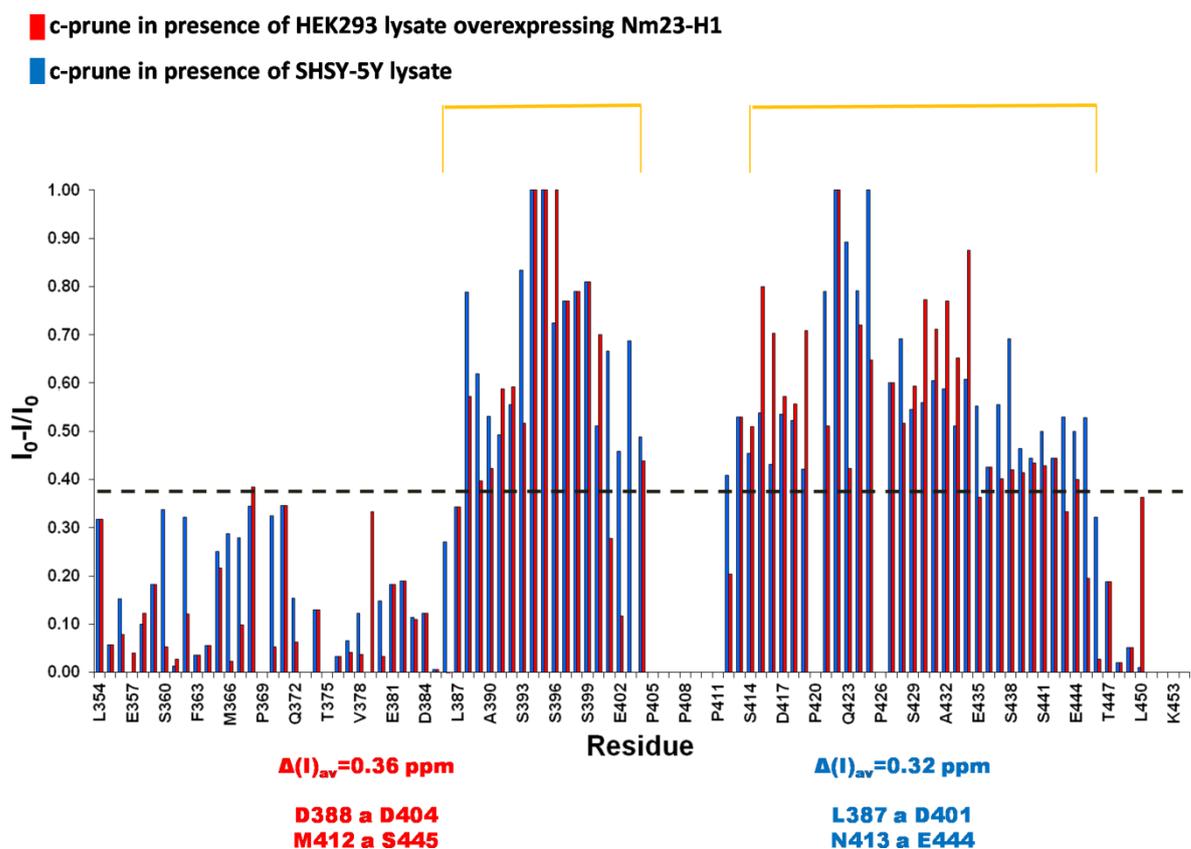


**Fig. 18** Overlay of the 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC correlation spectra of purified c-prune in NMR buffer solution (in red) and of c-prune in HEK-293 lysate (in blue).





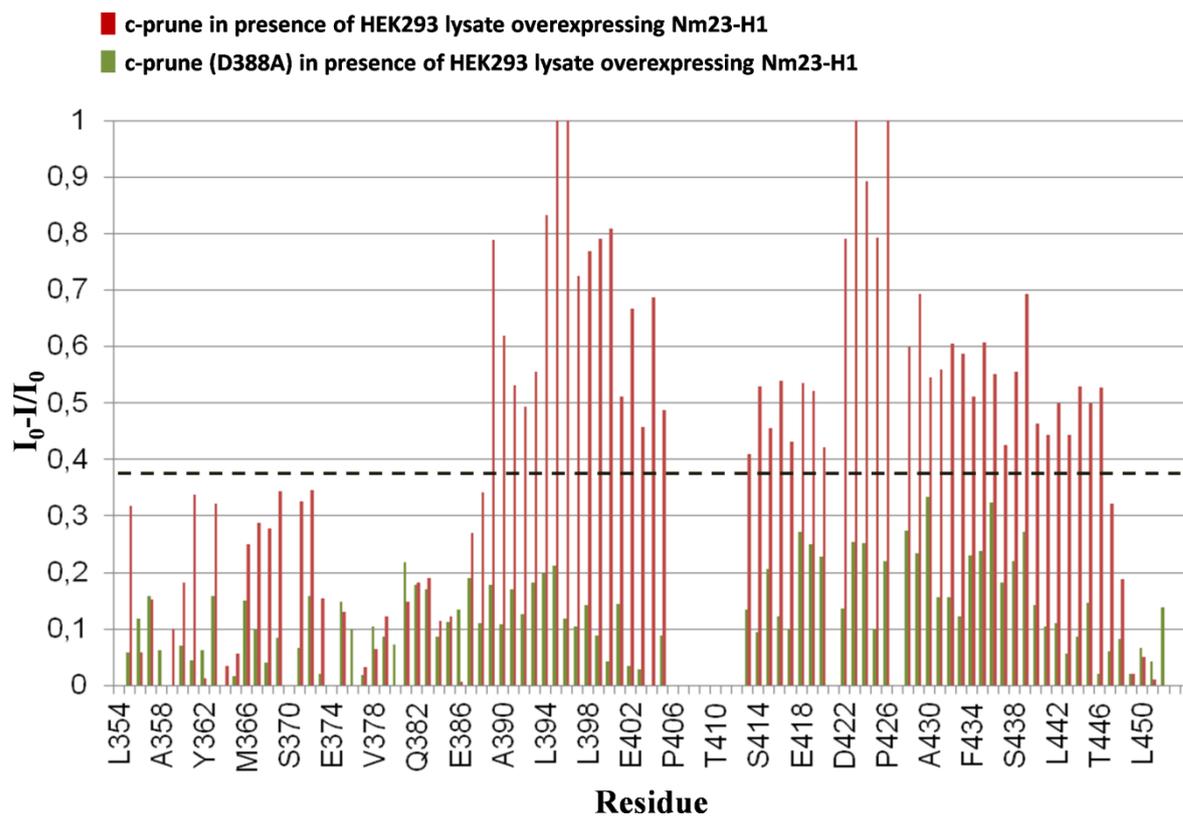
**Fig. 20** **A)** Intensity ratios for all c-prune residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in HEK-293 lysate overexpressing Nm23-H1. **B)** Intensity ratios for all c-prune residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments *in vitro* using recombinant Nm23-H1. **C)** The amino acids showing large variations upon complex formation with Nm23-H1 ( $I_0-I/I_0 \geq 0.41$ ) are mapped in cyan onto the h-prune surface. The region highlighted in red correspond to the polyproline sequence while the region coloured in yellow correspond to disulphide bridge between Cys419 and Cys437.



