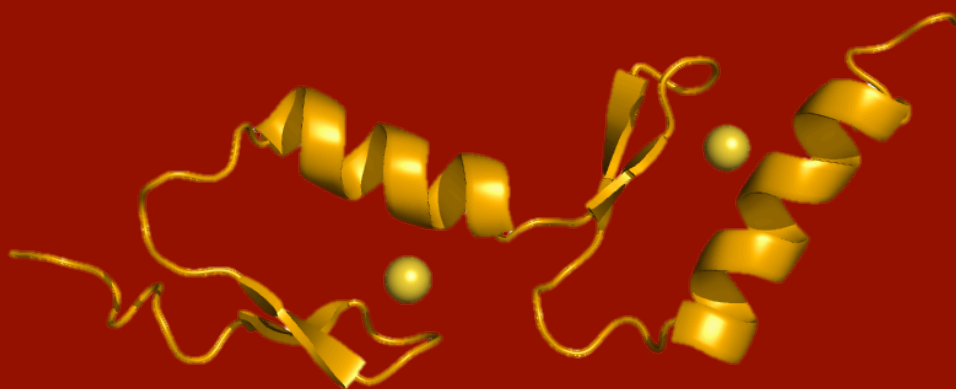


**UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA FÍSICA**



**VERSATILITY IN PROTEIN INTERACTIONS
INVOLVING MODULAR DOMAINS OF
ADAPTOR PROTEINS**



MARÍA DE LOS ÁNGELES CEREGIDO PÉREZ

**TESIS DOCTORAL
GRANADA
2013**

**UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA FÍSICA**



**VERSATILITY IN PROTEIN INTERACTIONS
INVOLVING MODULAR DOMAINS OF
ADAPTOR PROTEINS**

MARÍA DE LOS ÁNGELES CEREGIDO PÉREZ

**TESIS DOCTORAL
GRANADA
2013**

Editor: Editorial de la Universidad de Granada
Autor: María de los Ángeles Ceregido Pérez
D.L.: GR 491-2014
ISBN: 978-84-9028-797-2

**UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA FÍSICA**

**Memoria presentada para aspirar al título de Doctora en Química
(con la mención de “Doctora Europea”)**

Fdo. María de los Ángeles Ceregido Pérez

Licenciada en Química por la Universidad de Granada

Granada, a 24 de Junio de 2013

V^oB^o

DIRECTORES DE LA TESIS

Fdo: Nico A. J. van Nuland
VIB Group Leader BioNMR
Molecular Recognition Unit, VIB
Associate Professor
Structural Biology Brussels, VUB
Vrije Universiteit Brussel

Fdo: Ana I. Azuaga Fortes
Profesora Titular
Departamento Química Física
Facultad de Ciencias
Universidad de Granada

El doctorando María de los Ángeles Ceregido Pérez y los directores de la tesis Nico A.J. van Nuland y Ana I. Azuaga Fortes garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 24 de Junio de 2013

Director/es de la Tesis

Doctorando

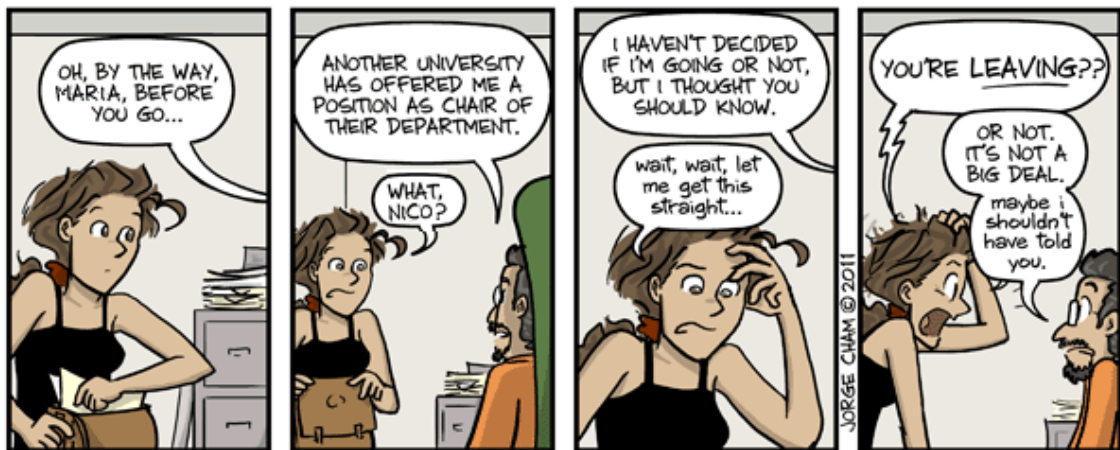
Fdo.:

Fdo.:

Nico A. J. van Nuland

Ana I. Azuaga Fortes

M. Ángeles Ceregido Pérez



Thanks for the opportunity

WWW.PHDCOMICS.COM

A mis padres
A Pako

AGRADECIMIENTOS

Esta tesis es el resultado de muchos momentos, tanto profesionales como personales; porque desde mi punto de vista, hacer una tesis no se limita a llevar a cabo un trabajo de investigación innovador, sino que forma parte de tu vida durante cuatro intensos años, y no sólo la haces tú, sino toda la gente que está a tu lado apoyándote en los momentos duros y alegrándose en los buenos, gente que te enriquece y que te hace más fuerte. Porque al fin y al cabo de eso se trata, de continuar aprendiendo. Y éste es el momento de daros las gracias, porque sin vosotros, nada habría sido posible.

En primer lugar, quiero agradecer a Pedro, al que le profeso una gran admiración, el haberme introducido en el mundo de la Química Física durante la carrera, es un privilegio haberle tenido como profesor, y por tanto, haber podido realizar mi tesis doctoral en su grupo de investigación, ha sido todo un honor.

A mis directores de tesis, Nico y Ana, por la confianza depositada en mí y la dedicación que habéis mostrado en formarme como científica durante estos años. Nico, tengo que agradecer muchas cosas, sobre todo que siempre has estado presente, ayudándome y guiándome, sin importar lo difíciles que fuesen las circunstancias; por abrirme las puertas de tu grupo de investigación en Bruselas, por enseñarme a mejorar cada día, por todos los conocimientos sobre RMN e interacción de proteínas que has compartido conmigo, y porque tu pasión por la investigación es contagiosa, gracias de corazón. A Ana, por introducirme en el mundo de la termodinámica y la biología molecular; por enseñarme a trabajar en el laboratorio de una manera rigurosa y crítica, muchas gracias.

También quiero expresar mi agradecimiento a Jero, porque aún sin ser mi director de tesis has colaborando en este proyecto de una manera muy activa, siempre dispuesto a ofrecer soluciones. Ha sido un placer poder contar contigo.

Por supuesto, quiero agradecer al Ministerio de Educación y Cultura de España, el haberme otorgado una beca FPU (Formación de Profesorado Universitario), gracias a la cual he podido realizar esta tesis doctoral, incluyendo las estancias en Bruselas.

I would like to express my gratitude to the members of my PhD tribunal Prof. Pedro L. Mateo Alarcón, Dr. Francisco Conejero Lara, Dr. Jerónimo Bravo, Dr. Malene Ringkjøbing Jensen, Dr. Abel García-Pino, Dr. Alex Volkov and Dr. Salvador Casares, for evaluating the work presented in this thesis. I also want to thank the external experts Dr. Sophie Vanwetswinkel and Dr. Rudi Beyaert for critically reading my thesis and all useful feedback.

A todos los profesores e investigadores del departamento de Química Física de la UGR: Quico, Obdulio, Enrique, Irene, Javier, Eva, Salva, Adela, Bertrand, José, Beatriz, Antonio, María del Mar... Gracias a todos por vuestras enseñanzas, apoyo y por estar siempre dispuestos a ayudarme en lo que hiciese falta.

A mis compañeros del laboratorio de Granada, por todos los buenos momentos compartidos. A Salva, por ser tan genuino y enseñarme a usar el ITC, pero sobre todo, por sus tartas de queso, que están buenísimas. A José Luis y Lorena, por vuestra amistad y por enseñarme todo lo que sabíais cuando entré en el laboratorio. A mis niñas, Sara y Ana Mari, gracias por todo, habéis sido un gran apoyo para mí y las mejores amigas que nunca pude imaginar. Sara, porque desde que te conozco sólo me has traído buenos momentos y alegría, y gracias a tí los cursos de doctorado fueron inolvidables, además, siempre me has ayudado en todo sin dudar, eres un sol. Ana Mari, por tu alegría y positivismo, porque con un abrazo eres capaz de recargarme de energía, y siempre he podido contar contigo. Y al resto de compañeros del labo: Javi, Carles, Crunchi, Fran, David, Álvaro, Inma, Roci, Ángel, Mari Carmen, Asun, ... Y por supuesto a José Miguel “el secre” y Araceli. Muchas gracias a todos por vuestra amistad y ánimos, sin vosotros nada habría sido igual.

I would like to thank the Molecular Recognition group (MoRe, SBB) of the Vrije Universiteit Brussel (VUB): Remy, Alex, Sophie, Abel, Wim, Lieven, Radu, Lucía, Yann, Karen, Anneleen, Ellen, Sarah, Valentina, San, Ariel, Steven, Alexandra... Thanks a lot for making me feel at home since the very first day that I entered in the lab! It has been very rewarding to work with you, thanks for your support and for all the wonderful moments!! Specially to the people I have more closely collaborated with: Alex, it has been an honor to work with you; you are an example of excellence, thanks for all the good advices, for helping me so much during this thesis and for introducing me in the “paramagnetic world”; Sophie, I am very grateful to you for all your support, for the fact that you took the time to read the entire thesis and for all the ideas and feedback; Abel, muchas gracias por todo lo que me has ayudado desde que entré al laboratorio, por los viajes al sincrotrón, por tu positivismo y porque tu amplia experiencia en el laboratorio y tu gran saber hacer, han sido muy importantes en el desarrollo de esta tesis; Lieven, thanks a lot for all your help in the process of NMR structure calculation; Karen, you are very sweet, thanks a lot for sharing with me your experience in paramagnetic labeling; Yann, it has been a pleasure to work with you in the lab, you are full of motivation and good ideas, thanks a lot for helping me with the sequence alignments and for making my tree!! And of course, my MARALU team, Radu and Lucía, you are great!! It has been marvelous to work with you and to share such a nice time together (and also the pipettes, the office, the bench, the lunch, the hug moments, the nitrogen refilling...), I will never forget you.

My friends in Brussels: Radu & Roxana (and recently Irene), Gabi & Xavier (and recently Iris), Lucía & Lennert, Sara & Luis, Ana & Santi and Cesyen & Natali. Thanks to all of you for always being there and for supporting me, because every time I spent with you, you made me feel at home. For all the incredible, funny moments that we have shared, THANKS!!

A mis padres, Ángeles y Eugenio, no habría palabras para describir todo lo que os tengo que agradecer, y si hoy estoy aquí es gracias a vosotros, a vuestro amor incondicional, porque siempre me habéis ayudado a conseguir mis sueños, poniendo todo y más de vuestra parte; porque cuando caía me levantasteis, y porque cuando volé no me cortasteis las alas. Os quiero mucho, y el mayor orgullo que tengo en la vida es ser vuestra hija. A mi hermano, el mejor periodista del mundo, un artista de los pies a la cabeza, que siempre me ha apoyado. A mis abuelos, que siempre han creído en mí, y me han dado tanta suerte!! A mis tíos, por alegrarse con cada logro, por sus visitas a Bruselas, y por quererme y enseñarme tanto desde chiquitita, sois una parte indispensable de mi vida. Isa esto va por ti, y por la promesa que te hice un día.

Y por supuesto a Pako, por compartir conmigo cada paso del camino sin importar la distancia o el tiempo, por brindarme su amor, comprensión y paciencia, pero sobre todo, por hacerme sentir la persona más feliz del mundo, animándome y aconsejándome siempre que lo necesito. Te quiero. Gracias también a la familia de Pako por su cariño y apoyo.

Finalmente, a todas aquellas personas que han participado en esta aventura y que no he nombrado anteriormente, gracias, gracias a todos.

CONTENTS

Chapter I. INTRODUCTION	3
1.1 Signaling pathways inside the Cell	3
1.1.1 The importance of protein-protein interactions	4
1.2 Adaptor proteins	5
1.2.1 Dynamic control of signaling by adaptor proteins	5
1.2.2 Modular domains	6
1.3 CIN85/CD2AP family of adaptor proteins and natural targets	7
1.4 SH3 domains	11
1.5 Ubiquitin and ubiquitin-binding domains	13
1.6 TAX1BP1 a novel ubiquitin-binding adaptor protein	17
1.6.1 Function of TAX1BP1	17
1.6.2 Structure of TAX1BP1	21
1.7 Zinc finger domains	23
1.7.1 Classical ZnF	25
Chapter II. AIMS OF THE THESIS	29
Chapter III. MATERIALS AND METHODS	33
3.1 Protein expression and purification	33
3.1.1 Transformation and Overexpression in <i>E. coli</i> cells	33
3.1.2 Purification protocol of the ¹⁵ N- ¹³ C isotopically labeled SH3A domain of CIN85	35
3.1.3 Purification protocol of ¹⁵ N- ¹³ C isotopically labeled UBZ1+2 domains of TAX1BP1	36
3.2 Isothermal Titration Calorimetry (ITC)	37
3.2.1 The ITC experiment: theory and practice	38
3.2.2 Mathematical analysis of the ITC experiment	41
3.3 Nuclear Magnetic Resonance (NMR)	44
3.3.1 Introduction to NMR	44
3.3.2 Multidimensional NMR spectroscopy	46
3.3.2.1 The ¹ H- ¹⁵ N and ¹ H- ¹³ C correlation spectra	47
3.3.2.2 Chemical exchange by NMR	48
3.3.2.3 Cross-relaxation and the nuclear Overhauser effect	50
3.3.3 Chemical shift assignment and structure calculation from NMR data	50
3.3.3.1 Sequential assignment using heteronuclear scalar coupling	50
3.3.3.2 Structure calculation from NMR data	51
3.3.3.3 Chemical shifts to map protein-protein interactions	52
3.3.4 Paramagnetic Relaxation Enhancement (PRE)	54
3.3.4.1 PRE experiments	55
3.3.4.2 Ensemble refinement against intermolecular PREs	55
3.3.5 Using relaxation to study ¹⁵ N backbone dynamics	56
3.4 Small Angle X-ray Scattering (SAXS)	59
3.5 Circular Dichroism (CD)	61
3.6 Bio-Layer Interferometry (BLI)	62
3.6.1 Experimental setups	62
3.6.2 Data analysis	64

Chapter IV. MOLECULAR BASIS OF THE VERSATILE INTERACTIONS OF SH3A DOMAINS OF CD2AP AND CIN85 WITH THE ATYPICAL PROLINE-RICH SEQUENCES PRESENT IN CD2 AND Cbl-b 67

4.1	Introduction.....	67
4.2	Experimental procedures	68
4.2.1	Purification of ¹⁵ N- ¹³ C-labeled SH3A domain of CIN85.....	68
4.2.2	Atypical proline-rich peptides	71
4.2.3	Isothermal Titration Calorimetry.....	71
4.2.4	Nuclear Magnetic Resonance Spectroscopy	72
4.2.5	Small Angle X-Ray Scattering.....	73
4.3	Results and discussion.....	74
4.3.1	Thermodynamic characterization of the interaction of CD2AP-SH3A with CD2	74
4.3.2	NMR titration of CD2AP-SH3A with CD2.....	78
4.3.3	Characterization of the interaction of CD2AP-SH3A with Cbl-b and mutant peptides using ITC and NMR.....	81
4.3.4	Characterization of the interaction of CIN85-SH3A with CD2 by ITC and NMR	87
4.3.5	Characterization of the interaction of CIN85-SH3A with Cbl-b and mutant peptides by ITC and NMR.....	89
4.3.6	Structural characterization of the interaction of CD2AP-SH3A and CIN85-SH3A with Cbl-b and CD2 using Small Angle X-ray Scattering (SAXS).....	98
4.4	Concluding remarks	102
4.4.1	Atypical polyproline recognition: the role of arginine residues for type I and type II binding	102
4.4.2	SH3A domains of CD2AP and CIN85 have different decoding mechanisms: biological implications.....	104

Chapter V. UBZ DOMAINS OF TAX1BP1 109

5.1	Introduction.....	109
5.2	Experimental Procedures	110
5.2.1	Purification of ¹⁵ N- ¹³ C-labeled UBZ1+2 of TAX1BP1.....	110
5.2.2	Purification of ubiquitin and its single-cysteine mutants	113
5.2.2.1	Introduction of a paramagnetic probe in ubiquitin	113
5.2.3	Isothermal Titration Calorimetry.....	115
5.2.4	Nuclear Magnetic Resonance Spectroscopy	115
5.2.5	Paramagnetic Relaxation Enhancement.....	116
5.2.6	Small Angle X-ray Scattering	116
5.2.7	Bio-Layer Interferometry	117
5.3	Results and discussion.....	117
5.3.1	Solution NMR structure of the free UBZ1+2.....	117
5.3.1.1	Sequential assignment and structure calculations.....	117
5.3.1.2	Backbone dynamics of free UBZ1+2.....	120
5.3.1.3	Comparison of the NMR and X-ray structures of the UBZ1+2 tandem.....	123
5.3.1.4	Coordination of the Zn ²⁺ atom and the hydrophobic core of the UBZ domains.....	125
5.3.2	Thermodynamic characterization of the interaction between the UBZ1+2 and ubiquitin	128
5.3.3	Mapping the interaction of UBZ1+2 with ubiquitin by NMR.....	131
5.3.3.1	Titration of ¹⁵ N-labeled ubiquitin with isolated UBZ1 and UBZ2 domains and the UBZ1+2 tandem	131

5.3.3.2	Determining the ubiquitin-binding interface on UBZ1+2 by chemical shift mapping	136
5.3.4	Backbone dynamics of UBZ1+2 tandem in complex with ubiquitin	150
5.3.5	Paramagnetic tagging for structural determination of the UBZ1+2-ubiquitin complex	152
5.3.5.1	PREs in the complex of ¹⁵ N UBZ1+2 with ubiquitin	153
5.3.5.2	Structure calculations of the UBZ1+2-ubiquitin complex	154
5.3.5.3	Experimental controls	157
5.3.6	Small Angle X-ray analysis of the free and bound UBZ1+2	158
5.3.7	Interaction of UBZ1+2 with Lys48-, Lys63- and linear polyubiquitin chains	159
5.4	Concluding remarks	162
5.4.1	The role of the Zn ²⁺ atom in the structure of the UBZ domains	162
5.4.2	UBZ1 and UBZ2 domains of TAX1BP1 bind to ubiquitin	164
5.4.2.1	UBZ1+2 recognition area	164
5.4.2.2	Ubiquitin recognition area	168
5.4.3	Electrostatic and hydrophobic forces involved in binding of ubiquitin to UBZ1+2 of TAX1BP1	169
5.4.4	Promiscuous interactions of UBZ1+2 with polyubiquitin chains	170
Chapter VI. GENERAL DISCUSSION.....		175
Chapter VII. RESUMEN Y CONCLUSIONES		191
7.1	Resumen	191
7.2	Conclusiones	193
7.3	Summary	196
7.4	Conclusions	198
ABBREVIATIONS.....		203
REFERENCES		208

CHAPTER I

INTRODUCTION

Learn from yesterday, live for today, hope for tomorrow. The important thing is to not stop questioning. Albert Einstein

Chapter I. INTRODUCTION

1.1 Signaling pathways inside the Cell

The cell is the smallest autonomous unit of life, and represents the midpoint between the molecular and macroscopic scales. Like in the macroscopic world, the flow of information and energy is essential for the proper functioning of the system. A cell in the body is continually exposed to many different stimuli, which it must integrate into a coherent response. Individual cellular processes, such as signaling, membrane trafficking, cytoskeletal dynamics, gene expression or cell division are coordinated events, responsible of the proper functioning of all living organisms. The control of cell homeostasis, morphology and growth, is vital for a correct function and development of tissues in multicellular organisms, and depends on a very stringent regulation of responses to extracellular signals modulating the level of cell activation. Failures in these mechanisms can lead to pathological or disease states, like uncontrolled cell growth or tissue invasion, characteristic features of cancer.