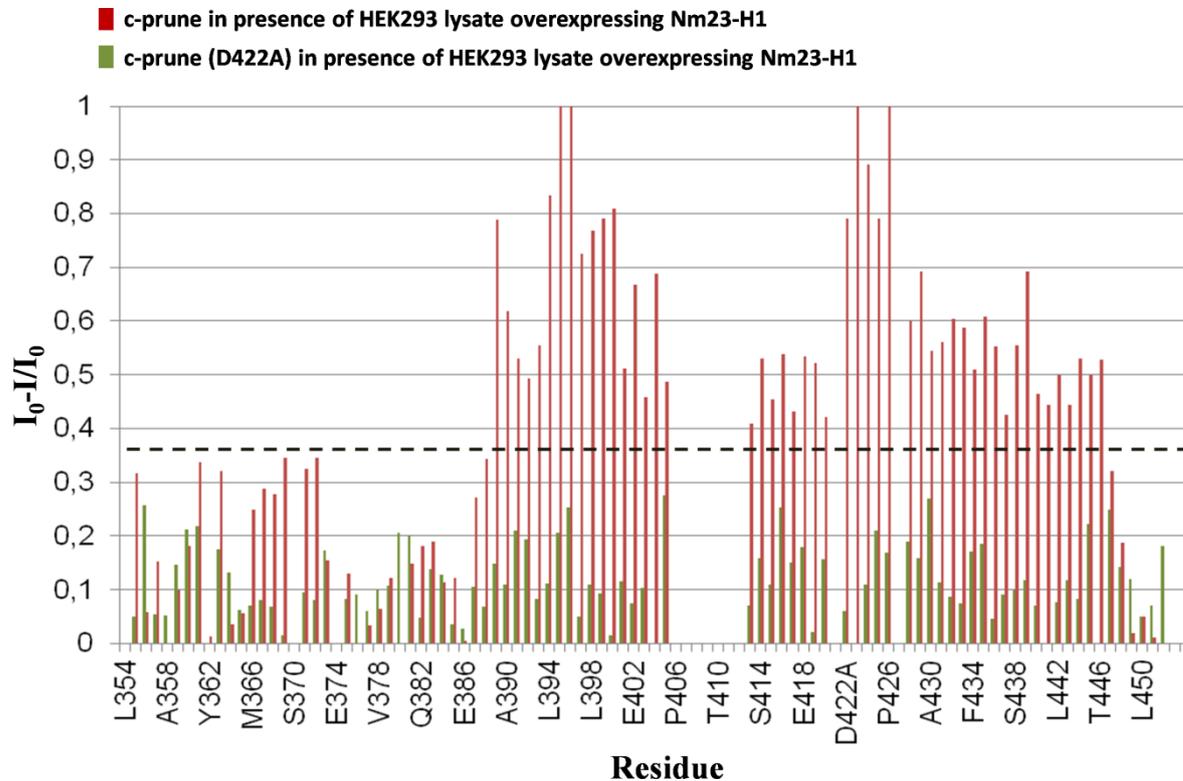
**Fig. 21** Intensity ratios for all c-prune residues ( $I_0-I/I_0 \geq 0.36$ ) extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in SH-SY5Y lysate (in blue) and compared to the data obtained in HEK-293 lysate (in red).

Following the NMR mapping interaction epitopes for c-prune, we used site-directed mutagenesis to further validate the binding site for Nm23-H1. Because negatively charged amino acids are thought to contribute to the c-prune/Nm23-H1 complex, we aimed to disrupt the interaction by uniformly replacing residues D388 and D422 with alanine; therefore, these two protein mutants, named D388A and D422A, were produced, purified and characterized. Earlier functional studies indicated that both h-prune mutants did not induce cell migration and that the mutants interacted weakly with Nm23-H1<sup>[41]</sup>. To validate the experimental functional studies we investigated the binding of the two c-prune mutants to Nm23-H1 by adding HEK-293 lysates overexpressing Nm23-H1 to a starting solution of <sup>15</sup>N-labeled D388A or <sup>15</sup>N-labeled D422A. Effectively, D388A and D422A did not show a significant change in peak intensities despite c-prune wild type, as indicated in the Figure 22 (for D388A) and Figure 23 (for D422A), confirming that just a single point mutation on the c-prune binding epitope greatly compromised its ability to interact with the Nm23-H1 *in vivo*.

Guided by the NMR interaction analysis (Figure 20 C), we designed another c-prune single point mutation affecting the Nm23-H1 binding site. In particular, we mutated L394 in alanine to evaluate whether uncharged residues also contribute to the complex formation.



**Fig. 22** Intensity ratios for all c-prune residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments (in red) compared to the data of c-prune (D388A) (in green) performed in HEK-293 lysate overexpressing Nm23-H1.