To understand signal transduction in a general sense, it is important to know whether different biochemical pathways use related molecular procedures to control cellular behavior. To identify specificity in signaling, we need to know which receptors interact with particular targets and how the proteins of one pathway can be separated from related signaling components. However, since the entire cell should function as a single unit whose different elements respond in an organized fashion to external signals, it is also crucial to understand the communication between distinct signaling pathways.

The control of positive and negative signals ensures that proper signal thresholds are reached and maintained for the appropriate length of time. Mapping and understanding molecular interaction networks and signaling pathways is the first step toward modeling how a cell operates in time and space. The next challenge is to incorporate structural information into this picture in order to fully understand the cell behavior.

Most of the essential structure and function of cells is mediated by proteins. And proteins are dynamic! These macromolecules exhibit a remarkable versatility allowing them to perform a large range of activities that are fundamental for the understanding of life and for the cure of disease. Protein dynamics (particularly at the domain level) is a controlling influence in the allosteric formation of protein complexes, as well as in catalysis, cell signaling and regulation, metabolic transport, and cellular locomotion.

To better understand the organization and regulation of these intracellular signaling pathways in normal condition and diseases, one has to look at the molecular interaction governing the interplay between these proteins.

1.1.1 The importance of protein-protein interactions

Membrane proteins cannot be redistributed in the cytosol, and it is therefore necessary to have a functional connection between external stimuli at the cell surface (such as growth factors, cytokines and hormones) and receptor internalization or downstream signaling. Nature has evolved as a clever cellular machinery consisting of millions of proteins acting in concert (Fig. I.1). Cellular networks require protein complexes to assemble and disassemble in a very dynamic way based on protein-protein interaction (Levy E.D. and Pereira-Leal J.B., 2008). Protein associations need precise spatial and temporal regulation. Just single mutations or environmental factors can interfere with protein-protein interactions, leading to pathology (Bader S. et al., 2008).

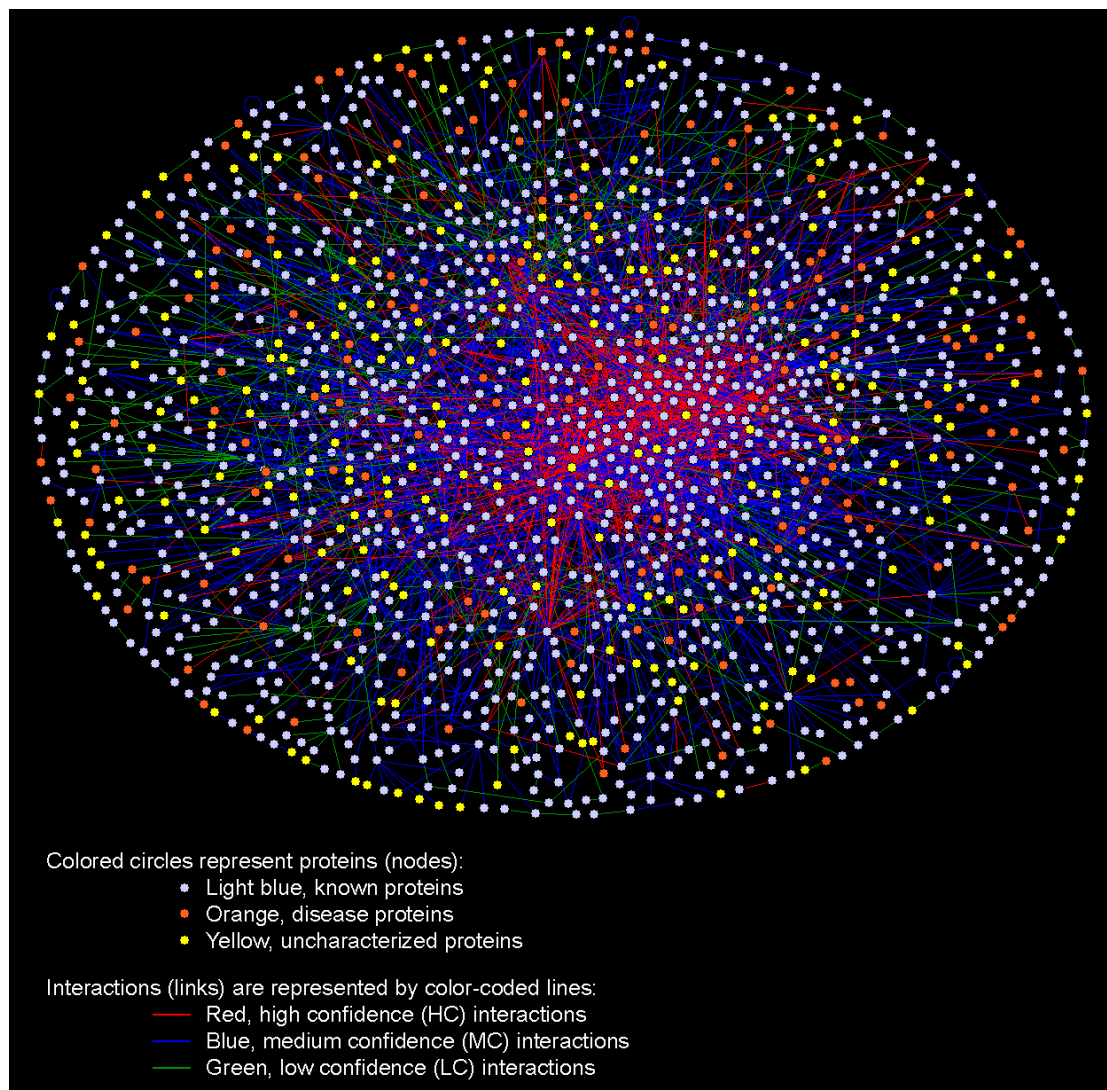


Figure I.1. Human protein-protein interaction network. Figure taken from MDC (MAX-DELBRÜCK-CENTRUM FÜR MOLEKULARE MEDIZIN, BERLIN-BUCH) website (www.mdc-berlin.eu/de/highlights/archive/2005/highlight11/index.html); reproduced from Stelzl U. et al., 2005.

The different kinds of protein-protein interactions can be classified depending on the type of protein chains involved, temporal stability of each component and lifetime of the complex (Nooren I.M and Thornton J.M., 2003; Nooren I.M., 2003). These

interactions can take place between identical, homo-oligomeric chains, or non-identical hetero-oligomeric chains. Protein complexes can also be classified into obligate and non-obligate depending if each interacting partner is stable on its own or not. Indeed, all interactions are driven by the free energy of binding and the concentration of the components. Consequently, protein-protein interactions can be regulated by influencing the binding affinity or local concentration of each component. There are three types of control taking place *in vivo*: (a) Encounter of two proteins, for which colocalization is required for an efficient contact between the interacting surfaces. (b) Control of the local concentration of each component. (c) Modification of the local physicochemical environment, for example by the presence of effectors molecules, change of the pH or concentration of ions.

Based on the interaction of the known three-dimensional (3D) structures stored in the Protein Data Bank (PDB) (Berman H.M. et al., 2000), protein interaction which involves one or more globular domains, can be divided in two main categories, based on their contact interfaces: domain-domain and domain-peptide interactions (Aloy P. and Russel R.B., 2006; Stein A. et al., 2011). In domain-peptide interactions a globular domain in one protein recognizes a short linear motif from another protein, creating a relatively small interface. In cell signaling, important protein interactions are frequently mediated by short, unstructured sequences, which specifically interact with peptide binding domains (Pawson T. and Scott J.D., 1997), and, due to their transient nature, are much more difficult to handle biochemically (Pawson T., 2003).

1.2 Adaptor proteins

One of the main objectives of the study of signal transduction pathways is to determine the mechanisms controlling the cross-talk between signaling cascades, and signaling specificity. Adaptor (or adapter) proteins are the major contributors to these processes.

Adaptor proteins are polypeptides with one or more domains able to bind other proteic or non-proteic ligands, and in general, without enzymatic activity. They have the capability to connect proteins to either other proteins, or to the plasma membrane or to intracellular organelles. This is how large signaling multimeric complexes are formed, localizing signaling molecules to specific subcellular localizations, and hence contributing to the specificity and efficiency of cellular responses in a spatio-temporal manner (Pawson T., 2002, Csiszár A., 2006). This allows the control of important cellular aspects, like migration, proliferation, differentiation and death.

1.2.1 Dynamic control of signaling by adaptor proteins

One of the most important characteristics of adaptor proteins is their multi-tasking capacity. A same adaptor protein can act at different levels in different signaling cascades, which means saving the metabolic cost that their delivery would suppose to the cell, but also allows dynamic signaling networks and integration of independent signaling events to happen through a limited number of unique proteins. For example, the Crk adaptor protein has a key role in a vast number of cellular processes involved in the relocalization and activation of several effector proteins, which include guanine nucleotide releasing proteins, protein kinases and small GTPases (Feller S.M., 2001).

In principle, such adaptors have a domain that selectively recognizes an activated receptor, and one or more domains that recruit cytoplasmic effectors, as in the case of the SH2/SH3 domains present in the adaptor proteins Crk and Grb2 (Feller S.M., 2001). The binding activity of adaptors can be regulated by conformational reorganization, and by the cooperative association of domains within the same adaptor. Furthermore, an individual adaptor can deliver different outputs by utilizing distinct combinations of binding partners in different cell types, or even at different locations in the same cell. Adaptors can also control the oligomerization of receptor signaling complexes and subcellular location and duration of signaling events, and act as coincidence detectors to enhance specificity in cellular responses.

Some results have also suggested that the binding properties of adaptors can themselves be dynamically regulated, for example by intramolecular interactions. Indeed the trafficking of signaling proteins is itself dependent on adaptor proteins, such as those associated with the endocytic machinery and protein sorting in endosomes (Slagsvold T. et al., 2006; Raiborg C., 2006; Teo H. et al., 2006).

For this reason, the study of adaptor proteins is key to understand why different kinds of cells behave differently in response to various external signals. A remarkable conclusion from studying physiological signaling complexes (Dueber J.E. et al., 2003; Bhattacharyya R.P. et al., 2006) is that adaptor proteins can confer sophisticated cooperative and gating properties, which are important for complex cellular behavior.

1.2.2 Modular domains

One of the main characteristics of adaptor proteins is that they are composed of domains and motifs that mediate molecular interactions, and can thereby link signaling proteins such as activated cell-surface receptors to downstream effectors. These modular domains (such as SH2 and SH3 domains, UBZ domains, WW domains, PHD domains, etc.) play a central role in the assembly of multiprotein complexes, and their specificity and affinity for a particular target (or targets) often determine the efficiency of cellular signaling (Nishida M. et al., 2001).

Many of these domains can be found in the human genome. Generally, modular domains are small (between 30 and 150 amino acids) motifs, folded into compact and stable structures, and use to appear in tandem of two or more modular domains which often present similar affinities and specificities for their natural targets. In general, each family of domains recognizes a specific and highly conserved sequence (no more than 3 or 6 amino acids) in the target protein. For example, SH2 and PTB domains recognize sequences with phosphorylated tyrosines; WW, SH3 and EVH domains recognize proline-rich sequences (Sudol M. et al., 1996; Niebuhr K. et al., 1997; Nguyen J.T. et al., 1998; Aghazadeh B. and Rosen M.K., 1999); PDZ domains bind short peptide motifs at the carboxy-termini of proteins, typically transmembrane receptors (Songyang Z. et al., 1997; Harris B.Z and Lim W.A., 2001); and EH domains recognize Asn-Pro-Phe sequences commonly found in polypeptides involved in protein trafficking (Salcini A.E. et al., 1997; Mayer B.J., 1999). These sequences are flanked by additional fragments that interact with variable elements in the binding site and seem to determine the binding specificity within the same family of domains.

Despite the central importance of motif-dependent interactions in cell signaling, there are few methods to identify peptide-binding partners. Phage display (Kay B.K. et al., 2001), array technologies (Espejo A. et al., 2002), and peptide libraries (Elia A.E. et al., 2003) are powerful strategies to learn about general features of optimal peptide sequences binding to a given domain or protein.

In many cases, cytoplasmatic signaling proteins possess multiple protein-protein and protein-phospholipid interaction domains, covalently linked in various combinations. The joining of different domains can serve a variety of functions. Two domains may interact with different sites on the same target, as commonly occurs with polypeptides that possess tandem SH2 domains, thereby increasing both the affinity and specificity of the interaction (Ottinger E.A. et al., 1998). Conversely, separate domains may interact with distinct partners, as observed for adaptors with SH2 and SH3 domains, such as Grb2 that links activated receptors to targets with proline-rich motifs, notably the Ras GEF Sos1 (Rozakis-Adcock et al., 1993).

In general, the interactions mediated by these domains imply small surfaces in comparison to other protein-protein interactions. This property favors the molecular design of compounds blocking protein-protein binding surfaces with high specificity (Kardinal C. et al., 1999; Kardinal C. et al., 2000; Lee K.Y. et al., 2002), which makes these domains attractive targets for the design of new-generation drugs.

1.3 CIN85/CD2AP family of adaptor proteins and natural targets

CIN85/CD2AP is a family of adaptor proteins with specific functions in higher eukaryotes. This family of ubiquitously expressed adaptors is involved in a wide variety of cellular processes, like cytoskeletal arrangement (Lynch D.K. et al., 2003; Bruck S. et al., 2006; Hernandez-Valladares M. et al., 2010; van Duijn T.J. et al., 2010), cell signaling (Gout I. et al., 2000; Huber T.B. et al., 2003; Wakasaki T. et al., 2010), degradative trafficking and endocytosis of receptors (Soubeyran P. et al., 2002; Kowanetz K. et al., 2003; Kobayashi S. et al., 2004; Havrylov S. et al., 2010; Schroeder B. et al., 2010; Tossidou I. et al., 2010; Yoon H.Y. et al., 2011). They carry out a crucial role in the formation of kidney glomeruli architecture, organization of specialized functions in T-cells (immunological synapse), apoptosis in neuronal cells (Dikic I., 2002), and can enhance the degradation of receptor tyrosine kinase (RTK) (Dikic I., 2002; Marois L. et al., 2011).

CD2AP (CD2 Associated Protein) is a cytosolic protein of 641 amino acids that belongs to the CIN85/CMS adaptor proteins family, being the mouse homologue of the human CMS (Cas ligand with Multiple SH3 domains). It was first found in T cells where CD2AP has been shown to be essential for CD2 (Cluster of differentiation 2) clustering and for inducing T cell polarization at the contact area between T cells and antigen-presenting cells (Dustin M.L. et al., 1998).

CD2 is a 50 kDa transmembrane glycoprotein that belongs to the immunoglobulin (antibody) family. CD2 presents a proline-rich cytosolic domain (well conserved in humans, rats and mice), which functions like a signaling domain. CD2 is expressed on

most human T-cells and natural killer (NK) cells (but not in B-cells). It plays an important role in mediating cell adhesion in both T-lymphocytes and in signal transduction. The immunologic role of CD2 includes recognition involving T-helper and antigen-presenting cells, cytolytic effector functions of natural killer cells and cytotoxic T-lymphocytes, and it is also related to the interaction of thymic epithelial cells and thymocytes (Singer K.H. and Haynes B.F., 1987).

Binding of the first SH3 domain of CD2AP (SH3A) to the proline-rich sequence at the C-terminus of CD2 is induced by T-cell activation and is required for CD2 clustering and T-cell polarization (Dustin M.L. et al., 1998). Both of these events are critical for the formation of an effective T-cell antigen-presenting cell function. But CD2AP plays other roles beside the immune system, for example, mice born with defects or deficient in CD2AP, develop a congenital nephrotic syndrome, dying six weeks after birth, suggesting that CD2AP is also involved in maintaining the integrity of the kidney glomerulus (Grunkemeyer J.A. et al., 2005). And in agreement with the multitask definition of adaptor proteins, CD2AP is also involved in the reorganization of actin in several cellular processes, including cytotogenesis (Dustin M.L. et al., 1998; Schafer D.A. et al., 2000; Welsch T. et al., 2001; Monzo P. et al., 2005).