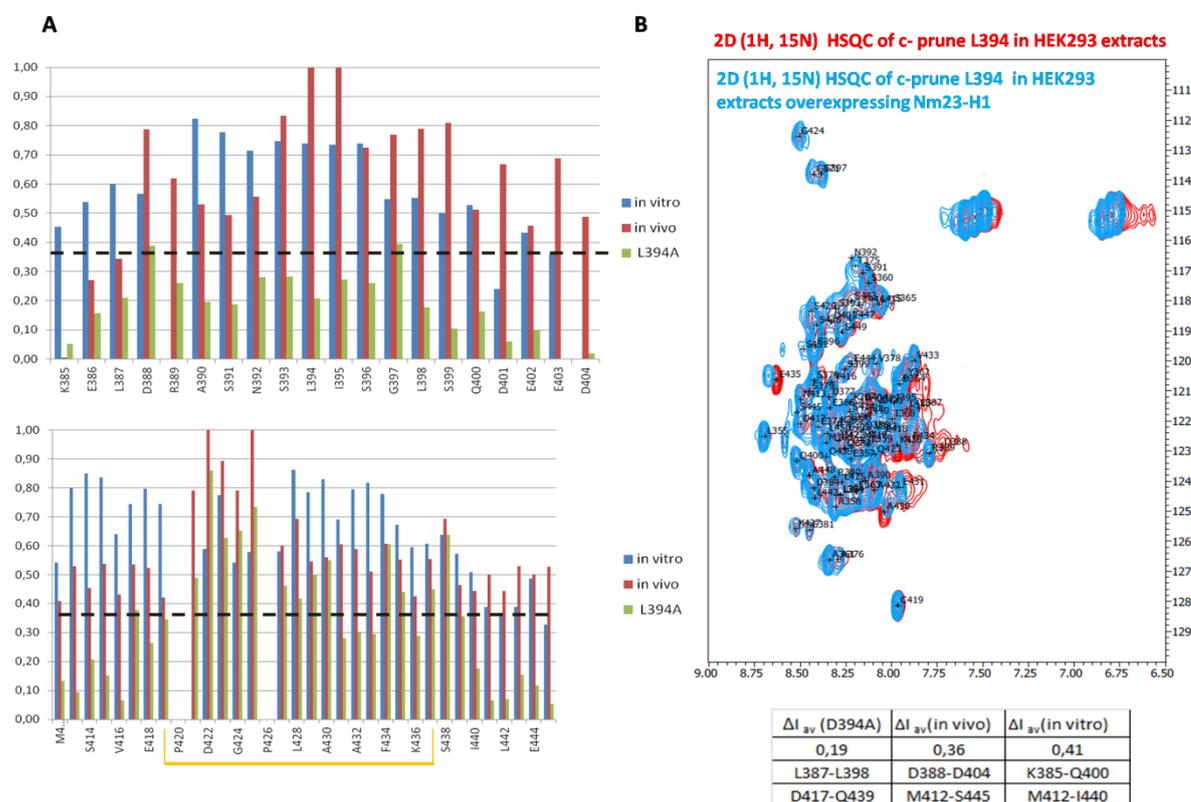


**Fig. 23** Intensity ratios for all c-prune residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments (in red) compared to the data of c-prune (D422A) (in green) performed in HEK-293 lysate overexpressing Nm23-H1.

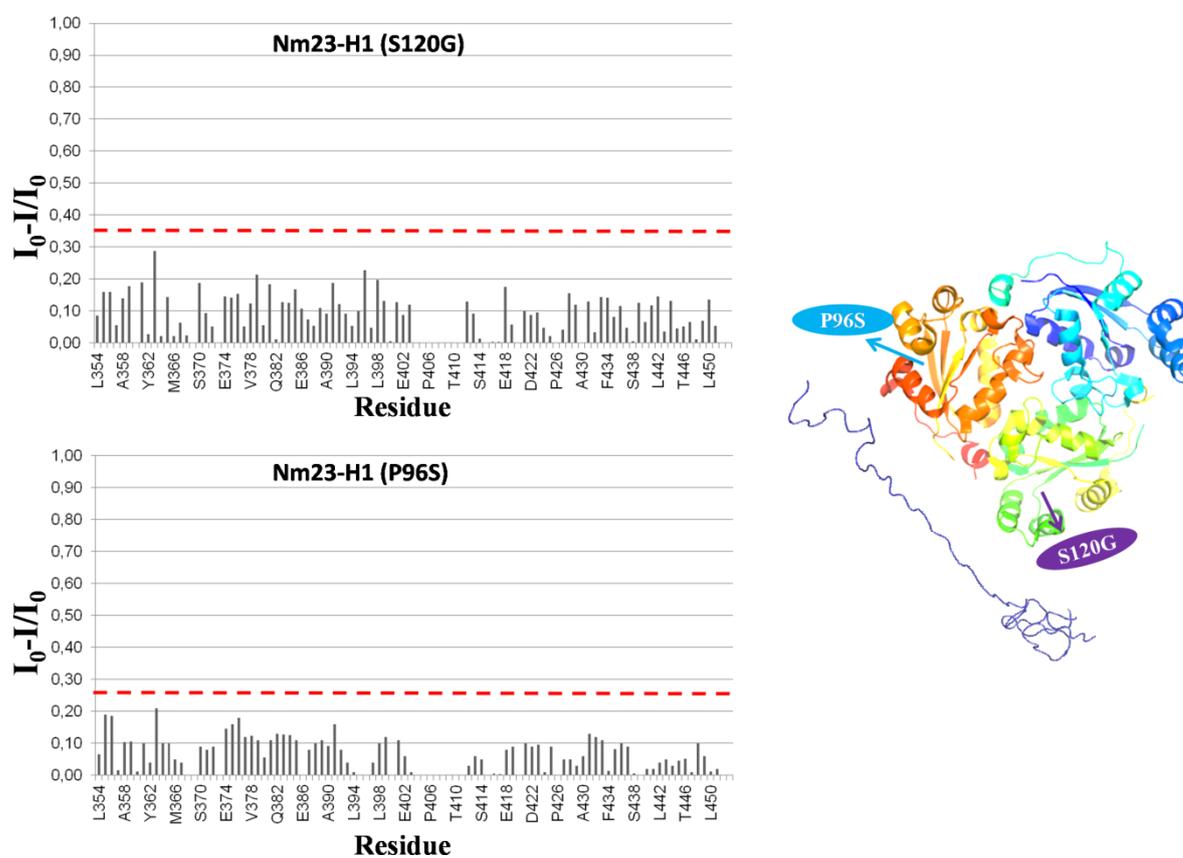
NMR binding experiments of the  $^{15}\text{N}$ -labeled L394A mutant in HEK-293 lysate show features of a remnant, substantially reduced interaction (Fig. 24 A & B).

To further validate our system we used cellular lysates transfected with two well characterized single point-mutants of Nm23-H1, P96S and S120G, both known to alter Nm23-H1 biochemical activity or structure affecting the cell motility. When oligomeric patterns of the recombinant Nm23-H1 proteins were examined by SDS-PAGE after cross-linking the protein, most oligomers in wild type Nm23-H1 were found to be hexamers, and dimers were less than 20% in the wild type. In contrast, no hexamers were found in P96S and S120G mutant Nm23-H1 proteins, and much higher proportions of dimers and monomers were observed in these mutant proteins compared to the wild type. The alteration in the oligomeric structure of Nm23-H1 protein by P96S and S120G mutations was also observed in gel filtration chromatography, in which most of the P96S and S120G mutant proteins eluted at a position corresponding to the molecular size of dimeric form, while a major protein peak in the wild type was detected at an elution volume similar to hexameric form. The altered oligomeric structure observed only in P96S and S120G mutant proteins suggests the intriguing hypothesis that the oligomeric structure of Nm23-H1 may be functionally linked to its metastasis suppressing activity. Reduction or loss of hexamers in Nm23-H1 by the mutations such as P96S and S120G could result in increased metastatic potential of cancer cells. P96S and S120G mutant proteins incapable of motility suppression exhibited reduced histidine protein kinase activity and autophosphorylation level, respectively. Since reduction of autophosphorylation level reflects decreased phosphoenzyme intermediate formation that is involved in histidine protein kinase activity as well as NDP kinase activity, these observations imply the fact that those mutations reducing histidine protein kinase activity of Nm23-H1, result in suppressive capacity on metastatic phenotypes <sup>[63]</sup>. We resuspended  $^{15}\text{N}$  labeled c-prune in HEK-293 extracts over

expressing each Nm23-H1 mutant. The signal intensities of the Nm23-H1 S120G/P96S extracts were comparable to those observed in the corresponding extracts not overexpressing Nm23-H1 (Fig. 25), demonstrating that a single mutation of Nm23-H1 can affect the interaction with c-prune.



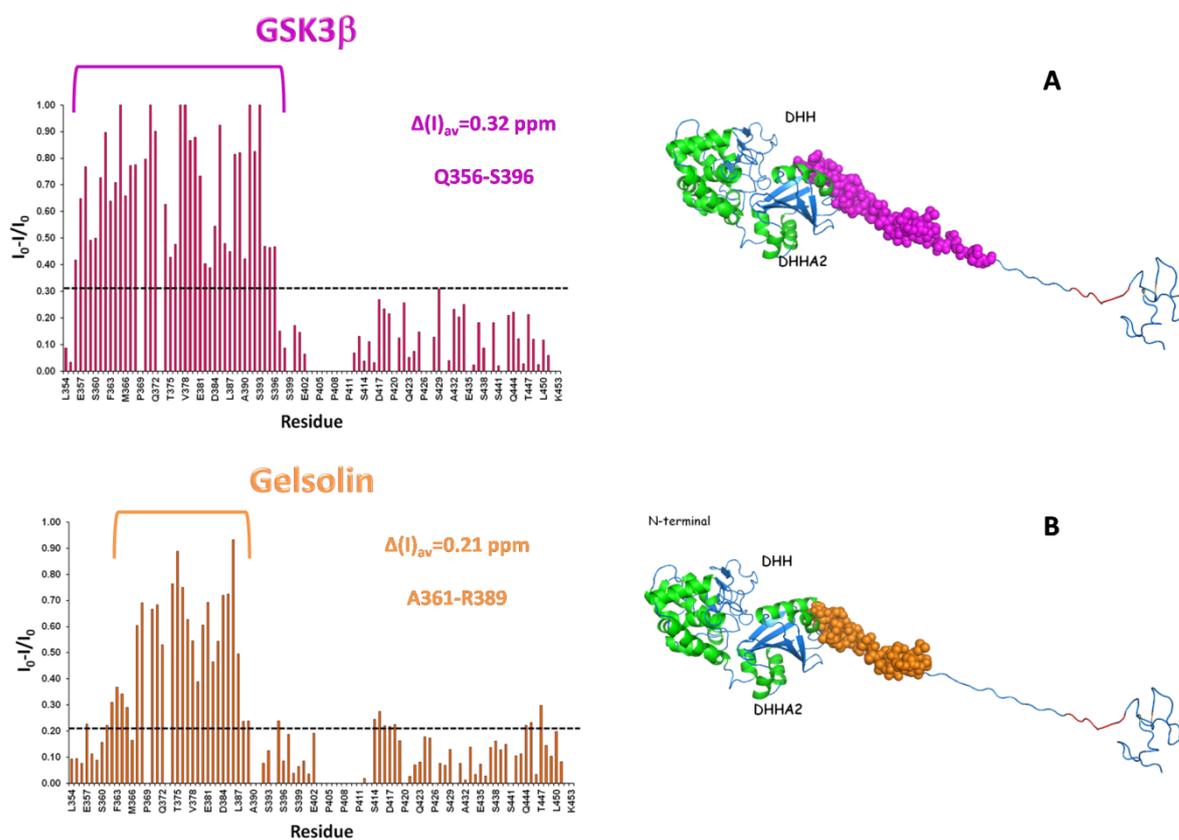
**Fig. 24** **A)** Intensity ratios for all c-prune (L394A) residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments (in green) compared to the data of c-prune wild type (in red) performed in HEK-293 lysate overexpressing Nm23-H1 and to the data of c-prune/Nm23-H1 interaction in vitro (in blue). **B)** Overlay of the 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC correlation spectra of purified c-prune (L394A) in HEK293 lysate (in red) and of c-prune (L394A) in HEK-293 lysate overexpressing Nm23-H1 (in blue). In table regions of interaction identified for each experiment.



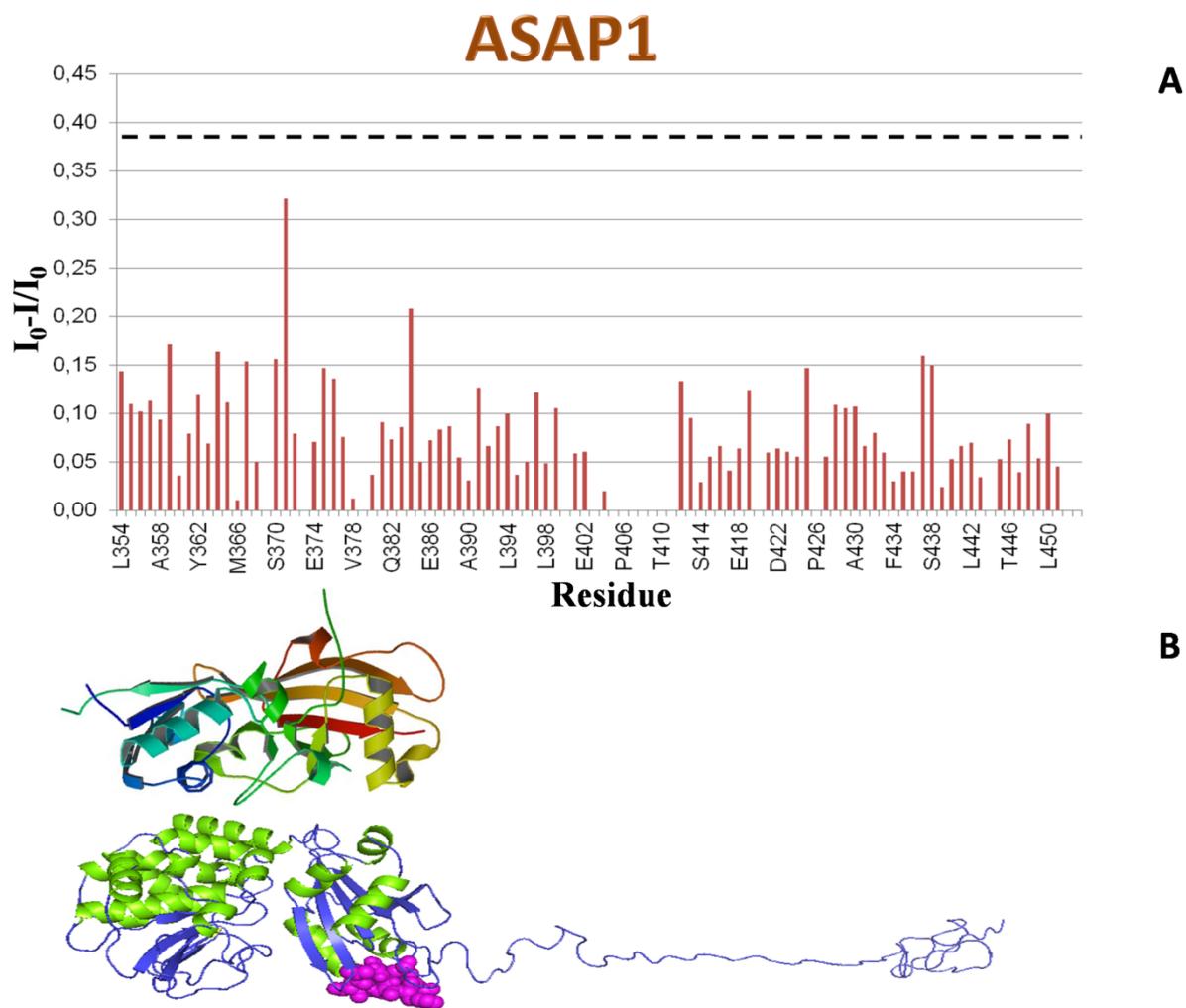
**Fig. 25** Intensity ratios for all c-prune residues extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in HEK-293 lysate overexpressing Nm23-H1 (S120G) up and Nm23-H1 (P96S) down.