The adaptor protein CIN85 (Cbl-interacting protein of 85 kDa) also known as Ruk (Regulator of ubiquitous kinase) and SETA (SH3-domain containing gene Expressed in Tumorigenic Astrocytes) in rat, and SH3KBP1 (SH3-domain Kinase Binding Protein 1) in mouse, is a multifunctional adaptor protein from the same family as CD2AP (Dikic I., 2002; Bian M. et al., 2008). Due to differential promoter and alternative splicing, CIN85 has multiple transcripts in cells, different to CD2AP, which is solely expressed in its full-length form. Some of the transcripts have the patterns of tissue-specific in developmentally regulated expression (Buchman V.L. et al., 2002; Finniss S. et al., 2004). CIN85 is implicated in Cbl-mediated down-regulation of receptor tyrosine kinase (RTK) (Kowanetz K. et al., 2003), and in immune cells, recent studies have shown that CIN85 is constitutively associated with c-Cbl, Cbl-b and BLNK (B-cell linker) in human B cells (Niuro H. et al., 2012), suggesting that CIN85 is required for Cbl-mediated regulation of BCR (B-cell receptor) signaling and for downstream events like survival, growth and B-cell differentiation. Recently, it has been shown that CIN85 links the B-cell antigen receptor to the activation of the canonical nuclear transcription factor (NF- κ B) pathway (Kometani K. et al., 2011).

Cbl (Casitas B-lineage lymphoma) is an E3 ubiquitin-protein ligase, which accepts ubiquitin from specific E2 ubiquitin-conjugating enzymes, and transfers it to substrates, generally promoting their degradation by the proteasome. However, Cbl can also promote receptor endocytosis via a pathway that is functionally separable from its ubiquitin ligase activity, and depend on the Cbl interactions with the family of adaptor proteins CIN85/CD2AP (Petrelli A. et al., 2002; Soubeyran P. et al., 2002). There are three described mammalian isoforms of Cbl: Cbl-3 (or Cbl-c), c-Cbl and Cbl-b (Fig. I.2). The first one is the shortest, lacking the C-terminal UBA/LZ domain and the CIN85/CD2AP binding region, and is mainly expressed in epithelial cells. The other two are ubiquitously expressed, and hold a high degree of sequence identity (48%) (Schmidt M.H. and Dikic I., 2005).

CIN85 associates with two members of the Cbl family, c-Cbl and Cbl-b, but not Cbl-3, via its SH3 domains. This interaction depends on the presence of a minimal binding domain at their distal carboxyl-terminal tails (Szymkiewicz I. et al., 2002), known as an atypical proline-rich motif (PxxxPR) (Fig. I.2) that can also be recognized by CD2AP (or CMS) (Feshchenko E.A. et al., 1998). The CIN85 association with Cbl is enhanced by growth factor stimulation in mammalian cells (Watanabe S. et al., 2000; Soubeyran P. et al., 2002), and initial studies have proposed that ligand-induced tyrosine phosphorylation of Cbl promotes a conformational change in its carboxy-terminus, exposing thereby the high affinity peptide to CIN85/CD2AP binding (Watanabe S. et al., 2000; Kirsch K.H. et al., 2001; Soubeyran P. et al., 2002; Szymkiewicz I. et al., 2002).

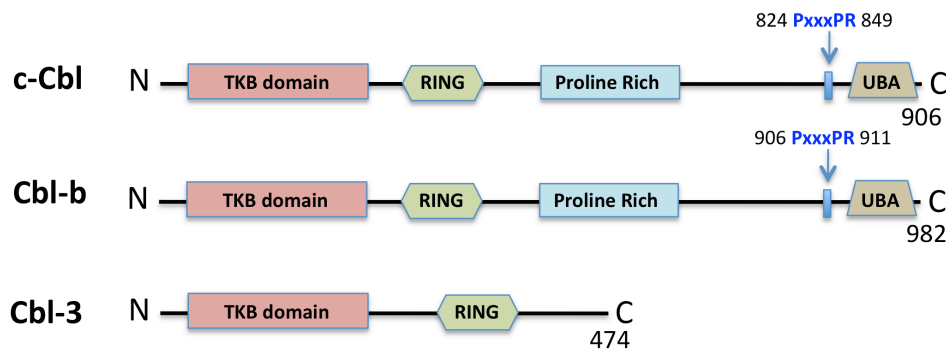


Figure I.2. Schematic representation of the domain structure of the three members of the mammalian Cbl family. The N-terminal tyrosine-kinase-binding (TKB) domain can bind to phosphorylated tyrosines on tyrosine kinases. The RING-finger interacts with ubiquitin-conjugated enzymes. The proline-rich region recruits Src-homology 3 (SH3)-binding proteins. In the C-terminus of c-Cbl and Cbl-b there is a PxxxPR motif that binds the SH3 domain of CIN85/CD2AP family of adaptor proteins (residues 824-849 in c-Cbl and 906-911 in Cbl-b). Cbl-b and c-Cbl have an ubiquitin-associated domain (UBA) that is lacking in the short Cbl-3.

This atypical polyproline motif is preferentially recognized by the SH3 domains of CIN85 and CD2AP (CMS) but not by other SH3 domain-containing proteins like Grb2, Src, Crk, Abl, CAP, or ArgBP2 (Kowanetz K. et al., 2003). Similarly, the amino-terminal SH3 domain of CIN85 has been shown to bind the cytosolic atypical proline-rich PxxxPR motif located in the CD2 receptor (Hutchings N.J. et al., 2003), natural target of CD2AP. Furthermore, interactions between CIN85/CD2AP and other trafficking proteins including AP-2, Dab2, Rab4, PAK2, ALIX, and ESCRT-1 have been described (Chen B. et al., 2000; Brett T.J et al., 2002; Cormont M. et al., 2003; Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003; Schmidt M.H. et al., 2003; Usami Y. et al., 2007).

In general, CIN85/CD2AP family of adaptor proteins, selectively controls the spatial and temporal assembly of multiprotein complexes that transmit intracellular signals in processes like cell growth regulation, differentiation, migration and survival (Chen B. et al., 2000). Both CIN85 and CD2AP are ubiquitously expressed, although with slight, but important differences in their expression pattern. CIN85 is expressed in most of the tissues, including the brain (Bogler O., 2000; Tibaldi E.V. et al., 2003), with high expression levels in skeletal muscle, and lowest in lung and pancreas (Watanabe O. et al., 2005). On the other side, CD2AP (CMS) is also broadly expressed in all tissues except the brain; it is expressed at high levels in the liver, spleen, pancreas, placenta,

colon, kidney and thymus, and at low levels in skeletal muscle, aorta, bladder and uterus (Kirsh K.H. et al., 1999). This might suggest an overlapping role of these proteins with a specialized location for each of them.

CIN85 and CMS/CD2AP share 39% identity and 54% similarity on amino acid sequence (Dikic I., 2002) and their overall domain organization is very similar (Fig. I.3). From the structural point of view, there are three Src homology 3 (SH3) in the N-terminal part (SH3A, SH3B and SH3C) known to recognize proline-rich sequences, a central proline-rich region providing potential recognition sites for other SH3 domain-containing proteins, and a Coiled-coil domain at the C-terminal through which they are capable to homo- and heterodimerize and hence promote clustering (Fig. I.3). A region rich in serine and threonine residues, which could be subject to phosphorylation, lies between the second and third SH3 domains. There are also three FxDxF sequences in the N-terminus that may serve as binding sites for the clathrin adaptor protein AP2 (Brett T.J. et al., 2002). The main difference in CD2AP sequence, compared to CIN85, is the presence of four putative actin binding sites in its carboxyl-terminus that are similar to the LKKTET motifs found in a number of actin binding proteins (Krisch K.H. et al., 1999).

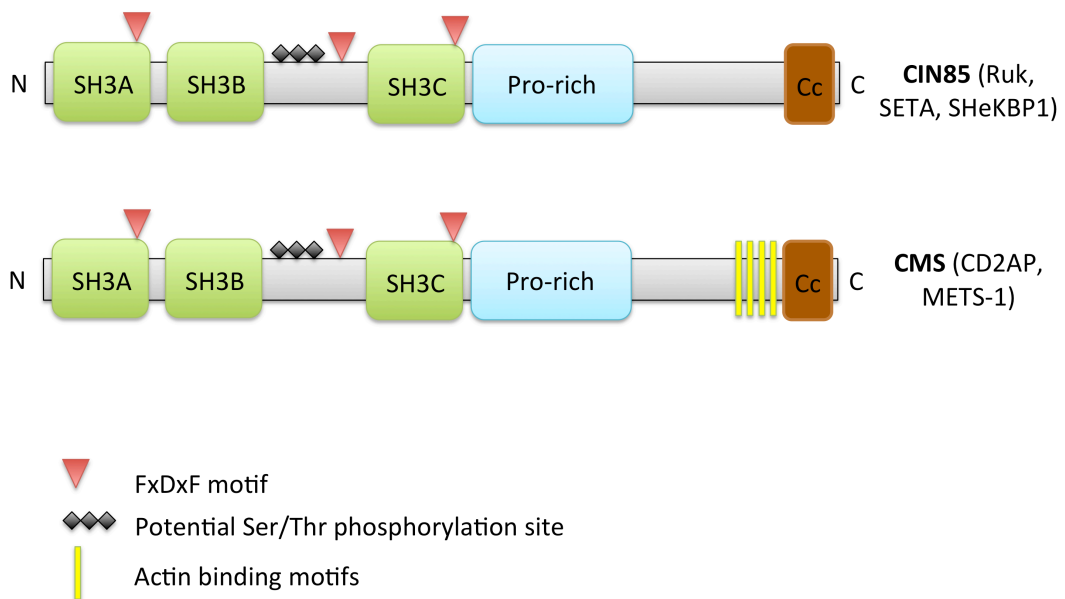


Figure I.3. Schematic representation of the domain structure of CIN85 and CD2AP adaptor proteins. Both adaptors contain three N-terminal SH3 domains, a centrally located proline-rich region (Pro-rich), a globular domain and a C-terminal Coiled-coil (Cc) domain. CD2AP additionally contains an actin-binding motif at the C-terminal, which is lacking in CIN85. Figure adapted from Dikic I., 2002.

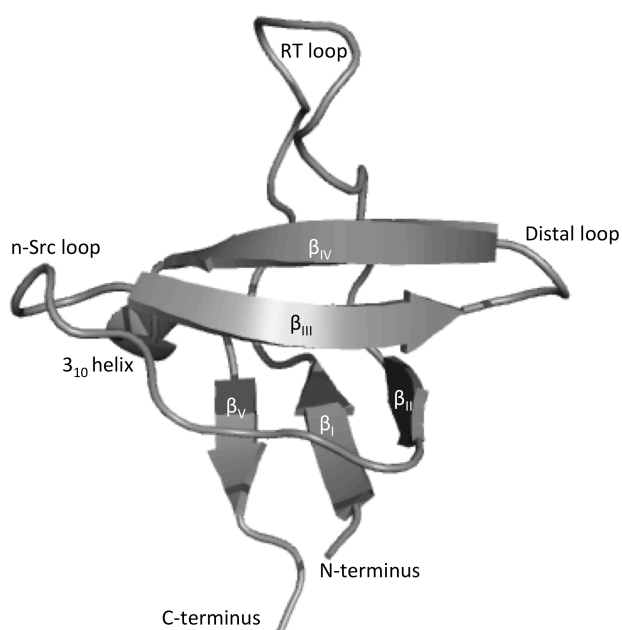
We can conclude that CD2AP and CIN85 act as anchoring proteins involved in several signaling pathways, controlling cellular adhesion, mobility and morphology. This versatility of roles in many cellular processes, make these proteins an interesting study-object from the medical and therapeutically point of view. Therefore, it is essential to understand the biophysics of the interactions mediated by the modular domains of these adaptor proteins.

1.4 SH3 domains

Since their discovery 25 years ago (Mayer B.J. et al., 1988), SH3 domains have been the focus of numerous studies. The human, drosophila and yeast genome contain 300, 90 and 28 different SH3 domains, respectively (Carducci M. et al., 2012). These small domains (generally composed of 60 - 70 amino acids) are parts of adaptor proteins that participate in intracellular communication networks, organization of the cytoplasm skeleton and membrane trafficking (Skorski T. et al., 1998; Stein R., 1998; McPherson P.S., 1999).

SH3 domains are protein modules of general importance in many signaling processes in all organisms, suggesting that they regulate basic aspects of cellular organization. Most of SH3 domain-mediated physiological interactions are moderate or weak, which facilitates the rapid response to activation or attenuation of cellular signals by the formation and dissociation of specific signaling complexes. Therefore, most protein-protein interactions occurring in cellular signal transduction are transient, making it a very dynamic process (Jia C.Y. et al., 2005).

From the structural point of view, all SH3 domains share the same 3D structure (Fig. I.4), consisting of a β -barrel made up of five antiparallel β -strands (generally named β through β e or β I through β IV). These β -strands form two β -sheets that pack against each other at approximately right angles. The first sheet is formed by β -strands I, IV and the first half of II; while the second sheet is formed by β -strands III and IV; a kink in β -strand II allows it to participate in both β sheets simultaneously. There are three variable loops: the RT-loop, which owes its name to its high content in arginine and threonine residues (located between β -strands I and II); the n-Src loop, named after the first protein in which the SH3 3D structure was described (between β -strands II and III); and the Distal loop, which is located at the opposite site of the ligand binding interface (between β -strands III and IV). The variability in the length of the SH3 domains results



generally from insertions or deletions in the n-Src and Distal loops. The residues separating β -strands IV and V are generally found in a 3_{10} helix conformation (Larson S.M. and Davidson A.R., 2000; Kishan K.V. et al., 2001).

Figure I.4. Cartoon representation of the high-resolution structure of the SH3A domain of CD2AP (PDB code 2J6K), used as example to highlight the secondary structure elements characteristic of SH3 domains. β -strands are labeled as β_I to β_{IV} . The RT-loop, n-Src loop, distal loop, 3_{10} helix and C- and N-terminus are marked.

SH3 domains recognize proline-rich sequences to which they generally bind with low affinity (K_d between 10^{-6} and 10^{-4} M). The arrangement of the SH3 domains forms a relatively flat and hydrophobic surface, consisting of one acidic and two hydrophobic shallow pockets between the two β -sheets (Fig. I.5). This surface is involved in the recognition and binding of proline-rich motifs (Larson S.M. and Davidson A.R., 2000; Cesareni G. et al., 2002; Ferreón J.C. and Hilser V.J., 2004). Most of SH3 domains bind to a proline-rich sequence with a ϕ Px ϕ P motif (also known as PxxP motif), where ϕ is usually a hydrophobic residue and x any residue (Mayer B.J., 2001). The ligand adopts an extended, left-handed helical conformation termed the polyproline type II (PPII) helix. Proline residues are in the trans conformation and the helical structure shows precisely three residues per turn (Adzhubei A.A. and Sternberg M.J., 1993), which makes this structure roughly triangular in the cross-section, with the base of this triangle sitting on the SH3 domain surface. The two hydrophobic ligand-binding pockets of the SH3 domain are occupied by two hydrophobic-proline (ϕ P) dipeptides, whereas the acidic pocket typically interacts with a basic residue in the ligand distal to the ϕ Px ϕ P core (Fig. I.5A). This basic residue binds the acidic pocket in the SH3 domain, which has been thought to provide the binding specificity and denoted as the “specificity pocket”. Proline-rich peptide ligands potentially bind a given SH3 domain in either one of two opposite directions governed by the location of such positively charged residue, usually Arg, which often precedes or follows the ϕ Px ϕ P core element (Feng S. et al., 1994; Lim W.A. et al., 1994). According to this, the ϕ Px ϕ P motif is usually further classified into +x ϕ Px ϕ P (class I) and ϕ Px ϕ Px+ (class II) (where + is generally a basic residue, usually an arginine). Thereby, the ϕ P dipeptides occupy different positions on the surface of the hydrophobic grooves, depending on the orientation of the ligand so that, relatively to the helix axis, class I binds conserved prolines on the motif in the same side as the basic residue, while class II does so on the opposite side.

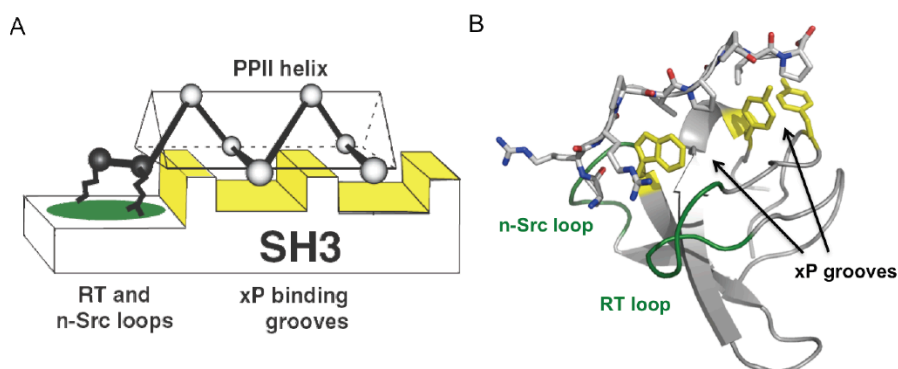


Figure 1.5. *A. Schematic representation of the interaction of proline-rich sequences with SH3 domains. B. Binding of the c-Src SH3 domain to a class II proline-rich ligand (PDB code 1QWF). Same colors are used as in A. Figure reprinted from Zarrinpar A. et al., 2003.*

Despite that the shallow binding surface of SH3 domains bears certain characteristics for the accommodation of the PPII helical structure, it has been shown that a given SH3 domain can interact with a few to several dozens of different peptide ligands (Li S.S., 2005).

It is common for SH3-interacting proteins to contain multiple weak to moderate binding sites. Computational methods have been developed to identify putative canonical SH3

domain binding sequences within proline-rich regions of proteins to narrow down the size of peptide libraries used for high throughput analysis (Ferraro E. et al., 2007). The SH3 domain can dissociate quickly from one site and subsequently be recaptured by a neighboring site in the same molecule. Due to that, the presence of multiple binding sites together with the capability of SH3 domains to recognize a collection of sequence motifs, increases the local concentration of SH3 domain and thereby promotes binding (Jia C.Y. et al., 2005). Therefore, the incidence of various SH3 binding motifs in the same molecule might be a mechanism used by the cell to regulate SH3-mediated interactions (Li S.S., 2005).