***h-prune interactions with GSK3beta and gelsolin in unlabeled lysates***

Moreover, we extended functional studies in cell lysates to other h-prune endogenous partners, such as GSK3 $\beta$ , gelsolin and ASAP1. Upon addition of HEK-293 extracts overexpressing GSK3 $\beta$  or gelsolin or ASAP1 to starting solutions of  $^{15}\text{N}$  labeled c-prune the NMR peaks weakened, indicating c-prune interaction with each of the proteins. In particular, residues between Q356 and S396 and residues from A361 to R389 of the  $^{15}\text{N}$  labeled c-prune show an extensive line broadening in the presence of HEK-293 lysate overexpressing GSK3 $\beta$  and gelsolin, respectively (Fig. 26 A & B). Interestingly, in cell lysate NMR experiments outline GSK3 $\beta$  and gelsolin binding sites that lie on the N-terminal portion of c-prune, indicating a remarkable binding versatility of this domain, likely due mostly to its intrinsically disordered nature. Indeed, residues 353-365 of c-prune constitutes the C-terminal region of h-prune DHHA2 domain and thus GSK3 $\beta$  and gelsolin interaction also with this domain cannot be excluded. On the other hand, in the presence of cellular extract overexpressing ASAP1 no changes in signal intensities in the c-prune [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectrum were observed corroborating the hypothesis that this partner does not interact with the C-terminal domain of h-prune but, probably, with the other two domains, DHH and DHHA2, involved in the catalytical activity and therefore able to stimulate h-prune phosphodiesterase activity. (Fig. 27 A & B)<sup>[47]</sup>.



**Fig. 26** **A)** Intensity ratios for all *c*-prune residues ( $I_0-I/I_0 \geq 0.32$ ) extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in HEK-293 lysate overexpressing GSK-3 $\beta$ . The amino acids showing large variations upon complex formation with GSK-3 $\beta$  are mapped in magenta onto the *h*-prune surface. **B)** Intensity ratios for all *c*-prune residues ( $I_0-I/I_0 \geq 0.21$ ) extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in HEK-293 lysate overexpressing Gelsolin. The amino acids showing large variations upon complex formation with gelsolin are mapped in orange onto the *h*-prune surface. The region of *h*-prune coloured in red correspond to the polyproline sequence.



**Fig. 27 A)** Intensity ratios for all *c*-prune residues ( $I_0 - I / I_0 < 0.36$ ) extracted from 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC experiments performed in HEK-293 lysate overexpressing ASAP1. **B)** Model of interaction between ASAP1 and *h*-prune.

## Discussion

## DISCUSSION

The aim of this PhD thesis was to design an appropriate system to study a protein in its physiological conditions. This methodology becomes more intriguing when the selected molecule is c-prune, a domain belonging to the family of Intrinsically Disordered Domain (IDD), whose folding may be induced by its interaction with the partners. We tried to obtain a fusion protein able to introduce c-prune in the cell to perform in cell-NMR spectra. For this purpose we have studied the mechanism of internalization of this domain in eukaryotic HeLa cell line using different CPPs. In particular, we focused on a peptide derived from Herpes Virus Simplex type 1, gH625. The peculiar properties of this peptide, encompassing 625-644 residues of the glycoprotein H (gH) of Herpes simplex virus type 1<sup>[24]</sup>, render it an optimal candidate to act as a transporter for net acidic charged molecules by comparison with the positively charged TAT. Recent studies indicated also that gH625 cellular uptake is associated with its ability to interact with membrane lipids and to form a transient helical structure that temporarily affects membrane organization, thereby facilitating insertion into the membrane and translocation bypassing the endocytotic pathway<sup>[29]</sup>. Structural studies highlighted that when gH625 is in helical conformation, the polar residues concentrate on one face of the helix, giving it an amphiphilic character common to fusion peptides of most fusion glycoproteins of class I enveloped viruses. In addition gH625 has been used to efficiently deliver in cell quantum dots, that do not significantly traverse the membrane bilayer on their own<sup>[30]</sup>; liposomes<sup>[31]</sup>; dendrimers<sup>[14]</sup>; or nanoparticles through the blood-brain barrier<sup>[15]</sup>.

Therefore, gH625 can be considered a promising tool to efficiently perform intracellular delivery and may represent an alternative strategy to the use of the well characterized positively charged TAT peptide.

In this study, we report the validation of gH625 as carrier of c-prune. We designed, cloned, expressed and purified the recombinant gH625-c-prune as a soluble product. Far-UV CD spectra revealed an unexpected alfa/beta structure of the protein when compared with the CD spectra of gH625 or c-prune alone, both typical of unfolded species. In spite of this change in structure, gH625-c-prune retained its activity binding to its partner Nm23-H1 as verified by Western blot analyses. The quenching of tryptophan fluorescence by brominated phospholipid probed the membrane insertion of gH625-c-prune indicating that the protein bound to the peptide is located inside the membrane bilayer. This experiments may indicate that in the cell, gH625-prune may be able to cross the cell membrane. In addition flow cytometric analyses revealed a significant signal of fluorescence corresponding to the internalization of gH625-c-prune NBD-labeled, confirmed by the quenching of external fluorescence following dithionite treatment. Finally confocal microscopy data performed on gH625-c-prune showed that gH625 is able to deliver c-prune in cells . Data obtained with different inhibitors showed that the entrance in the cell is strongly dependent on endocytotic pathway although a fraction of the protein is able to cross the bilayer through a translocation mechanism, hypothesized to be the main crossing system used by the peptide gH625. It is worth mentioning that not only this is the first time that gH625 was shown to be able to deliver intracellularly a protein but this is also the first time that an intrinsically disordered protein was shown to be easily transported across the membrane bilayer retaining its biological activity.