Despite these overlapping recognition profiles, selectivity is essential in the SH3 domain-mediated interaction network. Differences in binding motifs containing the sequence PxxP have been shown to provide various levels of specificity. In particular, the basic residue of the motif plays an important role at enhancing specificity in SH3 domain-mediated interactions (Zarrinpar A. et al., 2003). The evolution of a non-canonical recognition mechanism offers an alternative solution to the need of binding diversity of SH3 domains to achieve the required specificity. Several studies show that certain SH3 domains recognize non-PxxP motifs, like the SH3 domains of Eps8, recognizing PxxDY (Mongiovi A.M. et al., 1999), Gads, recognizing RxxK (Berry D.M. et al., 2002), or Fus1, recognizing Arg-Ser-rich sequences (Tong A.H. et al., 2002).

Interestingly, the SH3 domains of the adaptor proteins CD2AP and CIN85 recognize an atypical proline-arginine motif (PxxxPR) present in the C-terminal region of c-Cbl, Cbl-b and CD2 (Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003). Mutagenesis studies revealed the importance of the proline and arginine residues flanking the PxxxPR motif. Additional cellular CIN85 and CMS interacting proteins, such as Pak, Disabled 1 and 2, SETA-binding protein 1 (SB-1), SLP-65/BLNK, or Alg2-interacting protein (AIP1) and ASAP, also contain this atypical PxxxPR consensus sequence (Dikic I., 2002; Kowanetz K. et al., 2003).

Recent studies have shown that some SH3 domains are capable of binding to non-proline-rich sequences, as observed in the interaction between ubiquitin and Sla1-SH3-3 (He Y. et al., 2007), the three SH3 domains of CIN85 (Bezsonova I. et al., 2008) and SH3A and SH3C domains of CD2AP (Ortega-Roldan J.L. et al., 2009, Ortega-Roldan J.L. Ph.D thesis, 2010), using the same hydrophobic patch on the SH3 domain surface than the proline-rich peptides (Stamenova S.D. et al., 2007). SH3 domains can thus also be considered as ubiquitin-binding domains.

1.5 Ubiquitin and ubiquitin-binding domains

Ubiquitin is a highly conserved 76 residues protein whose conjugation to substrate proteins acts as a signal altering protein localization or function and/or regulating protein-protein interactions (Hershko A. and Ciechanover A., 1998; Hicke L., 2001; Pickart A.M. and Eddins M.J., 2004; Sun L. and Chen Z.J., 2004). Ubiquitin is conjugated to target proteins via an isopeptide bond formed between its C-terminal residue (Gly76) and a side chain lysine on the target protein (Fig. I.6). Conjugation of a single ubiquitin to target proteins is required for diverse processes such as receptor-mediated endocytosis, vesicular trafficking (Hicke L., 2001; Raiborg C. et al., 2006;

Staub O. and Rotin D., 2006; Dikic I. et al., 2009), cell-cycle control, stress response, DNA repair (Huang T.T. et al., 2006), signaling (Di Fiore P.P. et al., 2003; Haglund K. and Dikic I., 2005), transcription and gene silencing. Additionally, any of the seven lysines present in ubiquitin, can be used to form an isopeptide bond to another ubiquitin (Pickart C.M. and Fushman D., 2004), resulting in polyubiquitin chains (Fig. I.6) with different arrangements (Varadan R. et al., 2004). The best understood function of ubiquitination is proteolysis, whereby Lys48-linked polyubiquitin chains allow recognition by the 26S proteasome. However, proteins can also be monoubiquitinated or polyubiquitinated through alternative (e.g. Lys63) linkages. Lys63-polyubiquitin chains are associated with non-proteolytic roles in DNA repair, DNA damage tolerance, NF- κ B signaling and translation (Pickart C.M. and Eddins M.J., 2004; Sun L. and Chen Z.J., 2004; Chiu Y.H. et al., 2009; Chen Z.J. and Sun L.J., 2009). Likewise, monoubiquitination is associated with endocytosis, trafficking and transcriptional control (Hicke L. and Dunn R., 2003; Shilatifard A., 2006). The ubiquitination event can be viewed as a molecular zip code, which is used to direct different ubiquitination products to different destinations. Errors in delivery of ubiquitinated proteins to the proteasome or other destinations could be disastrous for cells.

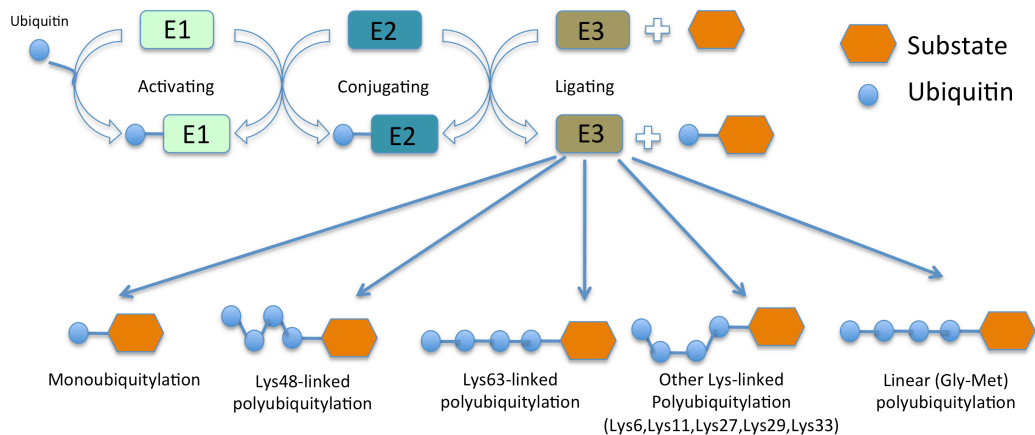


Figure I.6. Schematic representation of the ubiquitination processes, carried out by three enzymes, ubiquitin-activating enzyme E1, ubiquitin-conjugating enzyme E2, and ubiquitin ligase enzyme E3. The repetition of this enzymatic process leads to the formation of polyubiquitinated chains on the substrate, linked by an amide bond between the carboxylic acid of the C-terminal Gly of ubiquitin with the amine of the Lys of other ubiquitin. The residue number of that Lys will determine the name of the polyubiquitin chains, e.g. Lys48-linked polyubiquitin. In the case of linear polyubiquitin, the linkage takes place between the C-terminal Gly of one ubiquitin and the N-terminal Met of another ubiquitin molecule. Figure adapted from Dikic I. et al., 2009.

Recent progresses in the discovery of new biological roles of ubiquitin have gone hand in hand with the discovery of a multitude of ubiquitin-binding domains (UBDs) (Hicke L. et al., 2005; Harper J.W. et al., 2006) or motifs that non-covalently bind to ubiquitin. They are generally small domains (20 - 150 residues in length) found in enzymes that catalyze ubiquitination or deubiquitination, or in ubiquitin receptors that recognize and interpret signals from ubiquitinated proteins (Hicke L. et al., 2005). The characterization of these domains has become a major foundation for advancing the biology of ubiquitin-based regulatory mechanisms. UBDs can bind, and often distinguish, different types of ubiquitin modifications. The number of identified UBDs is constantly growing, with more than twenty different families identified to date (Table I.1).

Ubiquitin-binding domain	Representative protein	Function
<i>α-Helix</i>		
UIM	S5a (human) and Rpn19 (yeast), Vps27, STAM, epsins and RAP80 (UIMC1)	Proteasome degradation, endocytosis, MVB biogenesis and DNA repair
IUIM (also known as MIU)	RABEX5	Endocytosis
DUIM	HRS	MVB biogenesis
UBM	Polymerase iota and reversionless 1	DNA damage tolerance
UBAN	NEMO, ABIN1-ABIN3 and optineurin	Nuclear factor-κB signaling
UBA	Rad23(yeast)and R23A(human), Dsk2 and NBR1	Proteasome targeting, kinase regulation and autophagy
GAT	GGA3 and TOM1	MVB biogenesis
CUE	Vps9, TAB2 and TAB3	Endocytosis and kinase regulation
VHS	STAM and GGA3	MVB biogenesis
<i>Zinc finger (Znf)</i>		
UBZ	Polymerases eta and kappa; WRNIP1 and TAX1BP1	DNA damage tolerance and nuclear factor-κB signaling
NZF	NPL4, Vps36, TAB2 (MAP3k7IP2) and TAB3 (MAP3k7IP3)	ERAD, MVB biogenesis and kinase regulation
ZnF A20	RABEX5 (RABGEF1) and A20 (TNFAIP3)	Endocytosis and kinase regulation
Znf UBP (also known as PAZ)	Isopeptidase T (USP5) and HDAC6	Proteasome function, aggresome function and autophagy
<i>Pleckstrin homology (PH) domain</i>		
PRU	RPN13	Proteasome function
GLUE	EAP45(VPS36)	MVB biogenesis
<i>Ubiquitin-conjugating(Ubc)-like domain</i>		
UEV	UEV1(UBE2V1)and MMS2	DNA repair, MVB biogenesis and kinase regulation
UBC	UBCH5C(UBE2D3)	Ubiquitin transfer
<i>Others</i>		
SH3	Sla1 and CIN85 (SH3KBP1)	Endocytosis
PFU	Ufd3(Doa1)	ERAD
Jab1/MPN	Prp8	RNA splicing

Table I.1. Summary of the functional and structural diversity of ubiquitin-binding domains. Reprinted from Dikic I. et al., 2009.

UBDs diverge both in structure and in the type of ubiquitin recognition that they use. Several reviews have described the various classes of UBDs and their biological properties (Di Fiore P.P. et al., 2003; Hicke L. et al., 2005; Hurley J.H. et al., 2006; Harper J.W. and Schulman B.A., 2006; Hofmann K., 2009; Dikic I. et al., 2009).

Many of these UBDs include an α -helical domain (UBZ, UIM, MIU, DUIM, A20 ZnF, CUE, UBA, GAT, and so on), or an α/β -domain (UEV), or extended loops forming an ubiquitin binding interface (NZF motif). However, there are no obvious common topological features for ubiquitin-binding domains. For example, SH3 domains are structurally unrelated to any of these ubiquitin-binding domains.

UBDs use to recognize the hydrophobic patch on the surface of ubiquitin, known as the canonical or Leu8-Ile44-Val70 patch (based on the residues in which it is centered). However, some UBDs can recognize monoubiquitin by binding other regions on its surface (Fig. I.7). For example, the zinc finger domain (ZnF) of the nuclear protein localization 4 (NPL4) (Alam S.L. et al., 2004), which serves as a ubiquitin-binding adaptor protein in the endoplasmic reticulum-associated degradation (ERAD) pathway, and the ubiquitin-binding ZnF (UBZ) domain, which is present in translesion synthesis (TLS) polymerases (Bienko M. et al., 2005), bind to the canonical hydrophobic surface of ubiquitin. By contrast, the A20-type ZnF domain of RAB5 guanine nucleotide exchange factor (RABEX-5) recognizes a polar surface of ubiquitin that is centered around Asp58 (Lee S. et al., 2006), and the DUB isopeptidase T (Iso T), has a ZnF domain that binds to the C-terminal residues of ubiquitin (Reyes-Turcu F.E. et al., 2006).

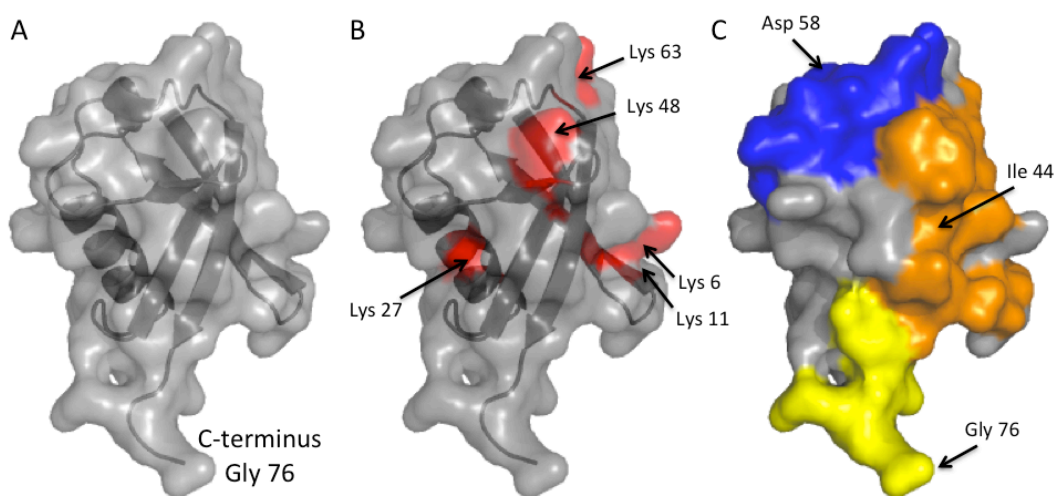


Figure I.7. Surface representations of ubiquitin: **A.** Combined cartoon and surface representation of the X-ray structure of ubiquitin (PDB code: 1UBQ). The C-terminal Gly76 is marked. **B.** Position of 5 lysine residues out of the 7 (red) in ubiquitin. **C.** Major recognition patches on ubiquitin. The hydrophobic patch centered around Ile44 (orange), the polar patch centered around Asp58 (blue) and the diglycine patch near the C-terminal Gly76 (yellow) are shown.

The structure and affinities of ubiquitin-binding domains seem to be very diverse. The early view that all ubiquitin-binding domains had a very low affinity for monoubiquitin appears now to be a too simple generalization. Indeed, the affinity of ubiquitin-binding domains for monoubiquitin can be as high as $K_d \approx 1 \mu\text{M}$, or lower than $100 \mu\text{M}$ (Hurley J.H. et al., 2006). The weak interactions between individual domains and monoubiquitin are leveraged into physiologically relevant high-affinity interactions via several mechanisms: ubiquitin polymerization, oligomerization of ubiquitinated proteins and binding domain proteins, tandem binding domains, binding domains with multiple ubiquitin-binding sites and cooperativity between ubiquitin binding.

Linker regions in tandem repeats of UBDs can also define linkage specificity. For example, receptor associated protein 80 (RAP80; also known as UIMC1) targets BRCA1 to DNA damage-induced foci (Yan J. et al., 2007; Sobhian B. et al., 2007, Wang B. et al., 2007) through its two UIMs (Ubiquitin Interacting Motives), which bind Lys63-linked ubiquitin chains but not Lys48-linked ones (Sims J.J. et al., 2009). Indeed, the sequence between the two UIMs of RAP80 positions them for efficient avid binding across a single Lys63-linkage, thus defining selectivity (Sims J.J. et al., 2009). Hence, the coordinated action of multiple UBDs can be used to direct substrates towards specific functional pathways according to their ubiquitin chain linkage type.

1.6 TAX1BP1 a novel ubiquitin-binding adaptor protein

Tax1-binding protein 1 (TAX1BP1; also known as TXBP151 or T6BP) was initially identified as a molecule that binds the human T-cell leukemia virus type I Tax oncoprotein, the HTLV-I viral oncoprotein (Chin K.T. et al., 2007), which participates in HTLV-I-mediated transformation of infected cells and adult T cell leukemia by inducing the persistent activation of NF- κ B through multiple pathways (Boxus et al., 2008). Tax is a potent NF- κ B activator, and it efficiently immortalizes mammalian peripheral blood lymphocytes *ex vivo* (Yoshida M. et al., 2001; Matsuoka M. and Jeang K.T., 2005; Matsuoka M. and Jeang K.T., 2007). Overexpression of TAX1BP1 inhibits (Iha H. et al., 2000), whereas TAX1BP1 deficiency enhances, Tax1-induced NF- κ B activation (Shembade N. et al., 2007), pointing to a role for TAX1BP1 in preventing Tax1-mediated oncogenic cell transformation.

Two human TAX1BP1 transcripts encoding the modular proteins of 789 (Isoform 1) and 747 (isoform 2) amino acids (De Valck D. et al., 1999) were identified. The isoform 1 (identifier: Q86VP1-1), with a molecular weight of 86.2 kDa, has been widely used to study this system. The human TAX1BP1 gene is located on chromosome 7p15 (Nagaraja G.M. and Kandpal R.P., 2004) and expressed in all tissues. The expression of TAX1BP1 can be modulated, e.g. the expression level of TAX1BP1 is increased due to the transcriptional activation of the bovine papillomavirus type 1 (BPV1) E2 protein (Wang X. et al., 2009), upregulation is also observed in hepatocytes upon expression of hepatitis C virus (HCV) core protein (Nguyen T.H. et al., 2006) and TAX1BP1 is downregulated after treatment with chemotherapy in epithelial ovarian cancer patients (L' Esperance S. et al., 2006). However, mechanisms contributing to the regulation of TAX1BP1 expression have not yet been described. The adaptor protein TAX1BP1 is localized in intra-nuclear speckles, diffusely in the cytosol and close to cell-matrix adhesion sites in the plasma membrane (Chin K.T. et al., 2007; Ulrich M. et al., 2007; Morriswood B. et al., 2007).

1.6.1 Function of TAX1BP1

The binding partners and function of TAX1BP1 have been recently described in an excellent review article (Verstrepen L. et al., 2011). As it is expected from an adaptor protein, TAX1BP1 is a multitask protein involved in different mechanisms, such as transcription regulation, cell growth and apoptosis. For a better understanding of the physiological function of TAX1BP1, the main processes in which this protein is involved are summarized as follows:

- Negative regulation of NF- κ B:

NF- κ B family of transcription factors is key in inflammatory responses. Inflammation is an important component of innate immunity, which provides the organism first-line defense by sensing and protecting against invading pathogens and endogenous dangerous signals. Its distinctive output is the result of concerted signaling cascades that trigger robust inflammatory and antimicrobial responses (Iwasaki A. and Medzhitov R., 2010).

NF- κ B signaling consists of two distinct pathways: the canonical and the non-canonical pathways (Fig. I.8). In the canonical NF- κ B pathway (Fig. I.8), the hetero-dimeric NF- κ B proteins such as RelA (p65) and p50 are retained in the cytoplasm as inactive subunits by physical association with members of the inhibitor of NF- κ B (I κ B) family, most notably I κ B α (Karin M. and Ben-Neriah Y., 2000). Proinflammatory cytokines, including tumor necrosis factor- α (TNF- α) and interleukin-1 (IL-1), viral-encode proteins, such as the human T-cell leukemia virus type I (HTLV-I) Tax, and bacterial lipopolysaccharides (LPS) all activate the I κ B kinase (IKK) complex, which is composed of two catalytic subunits, IKK α and IKK β and a regulatory component, IKK γ (also called NEMO or NF- κ B Essential MOdulator) (Häcker H. and Karin M., 2006). The activated IKK phosphorylates I κ B α , triggering its proteolysis through the ubiquitin/proteasome pathway, facilitating nuclear import of NF- κ B and activation of target genes (Karin M. and Ben-Neriah Y., 2000) encoding molecules regulating inflammation and survival.

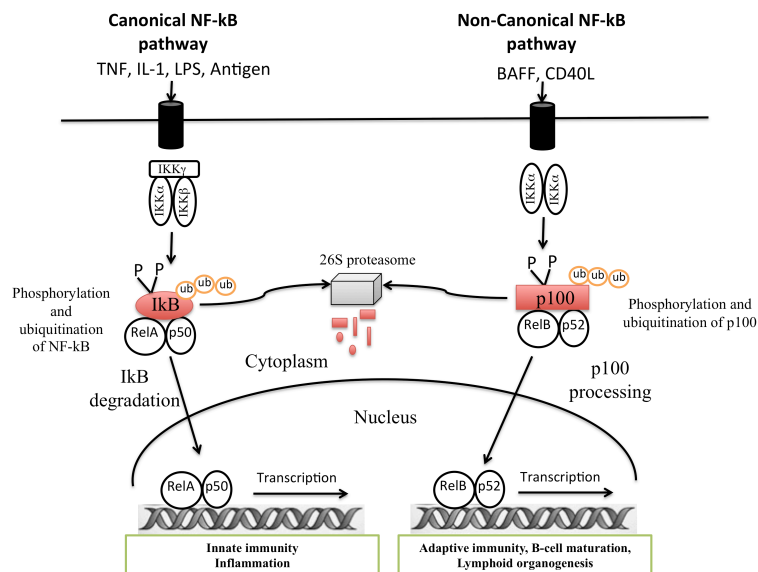


Figure I.8. Schematic representation of the activation pathways of NF- κ B. Activation of NF- κ B by external stimuli involves canonical and non-canonical pathways, which are based on the degradation of I κ B or processing of p100, respectively. Free NF- κ B (RelA-p50 in the canonical pathway; and RelB-p52 in the non-canonical) enters the nucleus and activates gene expression. Figure adapted from Lee M.S. and Kim K.A., 2006.

In the non-canonical NF- κ B pathway, various ligands of the TNF superfamily, such as BAFF (B-cell activating factor) and CD40L, trigger processing of the precursor protein p100 (NF- κ B2) to yield the NF- κ B subunit p52 (Coope H.J. et al., 2002; Claudio E. et al., 2002). RelB and p52 regulate genes encoding molecules involved in B cell survival

and lymphoid organogenesis. IKK α is activated in the non-canonical pathway by the kinase NIK (NF- κ B inducing kinase) (Senftleben U. et al., 2001; Xiao G., et al., 2001). In this pathway the complex IKK is only formed by the two catalytic subunits IKK α and NEMO (IKK γ) is not present (Israël A., 2010).

Aberrant activation of the NF- κ B signaling pathways leads among others to immunodeficiency, septic shock or induction of autoimmunity (Chen F. et al., 1999). For this reason NF- κ B is tightly regulated by numerous mechanisms maintaining its transient activation in order to prevent inflammation-induced tissue damage or malignancy (Karin M. and Greten F.R., 2005). Ubiquitination plays important regulatory roles in several steps of the NF- κ B signaling events.

The deubiquitinating enzyme A20 (also known as TNFAIP3) is a major regulator of NF- κ B and interferon regulatory factor 3 (IRF3) signaling (Verstrepen L. et al., 2010) (Fig. I.9). Polymorphisms in the gene encoding A20 have been linked to many inflammatory autoimmune diseases in humans (Coornaert B. et al., 2009). Studies indicate that A20 also functions as a tumor-suppressor gene in several subtypes of B cell lymphomas (Compagno M. et al., 2009; Kato M. et al., 2009). Therefore, understanding the mechanisms of A20-mediated downregulation of NF- κ B is of critical importance. A20 has been shown to function as a unique enzyme with dual-function deubiquitinase and E3 ligase activities. Its N-terminal domain [an ovarian tumor (OTU) domain] harbors deubiquitinating activity and can remove Lys63-linked ubiquitin chains from specific NF- κ B signaling molecules such as receptor interacting protein (RIP) 1, RIP2 and TNF receptor associated factor (TRAF) 6, thus blocking signaling initiated by TNF, IL-1, NOD2 (Nucleotide-binding oligomerization domain-containing protein 2) and TLR4 (Toll-like receptor 4) (Wertz I.E. et al., 2004; Boone D.L. et al., 2004; Hitotsumatsu O. et al., 2008). The C-terminal part of A20 contains seven zinc fingers that are involved (ZnF4 is essential) in the conjugation of Lys48-linked ubiquitin residues to RIP1 and the E2 ubiquitin conjugating enzyme, leading to their proteosomal degradation (Wertz I.E. et al., 2004; Bosanac I. et al., 2010; Shembade N. et al., 2010). Genetic alterations of A20 have been associated with the development of B-cell lymphomas (Kato M. et al., 2009), autoimmune disorders and hematologic malignancies (Plenge R.M. et al., 2007; Thomson K.J. et al., 2007; Musone S.L. et al., 2008; Compagno M. et al., 2009; Honma I. et al., 2009; Kato M. et al., 2009; Novak A.J. et al., 2009; Schmitz N. et al., 2009).

The NF- κ B and IRF3 inhibitory functions of A20 are dependent on several other proteins forming a complex with A20. This complex includes the E3 ubiquitin ligase ring finger protein 11 (RNF11) and itchy E3 ubiquitin protein ligase homolog (Itch), which have been shown to facilitate RIP1 degradation in an E3 ligase-dependent and non-redundant manner (Shembade N. et al., 2008; Shembade N. et al., 2009), and the protein TAX1BP1, that interacts with A20 forming the TAX1BP1-Itch-A20 ubiquitin-editing complex. ABIN-1 (A20-Binding Inhibitor of NF- κ B), a potential binding-partner of RNF11, also interacts with A20 and is required for A20 to deubiquitinate NEMO, suggesting that ABIN-1 is also a component of the A20 ubiquitin-editing complex (Mauro C. et al., 2006). TAX1BP1 works as a crucial ubiquitin-binding adaptor protein that recruits A20 to specific NF- κ B signaling molecules, leading to their A20-mediated ubiquitin editing and the termination of NF- κ B signaling (Fig. I.9). Indeed, TAX1BP1 depends of A20 for its NF- κ B inhibitory potential (De Valck D. et al., 1999; Iha H. et al., 2008). However, the formation of this complex is also regulated,

Tax acts as an antagonist of the A20 ubiquitin-editing complex by disrupting interactions among TAX1BP1, Itch and A20, to prevent downregulation of NEMO ubiquitination and IKK activation.

TAX1BP1 has also been identified as a TRAF6-binding protein. TRAF6 is an ubiquitin ligase that is involved in NF- κ B and c-Jun N-terminal kinase (JNK) signaling in response to multiple receptors, including IL-1 receptors and TLR4 (Ling L. et al., 2000). It was shown that TAX1BP1 contains at least two TRAF6-binding regions (Ling L. et al., 2000), one in the C-terminal and other in the N-terminal region. Moreover, Lys63-ubiquitinated TRAF6 is able to interact with the C-terminal second ubiquitin-binding domain (UBZ2) of TAX1BP1 (Iha H. et al., 2008). Similarly, TAX1BP1 could specifically interact with the Lys63-ubiquitinated RIP1 in the TNF signaling pathway (Iha H. et al., 2008).

These findings indicate that TAX1BP1 functions as an adaptor protein, which recruits A20 to the ubiquitinated signaling proteins TRAF6 and RIP1 (Fig. I.9), leading to their A20-mediated deubiquitination and the disruption of IL-1 and TNF-induced NF- κ B signaling, respectively (Shembade N. et al., 2007; Iha H. et al., 2008). The role of TAX1BP1 in this pathway is so important that mice genetically knocked out for TAX1BP1 develop inflammatory cardiac valvulitis, dermatitis and hypersensitivity to low doses of TNF and IL-1, and die prematurely (Iha H. et al., 2008).

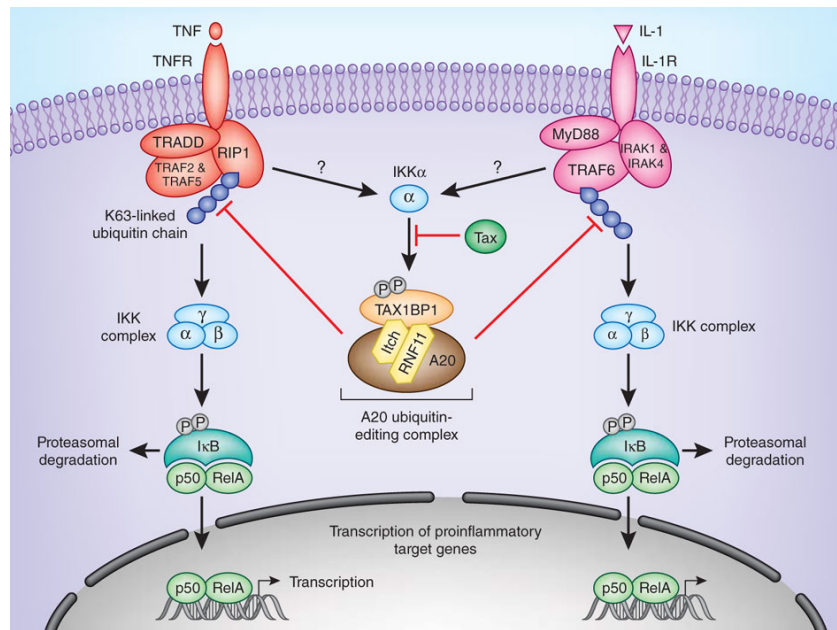


Figure I.9. Detailed scheme of the canonical pathway of NF- κ B in which TAX1BP1 is involved. The binding of cytokines, like TNF, or interleukins, like IL-1, to the surface receptors of the cell, induces the assembly of a signaling complex that promotes the activation of the IKK complex (this activation depends on the Lys63-linked polyubiquitination of RIP1 and TRAF6), which results in the degradation of I κ B. The A20-ubiquitin-editing complex, composed of A20, Itch, RNF11 and TAX1BP1, inhibits the activation of NF- κ B induced by TNF and IL-1, by cleaving the Lys63-linked polyubiquitin chains. Tax inhibits IKK α -mediated phosphorylation of TAX1BP1, preventing the assembly of the active A20 ubiquitin-editing complex, and favoring the persistent activation of NF- κ B (Shembade N. et al., 2007; Iha H. et al., 2008). Figure taken from Pelzer C. et al., 2011.

- Negative regulation of cell growth and apoptosis:

The binding of the cytokine TNF to its receptor initiates NF- κ B proinflammatory signaling but also caspase-mediated apoptosis; and both events are regulated by A20 (Verstrepen L. et al., 2010). In addition to regulating the A20 function in anti-inflammatory and antiviral signaling pathways, TAX1BP1 also mediates the anti-apoptotic activity of A20 (De Valck D. et al., 1999). Specifically, apoptosis induced by TNF is associated with proteolysis of TAX1BP1 and overexpression of TAX1BP1 inhibits TNF-induced apoptosis in specific cell lines (De Valck D. et al., 1999). However, the physiological role of the anti-apoptotic and cell growth regulatory effects of TAX1BP1 remains unclear. Upregulation of TAX1BP1 and A20 in hepatocytes by the inducible expression of the HCV core protein, indicates a possible role of TAX1BP1 in immune invasion and viral persistence by protecting infected hepatocytes from cell death (Nguyen T.H. et al., 2006).

- Positive regulation of gene expression

Till recently only the role of TAX1BP1 as negative regulator of gene expression by interfering with NF- κ B activation was reported. However, a study (Journó C. et al., 2009) demonstrated that TAX1BP1 can also increase Tax1-induced NF- κ B activation and in this way, positively regulate gene expression. This is possible because NEMO-related protein (NRP) also known as Opiterium, and TAX1BP1 can form a functional complex with Tax1 that contributes to Tax1-mediated NF- κ B activation (Journó C. et al., 2009). These results are in contradiction with the previous observations where TAX1BP1 was shown to inhibit Tax1-induced NF- κ B activation (Iha H. et al., 2000; Shembade N. et al., 2007). These discrepancies might indicate that TAX1BP1 function depends on the composition of the protein complex in which it acts.

TAX1BP1 also enhances the transcriptional activity of the human papilloma virus E2 protein (Wang X. et al., 2009), involving the binding of TAX1BP1 with E2 proteins as well as the transcriptional coactivator p300.

Furthermore, TAX1BP1 shares 27% identity and 45% similarity with the coiled-coil transcription coactivator protein CoCoA, known to enhance transcriptional activation by nuclear receptors (Kim J.H. et al., 2003). This homology suggests a potential role of TAX1BP1 as transcriptional coactivator.

1.6.2 Structure of TAX1BP1

The TAX1BP1 protein is highly conserved among species with human TAX1BP1, sharing 79% and 81% identity with rat and mouse orthologs, respectively. The N-terminal part of TAX1BP1 contains a SKICH (skeletal muscle and kidney enriched inositol phosphatase carboxyl homology) domain (Fig. I.10). SKICH domains have been described as membrane targeting domains, but this property has not been demonstrated for TAX1BP1 (Gurung R. et al., 2003). The central part of TAX1BP1 is predicted to form three coiled-coil structures and a homodimerization region (Ling L. et al., 2000). At its C-terminus, TAX1BP1 contains two highly conserved zinc finger domains (Fig. I.10). Notably, there is a conserved PPXY motif at the same position in each of the zinc fingers (Fig. I.11). The PPXY motifs mediate interactions with proteins containing WW domains (Sudol M. et al., 1995).

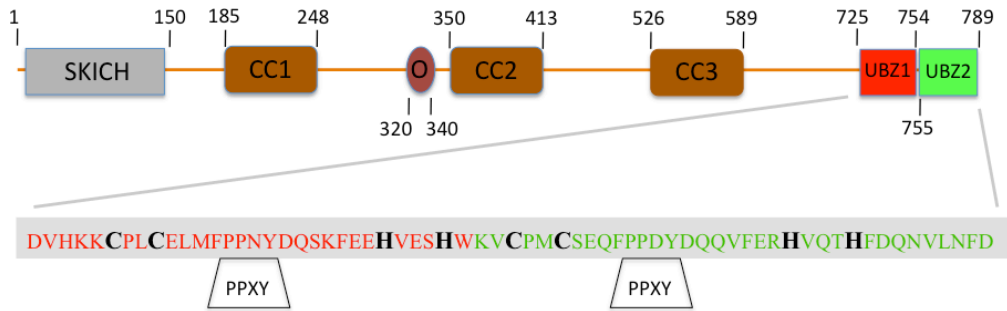


Figure I.10. Schematic representation of the domains present in human TAX1BP1. The numbers indicate the amino acids delimitating the structural motifs. The N-terminus contains a SKICH domain, SKIP (skeletal muscle and kidney enriched inositol phosphatase) carboxyl homology domain, the central part contain three coiled-coil (CC1, CC2 and CC3) structures and a region (O) which is responsible for homodimerization, and the C-terminal extreme contains the two zinc fingers domains UBZ1 (red) and UBZ2 (green), that are classified like classical (Cys₂His₂) zinc fingers domain. The sequences of the UBZ1 and UBZ2 domains are given, showing the conserved Cys and His in bold, the conserved PPXY motifs are also indicated. The color code red for UBZ1 and green for UBZ is kept throughout the whole thesis.



Figure I.11. Sequence alignment of different mammals (Human, Gorilla, mouse, dog, cow and sheep) and oviparous (Chicken) TAX1BP1 zinc fingers domains (UBZ1 and UBZ2). PPXY motifs are highlighted. Conserved Cys and His residues are in bold.

The TAX1BP1 zinc fingers function as novel ubiquitin-binding domains (UBZ) and can bind Lys63-ubiquitinated RIP1 and TRAF6 (Iha H. et al., 2008) as well as the E3 ubiquitin ligase Itch (Shembade N. et al., 2008). The PPXY motifs of TAX1BP1 (Fig. I.11) are critically important for binding to Itch, whereas the zinc-finger domains are more important for binding to TRAF6 and RIP1 (Shembade N. et al., 2008). The zinc fingers are also involved in the binding of TAX1BP1 to the motor protein myosin VI (Morriswood B. et al., 2007), which has been implicated in many cellular processes including endocytosis, secretion, membrane ruffling and cell motility. Indeed, TAX1BP1 colocalizes with myosin VI at the trans side of the Golgi complex, thus suggesting a role in secretion (Morriswood B. et al., 2007).

The interaction mechanism of TAX1BP1 with ubiquitin through the zinc finger domains, remain unknown. Identification of the binding motif of the UBZ domains of TAX1BP1 and structural characterization of their interaction with ubiquitin, would provide a molecular explanation on how TAX1BP1 acts as a scaffolding protein in the NF- κ B signal transduction pathway. This information could be useful to prevent chronic inflammation and other related diseases like, cardiac valvulitis, autoimmune disorders and hematologic cancers (Graham T.R. et al., 2008; Musone S.L. et al., 2008; Compagno M. et al., 2009; Honma I. et al., 2009; Kato M. et al., 2009).