In comparison with gH625-c-prune we started a characterization also of TAT-c-prune. The typical negative charge of a IDD<sup>[64]</sup>, such as c-prune, could represent a problem to obtain a fusion construct with TAT CPP. Indeed TAT-c-prune has been obtained as an unstable product in comparison with the stable gH625-c-prune achieving rather high levels of yield. Furthermore CD analyses of TAT-c-prune showed a spectrum typical of an unstructured

protein. In addition confocal microscopy data performed at the same concentration and incubation time for both proteins, gH625-c-prune and TAT-c-prune, revealed that only gH625 plays a role for c-prune delivery. We can hypothesize that the folding of gH625-c-prune, as monitored by CD, can represent an improvement and so facilitate the insertion into the membrane and subsequent translocation of the protein.

There is a natural abundance of intrinsically disordered proteins involved in the pathogenesis of numerous human diseases which may suggest that in the future these proteins may represent a huge reservoirs of novel drug targets. A key role is thus played by the availability of delivery tools for IDP. Altogether our results open a wider scenario on the potential applications of this membranotropic peptide fused with an IDP.

Due to the tendency to aggregate, it has not been possible to obtain gh625-c-prune at the concentration levels required to perform in cell-NMR experiments. Then, to overcome this problem, it was thought to use the NMR in lysates of eukaryotic cells in which is possible to use labeled recombinant c-prune. In this way it was feasible to study in detail the c-prune interactors in physiological environment.

So, in this PhD thesis we performed the NMR mapping of the c-prune functional sites, involved in the interactions with endogenous partners in human cell lysates. By the use of total extract of cells overexpressing proteins of interest, either naturally or transfected, we could map c-prune epitopes of key interactions with Nm23-H1, GSK-3beta and gelsolin, which are involved in relevant pathological cellular pathways. Furthermore, by site-directed mutagenesis we could identify residues involved in the molecular interactions whose mutation prevent the complex formation. Interestingly, the NMR analysis shows that the intrinsically disordered third domain of h-prune retains its conformational preferences in a cellular environment, at the same time showing a valuable capacity to recognize different cellular partners by means of appropriate protein portions. Overall, this

study shows that protein functions mediated by protein-protein interactions can be accurately followed by in cell lysates NMR fast experiments, which could be easily used for a very efficient NMR drug discovery strategy, either starting from fragments or developed ligands <sup>[65-68]</sup>. Finally, this functional characterization contributes to a better understanding of the h-prune biological activity and, moreover, by the means of NMR technique in cells lysates, it allows to study a protein under physiological conditions paving the way for drug discovery studies in intracellular environment.

## REFERENCES

- [1] Selenko P, Wagner G. *Looking into live cells with in-cell NMR spectroscopy*. J Struct Biol. 2007 May;158(2):244-53.
- [2] Serber Z, Keatinge-Clay AT, Ledwidge R, Kelly AE, Miller SM, Dötsch V.. *High-resolution macromolecular NMR spectroscopy inside living cells*. J Am Chem Soc. 2001 Mar 14;123(10):2446-7.
- [3] Reckel S, Löhr F, Dötsch V.. *In-cell NMR spectroscopy*. Chem biochem. 2005 Sep;6(9):1601-6.
- [4] Pervushin K, Riek R, Wider G, Wüthrich K.. *Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution*. Proc Natl Acad Sci U S A. 1997 Nov 11;94(23):12366-71.
- [5] Riek R, Wider G, Pervushin K, Wüthrich K. *Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules*. Proc Natl Acad Sci U S A. 1999 Apr 27;96(9):4918-23.
- [6] Riek R, Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K. *Solution NMR techniques for large molecular and supramolecular structures*. J Am Chem Soc. 2002 Oct 16;124(41):12144-53.
- [7] Hwang TL, van Zijl PC, Mori S.. *Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme*. J Biomol NMR. 1998 Feb;11(2):221-6.
- [8] Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Güntert P, Ito Y.. *Protein*

*structure determination in living cells by in-cell NMR spectroscopy*. Nature. 2009 Mar 5;458(7234):102-5

[9] Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y, Hiroaki H, Shirakawa M.. *High-resolution multi-dimensional NMR spectroscopy of proteins in human cells*. Nature. 2009 Mar 5;458(7234):106-9

[10] Liokatis S, Dose A, Schwarzer D, Selenko P. *Simultaneous detection of protein phosphorylation and acetylation by high-resolution NMR spectroscopy*. J Am Chem Soc. 2010 Oct 27;132(42):14704-5.

[11] Theillet FX, Liokatis S, Jost JO, Bekei B, Rose HM, Binolfi A, Schwarzer D, Selenko P. *Site-specific mapping and time-resolved monitoring of lysine methylation by high-resolution NMR spectroscopy*. J Am Chem Soc. 2012 May 9;134(18):7616-9.

[12] Selenko P, Serber Z, Gadea B, Ruderman J, Wagner G. *Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes*. Proc Natl Acad Sci U S A. 2006 Aug 8;103(32):11904-9.

[13] Takeuchi T, Kosuge M, Tadokoro A, Sugiura Y, Nishi M, Kawata M, Sakai N, Matile S, Futaki S. *Direct and rapid cytosolic delivery using cell-penetrating peptides mediated by pyrenebutyrate*. ACS Chem Biol. 2006 Jun 20;1(5):299-303.

[14] Carberry TP, Tarallo R, Falanga A, Finamore E, Galdiero M, Weck M, Galdiero S.. *Dendrimer functionalization with a membrane-interacting domain of herpes simplex virus type 1: towards intracellular delivery*. Chemistry. 2012 Oct 22;18(43):13678-85.

[15] Guarneri D, Falanga A, Muscetti O, Tarallo R, Fusco S, Galdiero M, Galdiero S, Netti PA. *Shuttle-Mediated Nanoparticle Delivery to the Blood-Brain Barrier*. Small. 2012 Nov 7.

[16] Zorko M, Langel U. *Cell-penetrating peptides: mechanism and kinetics of cargo delivery*. Adv Drug Deliv Rev. 2005 Feb 28;57(4):529-45.

- [17] Deshayes S, Morris MC, Divita G, Heitz F.. *Cell-penetrating peptides: tools for intracellular delivery of therapeutics*. Cell Mol Life Sci. 2005 Aug;62(16):1839-49.
- [18] Chang M, Chou JC, Lee HJ. *Cellular internalization of fluorescent proteins via arginine-rich intracellular delivery peptide in plant cells*. Plant Cell Physiol. 2005 Mar;46(3):482-8.
- [19] Suzuki T, Futaki S, Niwa M, Tanaka S, Ueda K, Sugiura Y.. *Possible existence of common internalization mechanisms among arginine-rich peptides*. J Biol Chem. 2002 Jan 25;277(4):2437-43.
- [20] Vivès E, Brodin P, Lebleu B. *A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus*. J Biol Chem. 1997 Jun 20;272(25):16010-7.
- [21] Schwarze SR, Dowdy SF. *In vivo protein transduction: intracellular delivery of biologically active proteins, compounds and DNA*. Trends Pharmacol Sci. 2000 Feb;21(2):45-8.
- [22] Duchardt F, Fotin-Mleczek M, Schwarz H, Fischer R, Brock R. *A comprehensive model for the cellular uptake of cationic cell-penetrating peptides*. Traffic. 2007 Jul;8(7):848-66.
- [23] Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF. *In vivo protein transduction: delivery of a biologically active protein into the mouse*. Science. 1999 Sep 3;285(5433):1569-72.
- [24] Galdiero S, Vitiello M, Falanga A, Cantisani M, Incoronato N, Galdiero M. *Intracellular delivery: exploiting viral membranotropic peptides*. Curr Drug Metab. 2012 Jan;13(1):93-104.

- [25] Galdiero S, Falanga A, Vitiello G, Vitiello M, Pedone C, D'Errico G, Galdiero M. *Role of membranotropic sequences from herpes simplex virus type 1 glycoproteins B and H in the fusion process*. *Biochim Biophys Acta*. 2010 Mar;1798(3):579-91.
- [26] Galdiero S, Falanga A, Vitiello M, Browne H, Pedone C, Galdiero M. *Fusogenic domains in herpes simplex virus type 1 glycoprotein H*. *J Biol Chem*. 2005 Aug 5;280(31):28632-43.
- [27] Galdiero S, Falanga A, Vitiello M, Raiola L, Russo L, Pedone C, Isernia C, Galdiero M. *The presence of a single N-terminal histidine residue enhances the fusogenic properties of a Membranotropic peptide derived from herpes simplex virus type 1 glycoprotein H*. *J Biol Chem*. 2010 May 28;285(22):17123-36.
- [28] Galdiero S, Falanga A, Vitiello M, Raiola L, Fattorusso R, Browne H, Pedone C, Isernia C, Galdiero M. *Analysis of a membrane interacting region of herpes simplex virus type 1 glycoprotein H*. *J Biol Chem*. 2008 Oct 31;283(44):29993-30009.
- [29] Galdiero S, Russo L, Falanga A, Cantisani M, Vitiello M, Fattorusso R, Malgieri G, Galdiero M, Isernia C. *Structure and orientation of the gH625-644 membrane interacting region of herpes simplex virus type 1 in a membrane mimetic system*. *Biochemistry*. 2012 Apr 10;51(14):3121-8.
- [30] Falanga A, Vitiello MT, Cantisani M, Tarallo R, Guarnieri D, Mignogna E, Netti P, Pedone C, Galdiero M, Galdiero S. *A peptide derived from herpes simplex virus type 1 glycoprotein H: membrane translocation and applications to the delivery of quantum dots*. *Nanomedicine*. 2011 Dec;7(6):925-34.
- [31] Tarallo R, Accardo A, Falanga A, Guarnieri D, Vitiello G, Netti P, D'Errico G, Morelli G, Galdiero S. *Clickable functionalization of liposomes with the gH625 peptide from Herpes simplex virus type 1 for intracellular drug delivery*. *Chemistry*. 2011 Nov 4;17(45):12659-68.

- [32] Okabe-Kado J, Kasukabe T, Honma Y, Hanada R, Nakagawara A, Kaneko Y. *Clinical significance of serum NM23-H1 protein in neuroblastoma*. *Cancer Sci.*2005 Oct;96(10):653-60.
- [33] Zollo M, Andrè A, Cossu A, Sini MC, D'Angelo A, Marino N, Budroni M, Tanda F, Arrigoni G, Palmieri G. *Overexpression of h-prune in breast cancer is correlated with advanced disease status*. *Clin Cancer Res.*2005 Jan 1;11(1):199-205.
- [34] Oue N, Yoshida K, Noguchi T, Sentani K, Kikuchi A, Yasui W. *Increased expression of h-prune is associated with tumor progression and poor survival in gastric cancer*. *Cancer Sci.*2007 Aug;98(8):1198-205.
- [35] Noguchi T, Oue N, Wada S, Sentani K, Sakamoto N, Kikuchi A, Yasui W. *h-Prune is an independent prognostic marker for survival in esophageal squamous cell carcinoma*. *Ann Surg Oncol.*2009 May;16(5):1390-6.
- [36] Massidda B, Sini M, Budroni M, Atzori F, Deidda M, Pusceddu V, Perra M, Sirigu P, Cossu A, Palomba G, Ionta M, Palmieri G. *Molecular alterations in key-regulator genes among patients with T4 breast carcinoma*. *BMC Cancer.*2010 Aug 24;10:458.
- [37] D'Angelo A, Garzia L, André A, Carotenuto P, Aglio V, Guardiola O, Arrigoni G, Cossu A, Palmieri G, Aravind L, Zollo M. *Prune cAMP phosphodiesterase binds nm23-H1 and promotes cancer metastasis*. *Cancer Cell.*2004 Feb;5(2):137-49.
- [38] D'Angelo A, Zollo M. *Unraveling genes and pathways influenced by H-prune PDE overexpression: a model to study cellular motility*. *Cell Cycle.*2004 Jun;3(6):758-61.
- [39] Virgilio A, Spano D, Esposito V, Di Dato V, Citarella G, Marino N, Maffia V, De Martino D, De Antonellis P, Galeone A, Zollo M. *Novel pyrimidopyrimidine derivatives for inhibition of cellular proliferation and motility induced by h-prune in breast cancer*. *Eur J Med Chem.*2012 Nov;57:41-50.

- [40] Tammenkoski M, Koivula K, Cusanelli E, Zollo M, Steegborn C, Baykov AA, Lahti R. *Human metastasis regulator protein h-prune is a short-chain exopolyphosphatase*. *Biochemistry*.2008 Sep 9;47(36):9707-13.
- [41] Carotenuto M, Pedone E, Diana D, de Antonellis P, Džeroski S, Marino N, Navas L, Di Dato V, Scoppettuolo MN, Cimmino F, Correale S, Pirone L, Monti SM, Bruder E, Zenko B, Slavkov I, Pastorino F, Ponzoni M, Schulte JH, Schramm A, Eggert A, Westermann F, Arrigoni G, Accordi B, Basso G, Saviano M, Fattorusso R, Zollo M. *Neuroblastoma tumorigenesis is regulated through the Nm23-H1/h-Prune c-terminal interaction*. *Sci Rep*. 2013 Mar 1;3:1351.
- [42] Kobayashi T, Hino S, Oue N, Asahara T, Zollo M, Yasui W, Kikuchi A. *Glycogen synthase kinase 3 and h-prune regulate cell migration by modulating focal adhesions*. *Mol Cell Biol*.2006 Feb;26(3):898-911.
- [43] Garzia L, Roma C, Tata N, Pagnozzi D, Pucci P, Zollo M. *H-prune-nm23-H1 protein complex and correlation to pathways in cancer metastasis*. *J Bioenerg Biomembr*.2006 Aug;38(3-4):205-13.
- [44] Garzia L, D'Angelo A, Amoresano A, Knauer SK, Cirulli C, Campanella C, Stauber RH, Steegborn C, Iolascon A, Zollo M. *Phosphorylation of nm23-H1 by CKI induces its complex formation with h-prune and promotes cell motility*. *Oncogene*.2008 Mar 20;27(13):1853-64.
- [45] Forus A, D'Angelo A, Henriksen J, Merla G, Maelandsmo GM, Flørenes VA, Olivieri S, Bjerkehagen B, Meza-Zepeda LA, del Vecchio Blanco F, Müller C, Sanvito F, Kononen J, Nesland JM, Fodstad Ø, Reymond A, Kallioniemi OP, Arrigoni G, Ballabio A, Myklebost O, Zollo M. *Amplification and overexpression of PRUNE in human sarcomas and breast carcinomas - a possible mechanism for altering the nm23-H1 activity*. *Oncogene*.2001 Oct 18;20(47):6881-90.