1.7 Zinc finger domains

Zinc-fingers domains (ZnFs) are small protein motifs extremely abundant in higher eukaryotes, whose name comes from the fact that zinc plays a structural role contributing to the stability of the domains. They were first identified in the *Xenopus laevis* (African clawed frog) transcription factor TFIIA (Miller J. et al., 1985). These TFIIA-type or “classical” ZnFs have proved to be common in complex organisms: for example, in humans, more than 15.000 of these domains are predicted to exist in approximately 1.000 different proteins (Rubin G.M. et al., 2000). Once considered to function exclusively as sequence-specific DNA-binding motifs, ZnFs are now known to have additional interacting targets such as RNA, lipids and other proteins. ZnFs are structurally diverse and present among proteins performing a broad range of functions in various cellular processes, such as replication and repair, transcription and translation, metabolism and signaling, cell proliferation and apoptosis. Structural studies of ZnF complexes revealed considerable diversity in terms of protein partners, binding modes and affinities, and highlighted the often underestimated versatility of ZnF structure and function.

The fact that ZnFs, in addition to their interactions with DNA and RNA, can also mediate protein–protein interactions (Mackay J.P. and Crossley M., 1998) makes them attractive targets for the study of signaling pathways. In the past 10 years, detailed structural and functional data have been reported for several ZnF–protein interactions showing unequivocally that several classes of ZnFs (including classical ZnFs) can function as protein recognition motifs. As it was previously mentioned (Table I.1), some ZnF domains as UBZ, NFZ and ZnF A20 are able to interact with ubiquitin, thus classified as ubiquitin-binding domains.

ZnF domains are often found as multiple copies in a single protein. The number of ZnF domains present within a protein together with their amino acid sequence compositions and the length of the linkers between them influence the binding properties of the protein. ZnF domains are often found in clusters, where fingers can have different binding specificities. There are many super families of ZnF motifs, varying in both sequence and structure. They display considerable versatility in binding modes, even amongst members of the same class (e.g. some bind DNA, others bind proteins), suggesting that ZnF motifs are scaffolds that have evolved toward specialized functions. For example, proteins containing ZnF domains have been found to participate in gene transcription, translation, mRNA trafficking, cytoskeleton organization, epithelial development, cell adhesion, protein folding, chromatin remodeling and zinc sensing (Laity J.H. et al., 2001). Zinc-binding motifs are stable structures, and they rarely undergo conformational changes upon binding their target.

In general, ZnFs maintain their 3D structure by coordinating Zn^{2+} ions through Cys and/or His residues, and occasionally Asp side chains or water molecules (Alberts I.L. et al., 1998; Matthews J.M and Sunde M., 2002; Zhang J. et al., 2012). The number and order of these residues were used to classify the different types of ZnF (e.g. Cys₂His₂ like ZnF of TAX1BP1, Cys₂Cys₂ like ZnF of A20, Cys₂HisCys, like ZnF of NEMO). Thus, Zn^{2+} atoms play a structural role in ZnF domains, and despite that in aqueous solution Zn^{2+} forms an octahedral structure with water, when is coordinated by amino acids, the tetrahedral arrangement is the most often observed (82% of the Zn^{2+} sites are 4-coordinated, 14% are 5-coordinated and 4% are 6-coordinated) (Zhang J. et al., 2012).

Based on the structural properties in the vicinity of the zinc-binding site, the available ZnF structures were classified into eight separate fold groups (Table I.2) (Krishna S.S. et al., 2003). This classification helps in understanding the relationship between the structure and function of these domains.

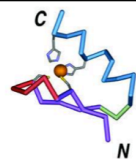
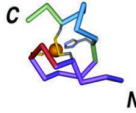
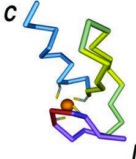
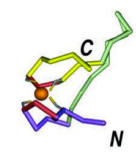
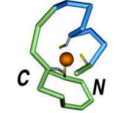
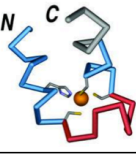
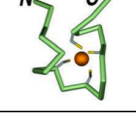
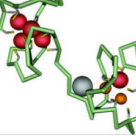
ZnF group	Ligand position	Structure
Classical Cys ₂ His ₂	Two ligands from a knuckle or β-sheet and two more from the C-terminus of the helix	
Gag Knuckle	Two ligands from a knuckle and two more from a short helix or loop	
Treble clef	Two ligands from a knuckle and two more from the N-terminus of a helix	
Zinc ribbon	Two ligands each from two knuckles	
Zn ₂ /Cys ₆	Two ligands from the N-terminus of the helix and two more from a loop	
TAZ2-domain-like	Two ligands each from the termini of two helices	
Zinc binding loops	Four ligands from a loop	
Metallothionein	Cysteine rich metal binding loop	

Table I.2. Structural classification of zinc finger domains. The ZnF domains have been divided into eight separate fold groups based on the structural properties around the zinc-binding site (zinc colored in orange). Color code: in cyan helices, in red the zinc knuckle connecting the two strands of the β-hairpin, in purple the primary β-strands that neighbor the knuckle, and other strands in yellow, in green loops, and in grey other parts of the structure that do not belong to the zinc-binding region. Table reprinted from Krishna S.S. et al., 2003.

1.7.1 Classical ZnF

The classical zinc finger proteins are those containing the Cys₂His₂ pattern, which represents one of the most abundant structural motifs involved in eukaryotic nucleic acid binding (Miller J. et al., 1985; Klug A. et al., 1987; Evans R.M. et al., 1988; Payre F. and Vincent A., 1988; Berg J.M., 1988; Lee M.S. et al., 1989; Pavletich N.P. et al., 1991). The consensus sequence of the Cys₂His₂ zinc finger motif is Tyr/Phe-X-Cys-X₂-₄-Cys-X₃-Phe/Tyr-X₅-Leu-X₂-His-X₃₋₄-His, where X is a variable amino acid (Cox E.H. et al., 2000). The Cys₂His₂ motif comprises 28 - 30 amino acid residues, including two conserved cysteine and two conserved histidine residues holding a zinc ion (Krishna S.S. et al., 2003), crucial to maintain the stability of the tertiary structure of most Cys₂His₂ domains. This tetrahedral coordination is very stable with the bond angles of the tetrahedron very close to the ideal values of 109.5° (Alberts I.L. et al., 1998).

The three dimensional structure of the Cys₂His₂ zinc finger motif is shown in figure I.12. The motif consists of two short antiparallel β-strands followed by an α-helix that forms a left-handed ββ_α-unit where the four conserved (Cys₂His₂) residues serve to hold one end of the antiparallel β-sheet to one end of the α-helix. The α-helix of each domain is often called “recognition helix” as it mediates DNA binding, through non-covalent interactions between three or four exposed residues of the helix and three or four adjacent bases within the DNA major groove (Klug A., 1999; Wolfe S.A. et al., 2000; Guillièrè F. et al., 2013).

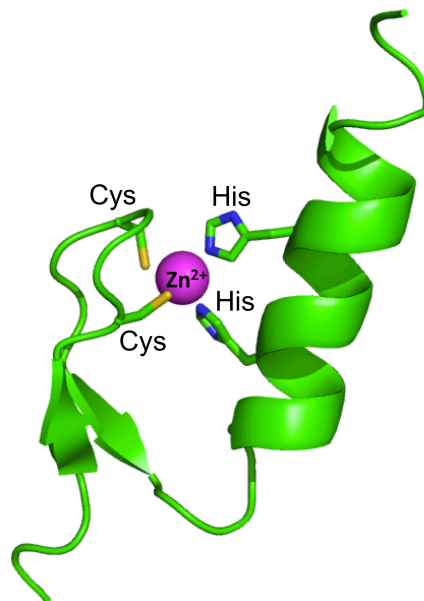


Figure I.12. Example of the structure of a classical Cys₂His₂ ZnF domain, corresponding to the ubiquitin-binding zinc finger of the human DNA Y-polymerase eta (PDB code 2150). The side chains of the His and Cys involved in binding are represented in sticks and colored by elements (S in yellow and N in blue) and the Zn²⁺ atom is represented as a pink sphere.

Despite classical ZnF domains are especially known for their role in sequence-specific DNA-binding (Wolfe S.A. et al., 2000), their functions extend beyond DNA binding to lipids, RNA and proteins (Fox A.H. et al., 1999; Polekhina G. et al., 2002). Usually, series of three or more classical ZnF are needed to recognize DNA motifs in target

genes (Matthews J.M and Sunde M., 2002) and the variation in zinc finger number and spacing may have an effect on the binding properties of the Cys₂His₂ zinc finger proteins (Iuchi S., 2001).

However, a single classical ZnF domain can mediate protein-protein interactions (Matthews J.M and Sunde M., 2002). That is the case of the transcriptional regulator FOG-1 (Friend of GATA) that contains nine ZnFs (four of them are Cys₂His₂ and five are a variant that present a Cys in the place of the final His) with most of them being able to bind independently to the erythroid transcription factor GATA-1 (Liew C.K. et al., 2000). In general, the interface of ZnF domains involved in protein recognition seems to differ from the one used in DNA-binding (Gamsjaeger R. et al., 2007).

The two ubiquitin-binding zinc finger domains located in tandem in the C-terminal region of the adaptor protein TAX1BP1 have been classified within the group of classical Cys₂His₂ zinc finger domains. However, their high-resolution structures in solution and their interacting mechanisms with ubiquitin are still unknown. The second zinc finger domain of TAX1BP1 (UBZ2) shares sequential similarities with the ubiquitin-interacting region of the ubiquitin-binding zinc finger of the human DNA Y-polymerase eta (DNA pol eta) (Bienko M. et al., 2005; Bomar M.G. et al., 2007) (Fig. I.12). The structure and mode of interaction of the UBZ domain of DNA pol eta is particularly different from other ubiquitin-associating zinc fingers, such as NZF, Znf-UBP, and RUZ domains (Alam S.L. et al., 2004; Lee S. et al., 2006; Penengo L. et al., 2006; Reyes-Turcu R.E. et al., 2006).

The structural and thermodynamic characterization of the interaction of UBZ1+2 with ubiquitin will set the basis to understand how TAX1BP1 acts as an ubiquitin-binding adaptor participating in the formation of the ubiquitin-editing complex with the A20 enzyme, which is involved in the negative regulation of the transcription factor NF- κ B pathway to limit inflammatory gene activation.

CHAPTER II

AIMS OF THE THESIS

.

Let me tell you the secret that has led me to my goal: my strength lies solely in my tenacity. Louis Pasteur.

Chapter II. AIMS OF THE THESIS

Proteins are key players in all biological events that take place within and between cells. And yet, proteins seldom act in isolation and often accomplish their function as part of large molecular machines whose action is co-ordinated through intricate regulatory networks of transient protein-protein interactions.

The presence of multiple ligand-binding sites within a protein, offers the possibility of multiple simultaneous interactions, enabling significant enhancement or reduction of ligand binding as compared with the binding to a single domain or motif. Therefore, to understand the molecular pathway in which adaptor proteins are involved, it is crucial to provide information at a structural level about the molecular details of the interactions with their targets.

The work presented in this thesis is focused on the study of modular domains of adaptor proteins. Modular domains play a central role in the assembly of multi-protein complexes and their specificity and affinity for particular targets determine the efficiency and outcome of cellular signalling. The detailed structural and thermodynamic characterization of the different types of intermolecular interactions between a adaptor protein and its targets might give us a better understanding of the signalling pathways in which these adaptor proteins are involved.

These small domains use to appear in tandem within the adaptor protein, often present similar affinities and specificities for a particular target, and the interactions with their targets are generally moderate to weak (micromolar range). These characteristics can promote cooperative and re-binding effects as well as versatile recognition modes, crucial for the proper functioning of adaptor proteins.

This thesis is divided in two main sections. The first one describes the characterization of the interaction of the N-terminal SH3 domains (SH3A) of the CIN85/CD2AP family of adaptor proteins with atypical proline-rich sequences present in CD2 and Cbl-b proteins. The versatile interaction of these SH3A domains is crucial for the protein clustering involved in endocytosis and regulation of the receptor tyrosine, cell adhesion or T-cell signaling pathways.

The second part of this thesis presents the determination of the high-resolution structure of the UBZ1+2 tandem of TAX1BP1 in solution, and the description at structural level of the, until present uncharacterized, interfaces involved in the binding of the UBZ domains of TAX1BP1 with ubiquitin. The structural determination of the interaction mechanism between UBZ1+2 and ubiquitin is crucial for the understanding of the role of TAX1BP1 within the negative regulation of the NF- κ B transcription factor.

Despite that CIN85/CD2AP and TAX1BP1 are non-related adaptor proteins, their modular domains, SH3 and UBZ respectively, share some interesting features: they are both ubiquitin-binding domains, small, appearing in tandem and interacting with their targets with moderate affinity. The latter characteristic probably enables quick dissociation from one site and subsequently recapturing by a neighboring site in the

same target protein, typical for modular domains present in adaptor proteins.

The interaction of ubiquitin with several SH3 domains of CD2AP and CIN85 were already studied in detail within our group. So we took advantage of the main difference between SH3 and UBZ domains. SH3 domains are specially known to interact with different proline-rich peptides; in the case of the SH3 domains of CD2AP and CIN85, the most representative examples are the peptides CD2 and Cbl-b. In contrast, UBZ domains are known to interact with proteins, like ubiquitin or WW domains. This gave us the possibility to study the interactions of modular domains in two different ways: protein-peptide and protein-protein interaction.

Both ways have served as a model to understand the role of modular domains in signal transduction pathways with special emphasis on the structural aspects that drive their biological functions, either in their free form or in complex with their natural targets.

The description at structural level on how these interactions take place in solution is a challenge that requires the use of multiple biophysical and biochemical techniques. Structural studies of moderate-weak complexes are compromised by serious practical limitations. In order to obtain low- and/or high-resolution structures of low affinity complexes, complementary methodologies are required. For this reason the study of both families of proteins have been carried out with the use of three complementary biophysical techniques in solution, ITC, NMR and SAXS.

With these premises, the main objectives of the presented study are:

- Characterize the interaction of the N-terminal domain (SH3A) of the CIN85/CD2AP family of adaptor protein with atypical proline-rich targets. Description at structural level of the differences between both members of the family, in order to shed light on the clustering mechanism in which they are involved, that till now remains unclear.
- Determination of the high-resolution structure of the UBZ1+2 zinc finger domains of TAX1BP1 in solution.
- Structural and thermodynamic characterization of the interaction of the UBZ1+2 tandem of TAX1BP1 with ubiquitin and compare it with the interaction of the isolated UBZ1 and UBZ2 domains, to identify possible cooperativity effects between both domains.
- Determination of a low-resolution model of the complex formed between the UBZ1+2 tandem of TAX1BP1 and ubiquitin by the use of NMR, Paramagnetic Relaxation Enhancements (PRE) and SAXS. Finally, it would be interesting to establish if UBZ1+2 can interact with different polyubiquitin linkages (Lys48-, Lys63- and linear polyubiquitin chains), which might have important physiological consequences.

CHAPTER III

MATERIALS AND METHODS

.

A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him as though they were fairy tales. Marie Curie.

Chapter III. MATERIALS AND METHODS

3.1 Protein expression and purification

The work presented in this thesis is based on the study of the modular domains of three different adaptor proteins (CD2AP, CIN85 and TAXBP1). To carry out the characterization of the interactions of these modular domains with their natural targets, it was necessary to prepare them in sufficient amount to achieve the requirements of the different biophysics techniques that have been used.

Seven different proteins were purified: SH3A of CD2AP, SH3A of CIN85, UBZ1+2 and ubiquitin wild type (WT) plus three Cys mutants of ubiquitin (E51C, D39C and D32C). All of them, except the mutants of ubiquitin, were also produced ^{15}N - and ^{15}N - ^{13}C isotopically enriched for NMR experiments.

As starting point, the plasmids used in protein expression are detailed in the Table III.1. All constructs were checked by DNA sequencing.

Vector	Antibiotic Resistance	Tag	Cleavage site	Gene inserted in the vector
pET-21a	Amp	No	No	CD2AP-SH3A ₂₋₆₂
pET-21b	Amp	No	No	Ubiquitin ₂₋₇₆ and mutants (D32C, D39C and E51C)
pET-28a	Kan	N-term His ₆	Protease preScission	CIN85-SH3A ₂₋₅₈ His-tag ₆ Tax1BP1-UBZ1+2 ₇₂₅₋₇₈₉ His-tag ₆

Table III.1. *E. coli* expression vectors

The pET-21b-based constructs encoding the ubiquitin Cys mutants (D32C, D39C and E51C) were purchased from GenScript (site-directed mutagenesis service, Piscataway, USA). The vector pET-21a containing the gene encoding the CD2AP-SH3A and the vectors pET-28a containing the CIN85-SH3A and TAX1BP1-UBZ1+2 coding sequences were kindly provided by Dr. Bravo (IBV-CSIC, Valencia, Spain).

The original plasmids were amplified in *Escherichia coli* (*E. coli*) Top10 or XL-10 strains and DNA was isolated using the QIAprep Spin miniprep kit (QIAGEN).

3.1.1 Transformation and Overexpression in *E. coli* cells

Chemically competent *E. coli* strains were transformed with plasmid DNA by heat shock. 100 μL of supercompetent cells (see Table III.2) were mixed with 10 ng of plasmid DNA and the sample incubated on ice for 15 min. A heat shock of 30 s at 42 °C was then applied followed by 5 min of incubation on ice. Cells were diluted with 1 mL

of LB medium and incubated at 37 °C for 45 - 60 min before plating on LB-Agar plates supplemented with the proper antibiotic (100 µg/mL Amp or 50 µg/mL Kan). Selected colonies were subsequently grown in the proper medium for protein expression.

Table III.2 summarizes the different *E. coli* strains used to perform the transformations (for convenience, the strains Top 10 and XL-10 used for DNA amplification and not protein expression, have been added to the same table).