- [46] Reymond A, Volorio S, Merla G, Al-Maghteh M, Zuffardi O, Bulfone A, Ballabio A, Zollo M. *Evidence for interaction between human PRUNE and nm23-H1 NDPKinase.* Oncogene.1999 Dec 2;18(51):7244-52.
- [47] Müller T, Stein U, Poletti A, Garzia L, Rothley M, Plaumann D, Thiele W, Bauer M, Galasso A, Schlag P, Pankratz M, Zollo M, Sleeman JP. *ASAP1 promotes tumor cell motility and invasiveness, stimulates metastasis formation in vivo, and correlates with poor survival in colorectal cancer patients.* Oncogene.2010 Apr 22;29(16):2393-403.
- [48] Plyte SE, Hughes K, Nikolakaki E, Pulverer BJ, Woodgett JR. *Glycogen synthase kinase-3: functions in oncogenesis and development.*BiochimBiophysActa.1992 Dec 16;1114(2-3):147-62.
- [49] Woodgett JR. *A common denominator linking glycogen metabolism, nuclear oncogenes and development.* Trends Biochem Sci.1991 May;16(5):177-81.
- [50] Galasso A, Zollo M. *The Nm23-H1-h-Prune complex in cellular physiology: a 'tip of the iceberg' protein network perspective.* Mol Cell Biochem.2009 Sep;329(1-2):149-59.
- [51] Schafer DA, Cooper JA. *Control of actin assembly at filament ends.* Annu Rev Cell Dev Biol.1995;11:497-518.
- [52] Furman C, Short SM, Subramanian RR, Zetter BR, Roberts TM. *DEF-1/ASAP1 is a GTPase-activating protein (GAP) for ARF1 that enhances cell motility through a GAP-dependent mechanism.* J Biol Chem. 2002 Mar 8;277(10):7962-9
- [53] Buffart TE, Coffa J, Hermsen MA, Carvalho B, van der Sijp JR, Ylstra B, Pals G, Schouten JP, Meijer GA. *DNA copy number changes at 8q11-24 in metastasized colorectal cancer.* Cell Oncol.2005;27(1):57-65.

- [54] Ehlers JP, Worley L, Onken MD, Harbour JW. *DDEF1 is located in an amplified region of chromosome 8q and is overexpressed in uveal melanoma*. Clin Cancer Res.2005 May 15;11(10):3609-13.
- [55] Onodera Y, Hashimoto S, Hashimoto A, Morishige M, Mazaki Y, Yamada A, Ogawa E, Adachi M, Sakurai T, Manabe T, Wada H, Matsuura N, Sabe H. *Expression of AMAP1, an ArfGAP, provides novel targets to inhibit breast cancer invasive activities*. EMBO J.2005 Mar 9;24(5):963-73.
- [56] Lin D, Watahiki A, Bayani J, Zhang F, Liu L, Ling V, Sadar MD, English J, Fazli L, So A, Gout PW, Gleave M, Squire JA, Wang YZ. *ASAP1, a gene at 8q24, is associated with prostate cancer metastasis*. Cancer Res.2008 Jun 1;68(11):4352-9.
- [57] Dyson HJ, Wright PE. *Intrinsically unstructured proteins and their functions*. Nat Rev Mol Cell Biol.2005 Mar;6(3):197-208.
- [58] Zhou HX. *Effect of backbone cyclization on protein folding stability: chain entropies of both the unfolded and the folded states are restricted*. J Mol Biol.2003 Sep 5;332(1):257-64.
- [59] Tompa P, Csermely P. *The role of structural disorder in the function of RNA and protein chaperones*. FASEB J.2004 Aug;18(11):1169-75.
- [60] Hope MJ, Bally MB, Webb G, Cullis PR. *Production of large unilamellar vesicles by a rapid extrusion procedure: characterization of size distribution, trapped volume and ability to maintain a membrane potential*. Biochim Biophys Acta. 1985 Jan 10;812(1):55-65.
- [61] Bolen EJ, Holloway PW. *Quenching of tryptophan fluorescence by brominated phospholipid*. Biochemistry.1990 Oct 16;29(41):9638-43.
- [62] De Kroon AI, Soekarjo MW, De Gier J, De Kruijff B. *The role of charge and hydrophobicity in peptide-lipid interaction: a comparative study based on tryptophan*

*fluorescence measurements combined with the use of aqueous and hydrophobic quenchers*. *Biochemistry*. 1990 Sep 11;29(36):8229-40.

**[63]** Kim YI, Park S, Jeoung DI, Lee H. *Point mutations affecting the oligomeric structure of Nm23-H1 abrogates its inhibitory activity on colonization and invasion of prostate cancer cells*. *Biochem Biophys Res Commun*. 2003 Jul 25;307(2):281-9.

**[64]** Receveur-Bréchet V, Bourhis JM, Uversky VN, Canard B, Longhi S. *Assessing protein disorder and induced folding*. *Proteins*. 2006 Jan 1;62(1):24-45.

**[65]** Pellecchia M, Becattini B, Crowell KJ, Fattorusso R, Forino M, Fragai M, Jung D, Mustelin T, Tautz L. *NMR-based techniques in the hit identification and optimisation processes*. *Expert Opin Ther Targets*. 2004 Dec;8(6):597-611.

**[66]** Fattorusso R, Jung D, Crowell KJ, Forino M, Pellecchia M. *Discovery of a novel class of reversible non-peptide caspase inhibitors via a structure-based approach*. *J Med Chem*. 2005 Mar 10;48(5):1649-56.

**[67]** Fattorusso R, Frutos S, Sun X, Sucher NJ, Pellecchia M. *Traditional Chinese medicines with caspase-inhibitory activity*. *Phytomedicine*. 2006 Jan;13(1-2):16-22. Epub 2005 Aug 18.

**[68]** Pellecchia M, Bertini I, Cowburn D, Dalvit C, Giralt E, Jahnke W, James TL, Homans SW, Kessler H, Luchinat C, Meyer B, Oschkinat H, Peng J, Schwalbe H, Siegal G. *Perspectives on NMR in drug discovery: a technique comes of age*. *Nat Rev Drug Discov*. 2008 Sep;7(9):738-45. doi: 10.1038/nrd2606.