Strain	comments	supplier
Top10 and XL-10	Used for DNA amplification. High-efficiency cloning and plasmid propagation.	Invitrogen
BL21 (DE3)	Used for protein expression T7 promoter driven expression is repressed until IPTG induction	Invitrogen
Rosetta (DE3)	Rosetta host strains are BL21 lacZY designed to enhance the expression of eukaryotic proteins containing codons rarely used in <i>E. coli</i>	Novagen

Table III.2. *E. coli* strains

For the unlabeled proteins, the expression conditions are indicated as follows: SH3A domain of CD2AP was overexpressed without His-tag in *E. coli* Rosetta (DE3) cells at 37 °C in LB medium; SH3A of CIN85 His-tag₆ and TAX1BP1-UBZ1+2 His-tag₆ constructs were overexpressed in *E. coli* BL21 (DE3) cells and grown in 2xTY medium at 37 °C; Ubiquitin WT and mutants were overexpressed in *E. coli* BL21 (DE3) cells and grown in LB at 37°C. In all these cases, protein expression was induced with 1 mM IPTG at 37 °C during 18 h. The purification protocols of the unlabeled and ¹⁵N-labeled SH3A domain of CD2AP, the unlabeled SH3A domain of CIN85 and the UBZ1+2 domains of TAX1BP1 were already described (Jozic D. et al., 2005; Ortega-Roldan J.L. et al., 2007; Spinola M., Ph.D. thesis, 2010). Ubiquitin and its Cys mutants (D32C, D39C and E51C) were purified as described (Chapter V, section 5.2.2). ¹⁵N- and ¹⁵N-¹³C-labeled ubiquitin were expressed in M9 minimal media with ¹⁵NH₄Cl and ¹³C-enriched glucose (Spectra Stable Isotopes and Cortecnet) as the sole source of nitrogen and carbon, and purified using the same protocol.

However, for the expression and purification of the isotopically ¹⁵N -¹³C and ¹⁵N-labeled SH3A domain of CIN85 and the UBZ1+2 domains of TAX1BP1, it was necessary to optimize the expression in minimal medium and to set up a purification protocol.

To perform small-scale expression tests, different *E. coli* strains were transformed with the plasmids, and selected colonies were grown in different rich media (supplemented with 50 µg/mL Kan): **Lysogeny Broth** (LB) (10 g/L Tryptone, 5 g/L Yeast extract, 10 g/L NaCl, pH 7.2); **Two times Tryptone-Yeast extract** (2xTY) (16 g/L Tryptone, 10 g/L Yeast extract, 5 g/L NaCl, pH 7.2) and **Terrific Broth** (TB) (12 g/L Tryptone, 24 g/L Yeast extract, 0.4% Glycerol, 2.3 g/L KH₂PO₄, 12 g/L K₂HPO₄), and also directly in **M9 minimal media** (6 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 0.5 g/L NaCl) containing 1 g/L of ¹⁵NH₄Cl and 3 g/L of ¹³C-enriched glucose (Spectra Stable Isotopes and Cortecnet)

and supplemented with (amounts required per liter of medium): 1 mL of 1 M MgCl₄, 200 µL of 1 M CaCl₂, 1 mL of a vitamins mix (BME vitamins solution 100x from Sigma-Aldrich) and 10 mL of trace element mix 100x (6 g/L of FeSO₄, 1.2 g/L of MgCl₂, 0.8 g/L CoCl₂, 0.7 g/L of ZnSO₄, 0.3 g/L of CuCl₂, 0.08 g/L of H₃BO₃ and 0.5 g/L of EDTA). The M9 minimal media used to grow the cells containing the zinc finger domains (UBZ1+2) of TAX1BP1 was also supplemented with 50 µM ZnCl₂. All chemicals used for the preparation of buffers and media were of the purest grade available and purchased from either Sigma-Aldrich (St. Louis, MO, USA) or Merck (Darmstadt, Germany).

When the cell cultures grown directly in M9 minimal media reached the optical density at 600 nm (OD₆₀₀) of 0.6 - 0.8, expression was induced with different concentrations of Isopropyl β-D-1thiogalactopyranoside (IPTG), 0.2 - 1 mM. For the cultures grown in rich media (LB, 2xTY and TB), two washing steps before resuspending the cells in minimal medium were included. In these cases, when the OD₆₀₀ was 0.6 - 0.8, the cultures were centrifugated at 4,000 rpm (Eppendorf Micro Centrifuge 5430 R, for 50 mL falcon) at 4 °C during 10 min and the rich medium was removed. The cells were resuspended in isotopically enriched M9 minimal medium, in a proportion 4:1 (the cells contains four times the volume of the expression test, were resuspended in one time the volume in minimal medium), and after 1 hour of incubation at 37 °C, the expression was induced by adding different concentrations of IPTG, from 0.2 to 1 mM. After induction, cells were incubated at different temperatures (37 °C, 20 °C and 15 °C). Samples were collected every two hours and the expression was checked by SDS-PAGE analysis.

The best results from the expression tests were:

- For the SH3A domain of CIN85: induction with 1 mM IPTG and incubation at 37 °C with shaking at 120 rpm, during 18 hours.
- For the UBZ1+2 domains of TAX1BP1: induction with 0.2 mM IPTG and incubation at 20 °C with shaking at 120 rpm, during 24 hours.

The conditions were adopted for the preparation-scale expression cultures (before induction with IPTG, cells grown in 4 liters of 2xTY medium were resuspended in 1 liter of enriched M9 minimal medium after two washing steps).

3.1.2 Purification protocol of the ¹⁵N-¹³C isotopically labeled SH3A domain of CIN85

Cells from expression cultures were recovered by centrifugation at 5,000 rpm (BECKMAN COULTER, J-LITER JLA-8.1000 rotor) for 20 min at 4 °C, resuspended in the lysis buffer without urea (see Table III.3) containing one tablet of protease inhibitor cocktail (cOmplete, EDTA-free, Roche) and lysed using a cell cracker (HAIVA of CONSTANTSYSTEMS Ltd.) at a pressure of 20 kpsi (1350 bar). However, after centrifuging the lysate at 20,000 rpm (BECKMAN COULTER, JA-20 rotor) during 30 min at 4 °C, the protein was found to be in the insoluble fraction (inclusion bodies) and needed therefore to be purified in denaturing conditions (buffers containing 8 M urea, Table III.3) and refolded afterwards.

The insoluble protein fraction was solubilized in lysis buffer containing 8 M urea (Table III.3). To break the inclusion bodies, it was necessary to perform three sonication cycles (MSE Ultrasonic disintegrator Soniprep-150) of 45 seconds at 60% power, with 45

seconds intervals (between each pulse) and then to stir up the sample overnight at 37 °C. When the inclusion bodies were totally redissolved, the sample was centrifuged at 20,000 rpm (BECKMAN COULTER, JA-20 rotor) and the supernatant was loaded on a HisTrap nickel (Ni)-sepharose resin (GE Healthcare), equilibrated with the lysis buffer. The protein was eluted using an imidazole gradient (from 20 mM to 500 mM imidazole) in presence of 8 M urea. Afterwards, the protein was refolded by successive dialysis (G2 dialysis cassettes molecular weight cut-off 3.5 kDa, Thermo Scientific) in buffer A (see Table III.3) progressively depleted in urea (8 M, 3 M, 1.5 M and 0 M urea). Finally, the His-tag was cleaved by incubation with the enzyme Protease PreScission (GE Healthcare) (5 units/mg of protein) at room temperature during 16 h.

To separate the cleaved CIN85-SH3A from the uncleaved one, the protein was passed again through a Ni-chelating column (GE Healthcare), using the same buffer A and buffer B (see Table III.3) as before, but without urea. The sample was further purified on a Hiload 16/60 Superdex-75 gel filtration column (GE Healthcare) in 50 mM buffer sodium phosphate (NaPi) at pH 6.0 (see Table III.3). Sample concentration was determined spectrophotometrically (UV-VIS spectrophotometer Cary 100Bio, Agilent) at 280 nm, using an extinction coefficient (ϵ_{280}) of 12,490 M⁻¹cm⁻¹, determined by the ProtParam tool (Expasy). The purified domain was concentrated to 1 mM final concentration on Amicon Ultra centrifugal units (Amicon Ultra-4, PLBC Ultracel-3 Membrane, 3 kDa, Milipore) and stored at -20 °C after fast freezing in liquid nitrogen.

Purification Buffers			
Lysis buffer	Nickel column		Superdex buffer
50 mM Tris-HCl pH 8.0, 400 mM NaCl, 20 mM imidazole with 8 M urea.	Wash buffer (A) 50 mM Tris-HCl pH 8.0, 400 mM NaCl, 20 mM imidazole with 8 M urea.	Elution buffer (B) 50 mM Tris-HCl pH 8.0, 400 mM NaCl, 500 mM imidazole with 8 M urea.	50 mM NaPi buffer pH 6.0.

Table III.3. Buffers used during the purification of the ¹⁵N-¹³C isotopically labeled SH3A domain of CIN85.

3.1.3 Purification protocol of ¹⁵N-¹³C isotopically labeled UBZ1+2 domains of TAX1BP1

The UBZ1+2 tandem of TAX1BP1 is composed of two zinc fingers and each presents two Cys and two His residues that coordinate a zinc (Zn²⁺) atom, which is crucial to maintain the tertiary structure of the domains. To avoid precipitation and formation of disulfide bonds, the lysis buffer was supplemented with ZnCl₂ and a reductor agent, like TCEP [Tris(2-carboxyethyl)phosphine].

Cells were recovered from the expression cultures by centrifugation at 4,000 rpm (BECKMAN COULTER, J-LITER JLA-8.1000 rotor) for 20 min at 4 °C, resuspended in lysis buffer (see Table III.4) containing one tablet of protease inhibitor cocktail (cOmplete, EDTA-free, Roche) and disrupted using a cell cracker (HAIVA of CONSTANTSYSTEMS Ltd.) at a pressure of 20 kpsi (1350 bar).

After centrifugation at 20,000 rpm (BECKMAN COULTER, JA-20 rotor) during 30 min at 4 °C, the supernatant was loaded on a 5 mL HisTrap Ni-chelating column (GE Healthcare) equilibrated with buffer A (Table III.4). The protein was eluted with a gradient from 20 to 500 mM imidazole (buffers A and B in Table III.4). Fractions containing UBZ1+2 were incubated overnight at room temperature with PreScission protease (GE Healthcare), 10 units/mg of protein, to remove the His-tag.

The cleaved UBZ1+2 was further purified on a Hiload 16/60 Superdex-75 gel filtration column (GE Healthcare) in 50 mM Tris-HCl pH 7.5, 200 mM NaCl and 1 mM TCEP. Sample concentration was measured in an UV-VIS spectrophotometer (Cary 100Bio, Agilent) using an $\epsilon_{280} = 8,480 \text{ M}^{-1}\text{cm}^{-1}$ determined by the ProtParam tool (ExPasy). After purification, two equivalents of ZnCl_2 were added to the sample, always under reducing conditions. The sample was then concentrated on Amicon Ultra centrifugal units (Amicon Ultra-4, PLBC Ultracel-3 Membrane, 3 kDa, Milipore); flash-frozen in liquid nitrogen and stored at -20 °C. Proper folding of the UBZ1+2 tandem was checked by NMR using 1D ^1H and 2D ^1H - ^{15}N -HSQC spectra (see Chapter III, Section 3.3) and circular dichroism (CD).

Purification Buffers			
Lysis buffer	Nickel column		Superdex buffer
50 mM Tris-HCl pH 7.5, 200 mM NaCl, 20 mM Imidazole, 50 μM ZnCl_2 , 1 mM TCEP, 0,5 % Tween-20 (v/v)	Wash buffer (A) 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 20 mM Imidazole, 1 mM TCEP	Elution buffer (B) 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 500 mM Imidazole, 1 mM TCEP	50 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM TCEP

Table III.4. Buffers used during the purification of the ^{15}N - ^{13}C isotopically labeled UBZ1+2 of TAX1BP1.

3.2 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a powerful technique for the study of the thermodynamics of macromolecular interactions. In an ITC experiment, two reactants are titrated against each other and the extent of binding is determined by direct measurement of the heat exchange with the environment. It is the only technique where the binding or dissociation constant ($K_d = 1/K_b$), Gibbs free energy of binding (ΔG_b), enthalpy (ΔH_b) and entropy (ΔS_b) can be directly and accurately determined from a single experiment (Leavitt S. and Freire E., 2001), using the following equations (Eq.1 and Eq.2, where R is the universal gas constant, $1.987 \text{ cal}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$, and T is the absolute temperature in Kelvins) that provide the most relevant thermodynamic relationships:

$$\Delta G_b = -RT \ln K_b \quad \text{Eq.1}$$

$$\Delta S_b = \frac{\Delta H_b - \Delta G_b}{T} \quad \text{Eq.2}$$

To provide a full description of the energetics governing molecular interactions, it is also possible to determine the heat capacity change (ΔC_p), which is defined as the temperature derivative (at constant pressure) of the enthalpy (Eq.3), by performing the

same experiment at several temperatures. Indeed, the slope of the enthalpy versus temperature plot gives the heat capacity of the reaction. ΔC_p has been shown to originate from surface dehydration upon binding or hydration upon dissociation and, to a less extent, from the difference in vibrational modes between the complex and the free species (Gomez J. et al., 1995; Murphy K.P. and Freire E., 1992).

$$\Delta C_p = \left(\frac{\Delta H_b}{\Delta T} \right)_P \quad \text{Eq.3}$$

Furthermore, ITC allows an accurate determination of the stoichiometry of the interaction from the titration equivalence point (Fig. III.1C). Therefore, ITC is increasingly used in the analysis of systems involving multiple binding events, such as the formation of multiprotein complexes (Rickert M. et al., 2004) or the binding of multivalent ligands (Dam T.K. et al., 2000), which can be influenced by allosteric or cooperative effects. Thanks to these facts, together with other factors like ITC experiments are non-destructive and non-invasive, the technique has found widespread applicability in the study of biological systems (Velazquez-Campoy A. and Freire E., 2005).

3.2.1 The ITC experiment: theory and practice

In a titration experiment, one binding partner (called ligand, L, or “titrant”) in the syringe is added in small aliquots to a solution containing the other binding partner (called macromolecule, M, or “titrand”) in the calorimeter cell (Jelesarov I. and Bosshard H.R., 1999). At the beginning of the experiment the calorimetric cell is filled with the macromolecule, with an effective volume (V_c) that is sensed calorimetrically. During this thesis, the ITC experiments have been carried out on a VP-ITC or on a VP-ITC200 instrument (MicroCal Inc., GE Healthcare), which cell volumes are 1.4 mL and 200 μ L, respectively.

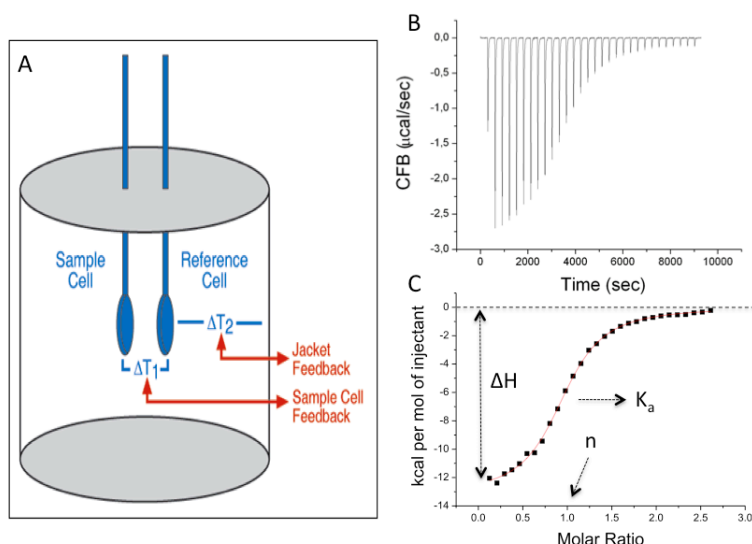


Figure III.1. *A. Diagram of the sample cell, where the ligand is injected through a syringe, and the reference cell of an ITC. B. Typical ITC exothermic thermogram obtained during the titration of the CD2AP-SH3A domain with the peptide CD2 at 20 °C. The area underneath each injection peak is equal to the heat released for that injection. C. Binding isotherm resulted from the plot of the integrated heats (black dots) against the molar ratio between the ligand added*

and the macromolecule in the cell. A model of n independent and identical binding sites was used to fit the data (section 3.2.2). The solid red line is the calculated curve using the best fitting parameters. The arrows indicate from where are determined the principal parameters.

A typical experiment is shown in Figure III.1, panel B. Upon titration of a ligand into the sample cell, heat is either released (exothermic) or absorbed (endothermic), and the heat is measured from the amount of power (μcal per second) required to maintain the temperature balance between the sample and reference cells. MicroCal's ITC systems use a cell feedback network (CFB) to differentially measure and compensate for heat produced or absorbed between the sample and reference cell (Fig. III.1). The heat after each injection is derived by calculating the area under each peak. The size of the heat event is directly proportional to the amount of binding that occurs. As the titration progresses, the macromolecule becomes saturated with the ligand and the magnitudes of the peaks decrease until the peak size only reflects dilution and mechanical effects, resulting in a classical sigmoidal curve (Fig III.1B). The successful extraction of thermodynamic parameters relies on the use of non-linear least squares curve fitting while employing an appropriate model that describes the interaction under study.

The setup of an ITC experiment is largely dependent on the expected binding affinity and the heat effect of the interaction. The shape of the binding curve is dependent on the product of the binding constant K_b (in M^{-1}) and the molar concentration of the macromolecule M_T being titrated (Wiseman T. et al., 1989) $c = K_b M_T$. Isotherms can only be accurately deconvoluted for c value ranging from approximately 1 to 1,000 (Wiseman T. et al., 1989).

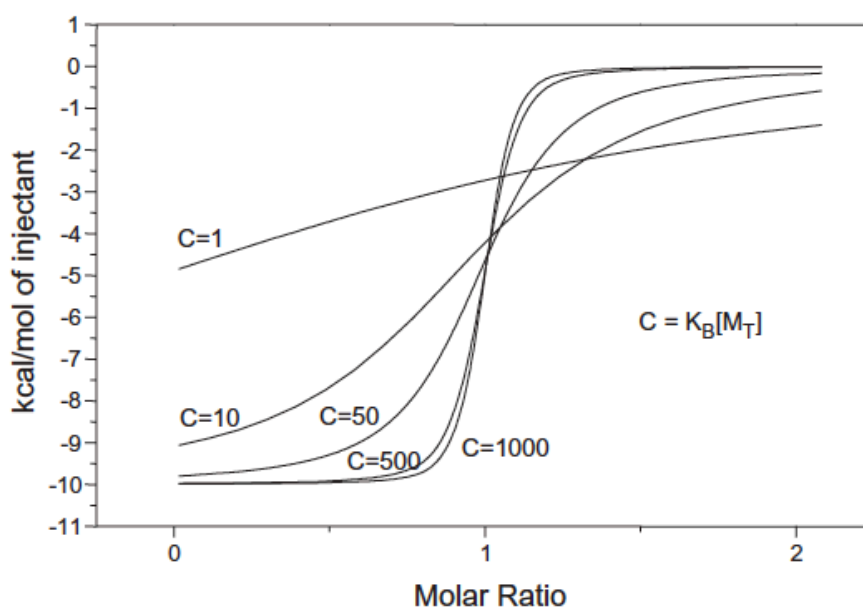


Figure III.2. The graph shows the effect of c value on the binding isotherm. High c values result in titration curves that are too steep to resolve K_b accurately (however, n and the ΔH are well resolved, whereas low c values result in shallow titration curves from where the three parameters (n , K_b and ΔH) are poorly resolved. Figure reprinted from Perozzo R. et al., 2004.

As shown in figure III.2 the shape of the binding isotherm, is crucial for the determination of the binding constant. At high values of c ($c > 500$), the shape of the curve approaches a step function and becomes increasingly insensitive to changes in K_b . At low c values ($c < 10$), the binding curve forms a horizontal trace that again yields very little information about K_b . Experience shows that conditions should be chosen to have a c -value in the range of 20 - 100 for an accurate determination of K_b (Wiseman T. et al., 1989; Myszka D.G. et al., 2003). For low-affinity systems, it is not always possible to achieve c values in the preferred range due to limited receptor and/or ligand solubility. In these cases, the ITC experimental window can sometimes be extended to lower c values if n is accurately known and fixed (Turnbull W.B. and Daranas A.H., 2003).

Taking this into account, the first step before running an ITC experiment is to determine the concentrations of the macromolecule and ligand in solution that should be used. To simultaneously determine K_b , ΔH and n for a binding event, a complete binding isotherm must be recorded and the titration experiment should be designed to reach the complete saturation of the binding site at the end of the experiment.

Sometimes, in order to fully resolve the binding, especially in the case of protein complexes where more than one binding event is taking place, is necessary to perform a reverse titration, where the “titrant” and the “titrand” are reversed. Reverse titrations allow to check the stoichiometry and the suitability of the binding model (Velazquez-Campoy A. and Freire E., 2006). For 1:1 biomolecular reactions it is expected that the measured thermodynamic parameters of the direct and reverse experiments are invariant.

To design the ITC experiments that were carried out with the CD2AP-SH3A, CIN85-SH3A and TAX1BP1-UBZ1+2 domains and their respective ligands, we used a simulation script implemented in MATHCAD (PTC MATHCAD) and designed by Dr. Obdulio López Mayorga (Department of Physical Chemistry, UGR), to find the optimized experimental conditions (macromolecule and ligand concentration, and volume profile of the injections). Based on these simulations it was possible to obtain isotherms which shapes could be accurately fitted by non-linear regression analysis

The time interval between successive injections is another parameter to take into account, and depends on the response time of the calorimeter and the kinetics of the reaction. During this interval, the DT (Differential Temperature between the adiabatic jacket and the calorimeter cells) signal return to the baseline after an injection peak deflection. An interval of 4 - 6 min between injections was used for all the ITC experiments performed during this thesis.

The selection of the buffer for an ITC experiment is an important aspect. We have taken into account three facts: *i.* the buffer should have good buffering capacity at the pH studied, *ii.* the protein and ligand should be stable in that buffer, and *iii.* the buffer should have a low enthalpy of ionization in order to minimize artifactual heats. It is moreover crucial that the protein and ligand are in identical buffers, otherwise large heats of dilution could mask the event to observe. All the ITC experiments carried out during this thesis have been done at pH 6.0 in 50 mM cacodylate buffer, which has one of the lowest protonation enthalpies ($\Delta H_i = -0.47$ kcal/mol at 25 °C), and presents a good buffering capacity at pH 6.0. Some ITC experiments of the CD2AP-SH3A domain were also done in 50 mM MES buffer ($\Delta H_i = 3.71$ kcal/mol at 25 °C), to calculate the

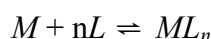
contribution of de-solvation to the binding process (Murphy, K.P. and Freire E., 1992; Gomez J. et al., 1995). Before every ITC experiment, the samples were filtered and degassed, the pH was carefully checked and the concentration was accurately measured.

Specific experimental conditions for the ITC experiments are provided in Chapter IV for the SH3A domain of CD2AP and CIN85, and in Chapter V for the experiments with the UBZ1+2 domains of TAX1BP1.

3.2.2 Mathematical analysis of the ITC experiment

The simplest and most widely used mathematical model to fit the ITC data considers the existence of a set of n identical and independent binding sites on the macromolecule. The mathematic bases of this model are explained below.

- Model of n independent and identical binding sites



The thermodynamic analysis of a binding process where the macromolecule (M) possess n identical an independent binding sites, presents a partition function (Z) that relates the addition of concentrations of all species to the concentration of an arbitrarily chosen references species (Eq.4), conveniently the free ligand (L) (Wyman and Gill, 1990).

$$Z = (1 + K[L])^n \quad \text{Eq.4}$$

where K is the apparent constant of the binding equilibrium, $[L]$ is the concentration of free ligand and n is the number of binding sites within the macromolecule.

The relation between the bound ligand $[L_b]$ and the total concentration of macromolecule $[M_T]$, is known like the union parameter, and is related with the partition function as in equation Eq.5

$$\bar{v} = \frac{[L_b]}{[M_T]} = \frac{\partial \ln Z}{\partial \ln [L]} \quad \text{Eq.5}$$

in this way, we can define the union parameter like:

$$\bar{v} = \frac{nK[L]}{1+K[L]} \quad \text{Eq.6}$$

The heat after each injection can be determined by:

$$q_i = \Delta H_b \left(\frac{\text{kcal}}{\text{mol} \cdot \text{L}} \right) \cdot \Delta(\text{mol} \cdot \text{L}_b) \quad \text{Eq.7}$$

where L_b represent the bound ligand and ΔH_b the enthalpic change per mol of bound injectant. We can express this equation to take into account that the volume of the cell, V_c , matches the initial sample volume:

$$q_i = \Delta H_b \cdot V_c \cdot ([L_b]_i - [L_b]_{i-1}) = \Delta H_b \cdot V_c \cdot (\bar{v}_i \cdot [M]_i - \bar{v}_{i-1} \cdot [M]_{i-1}) \quad \text{Eq.8}$$

During the titration, each injection drives a volume of liquid out of the calorimetric cell, for this reason the concentrations of the macromolecule and ligand should be slightly corrected after each injection. The real concentration of macromolecule (M_i) in the cell, can be calculated as:

$$[M]_i = [M]_{i-1} \cdot \frac{V_c - V_{in}}{V_c} \quad \text{Eq.9}$$

And the concentration of the ligand in the calorimetric cell as:

$$[L_T]_i = \frac{(V_c - V_{in}) \cdot [L_T]_{i-1} + V_{in} \cdot [L_0]}{V_c} \quad \text{Eq.10}$$

where V_{in} and $[L_0]$ are the injected volume and the concentration of the ligand in the syringe respectively.

In an ITC experiment the observable is the heat involved in the binding of a certain amount of ligand when injected to the protein in the cell. The heat accumulated after N injections can be expressed like:

$$Q = \sum_{i=1}^N q_i = V_c \cdot [M]_N \cdot \Delta H_b \cdot \bar{v}_N = V_c \cdot [M]_N \cdot \Delta H_b \cdot \frac{nK[L]}{1 + K[L]} \quad \text{Eq.11}$$

However the values of $[L]$ is unknown, and should be expressed based on the experimental variables $[L_T]$ and Q :

$$[L] = [L_T] - [L_b] = [L_T] - \frac{Q}{V_c \cdot \Delta H_b} \quad \text{Eq.12}$$

If we substitute the value of $[L]$ in the Eq.11, we obtain a second order equation which solution is:

$$Q = \frac{V_c \cdot \Delta H_b}{2K} \left[1 + K[L_T] + nK[M_T] - \sqrt{(1 + K[L_T] + nK[M_T])^2 - 4nK^2[M_T][L_T]} \right] \quad \text{Eq.13}$$

This equation relates the total heat accumulated, after N injections, with the experimental variables V_c , $[L_T]$ and $[M_T]$, and the parameters n , K and ΔH_b .

However, exists other way to treat the experimental heats. In this case, we can write Eq.12 as:

$$[L] = [L_T] - [L_b] = [L_T] - [M_T] \cdot \frac{nK[L]}{1 + K[L]} \quad \text{Eq.14}$$

Based on this, we can express $[L]$ in terms of n , K , $[L_T]$ and $[M_T]$:

$$[L] = \frac{1}{2K} \left[-(1 + nK[M_T] + K[L_T]) + \sqrt{(1 + nK[M_T] - K[L_T])^2 + 4K[L_T]} \right] \quad \text{Eq.15}$$

The $[L_T]$ derivated from Eq.13 gives us an expression considering the heat per mol of ligand injected during each injection:

$$\frac{1}{V_c} \cdot \frac{dQ}{d[L]_T} \approx \frac{1}{V_c} \cdot \frac{\Delta Q}{\Delta[L]_T} = \frac{\Delta H_b}{2} \left[1 - \frac{1 + K[M_T]\theta - nK[M_T]}{\sqrt{(1 + K[M_T]\theta - nK[M_T])^2 - 4nK^2[M_T]^2 \theta}} \right] \quad \text{Eq.16}$$

$$\text{where } \theta = \frac{[L_T]}{[M_T]} \quad \text{Eq.17}$$

Using these equations, the thermodynamic parameters n , K and ΔH_b , can be determined by non-linear regression analysis.

Thus, there are two different possible ways to treat the experimental heat. In this thesis, the last model (that considers the heat per mol of ligand during each injection) implemented by Dr. Mayorga (UGR) in the program Origin 7.0 (OriginLab Corps) was used to fit most of the experimental data.

However, due to the fact that the UBZ1+2 tandem is composed of two modular domains, able to bind ubiquitin in a different manner, we have also used the global data analysis platform SEDPHAT (Vistica J. et al., 2004; Balbo A. et al., 2005; Dam J. and Schuck P. 2005; Dam J. et al., 2005), which gives the option to fit the experimental data using different models. One of them considers the presence of two non-symmetric sites in the macromolecule and takes into account possible cooperativity effects (Houtman J.C. et al., 2007). That particular model gave a good fit of the isotherm obtained from the titration of the UBZ1+2 domains with ubiquitin. The software SEDPHAT, directly imports the multi-column tables with the integrated heat per injection from the MicroCal Origin software. In the Origin table the concentration and the normalized heats are corrected for the effects of dilution of the cell contents due to the ligand addition. To avoid the truncated error of the concentration values exported from MicroCal software, SEDPHAT can recalculate the cell concentration values based on the provided cell and syringe concentrations, which results in a smooth fitted curve.

3.3 Nuclear Magnetic Resonance (NMR)

3.3.1 Introduction to NMR

Nuclear magnetic resonance (NMR) spectroscopy is a high-resolution technique able to reveal macromolecule structures in solution. In contrast to other methods, NMR spectroscopy reports chemical properties by studying individual nuclei. This technique is based on the fact that certain atomic nuclei present a non-zero spin (S), which is associated with a non-zero magnetic moment (μ) by the equation $\mu = \gamma S$, being γ the gyromagnetic ratio. Only a limited number of isotopes display this property (^1H , ^2H , ^{13}C , ^{15}N , ^{17}O , ^{31}P , ^{35}Cl , and so on) and exhibit NMR absorption. For example, the spinning of a proton (^1H) generates a magnetic moment, which can take two orientations (or spin states called α and β), when an external magnetic field is applied. The energy difference between these states is proportional to the strength of the imposed magnetic field. The α state has a slightly lower energy and hence, is slightly more populated (by a factor of the order of 1.00001 in a typical experiment), because it is aligned with the field. A spinning proton in α state can be raised to an excited state (β state) by applying a radio-frequency pulse that provided the frequency corresponds to the energy difference between the α and β states. In these circumstances, the spin state will change from α to β (this is a single quantum coherence) and the resonance will be obtained. Applying such a pulse to a set of nuclear spins will simultaneously excite all the NMR transitions, tilting the magnetization vector away from its equilibrium position and aligning it along the external magnetic field. The magnetization vector precesses around the external magnetic field vector at the NMR frequency of the excited spins (Larmour frequency). This oscillating vector will induce a current in a detector coil, which creates an electrical signal. The oscillating signal decays as a function of time, due to the action of relaxation. This signal is the free induction decay (FID), which is a time-domain signal, and contains the sum of the vectors from all the excited spins. To transform this signal in a frequency domain NMR spectrum, a mathematical process known as Fourier Transformation is applied.

It is important to understand that different spins in a molecule do not precess independently. If one spin can have different energy states depending on which state a nearby spin is in, these two spins are coupled. There are two basic mechanisms by which information between coupled spins is exchanged:

- **J-coupling:** due to the indirect interaction between two nuclear spins mediated by the electrons that participate in the bonds connecting them. J coupling is also referred to as spin-spin or scalar coupling. Two nuclei in a molecule, which are connected by one, two and three bonds, can be coupled in the NMR spectra. The coupling is observed by a splitting of the NMR signal and contains information about bond lengths and angles.
- **Dipolar coupling:** results from each spin generating a magnetic field that is oriented parallel to the nuclear spin vector. Two spins that are close to each other in space experience each other's magnetic field. This leads to a slightly different effective magnetic field at one spin depending on the orientation of both magnetic dipoles. The strength of the dipole-dipole coupling depends strongly on the gyromagnetic ratios of the coupled spins and the spin-spin distance.

These properties reflect the chemical surroundings of the magnetic nucleus. The flow of electrons around a magnetic nucleus generates a small local magnetic field that opposes the applied field. The degree of such shielding depends on the surrounding electron density. Consequently, nuclei in different environments resonate at slightly different field strengths or radiation frequencies. The nuclei of the perturbed sample absorb electromagnetic radiation at a frequency that can be measured. The different frequencies, called *chemical shifts*, are expressed in fractional units δ (parts per million, or ppm) relative to the shifts of a standard compound added to the sample.

In a protein structure the individual residues are packed and experience specific chemical environments. The figure III.3 (^1H -NMR spectrum of the SH3A domain of CD2AP) shows the one-dimension (1D) NMR spectrum of a folded protein, which is more than the sum of the random coil spectra of the amino acid residues present in the protein, as it also reflects that nuclei in the folded form are subject to different microenvironments.

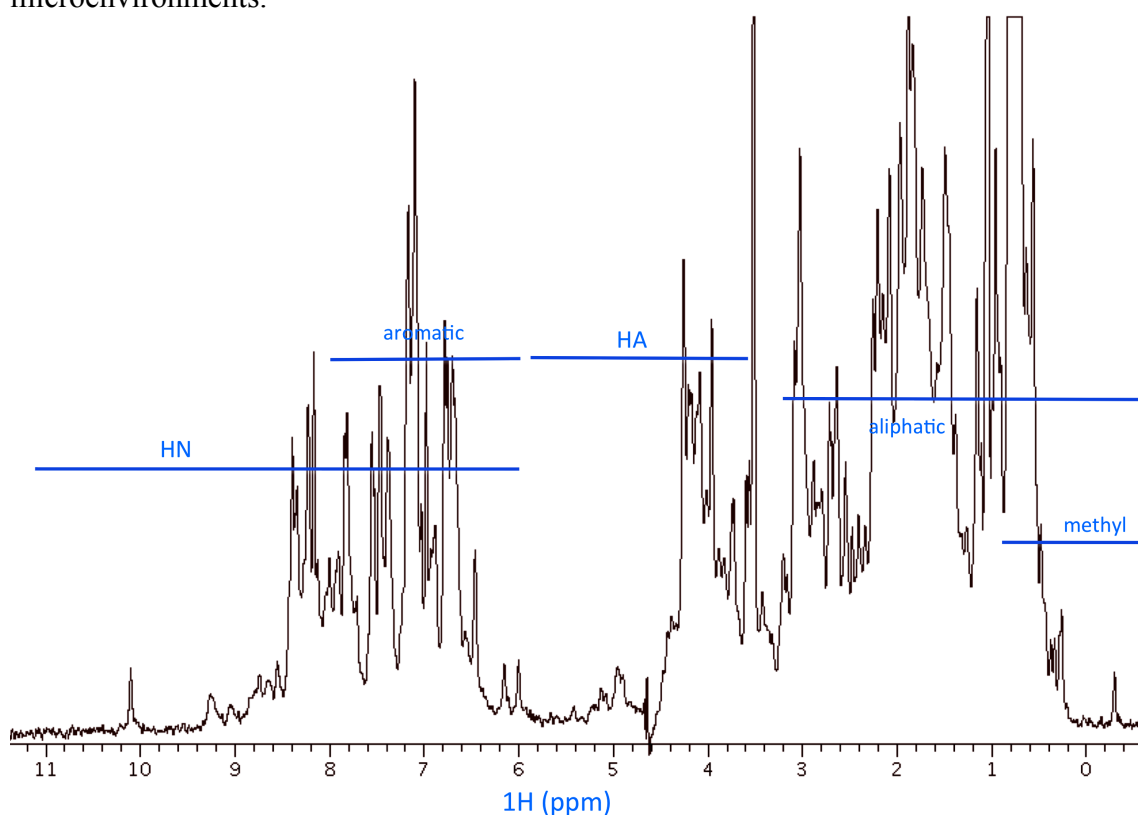


Figure III.3. One-dimension ^1H -NMR spectrum of the SH3A domain of CD2AP used as example to indicate the principal regions where the different types of protons are found.

The 1D spectra contain structural information of the protein, which is difficult to extract because of the signal overlapping. For this reason, it is necessary to resort to isotopic labeling of the proteins and carry out multidimensional heteronuclear NMR experiments.

3.3.2 Multidimensional NMR spectroscopy

The introduction of two, three and four-dimensional NMR experiments between the late 1980s and early 1990s (Clore G.M. and Gronenborn A.M., 1998) and the availability of ^{15}N - ^{13}C labeled proteins (McIntosh L.P. et al., 1987; Muchmore D.C. et al., 1989; McIntosh L.P. and Dahlquist F.W., 1990; Hansen A.P. et al., 1992; Lustbader J.W. et al., 1996) have been crucial for the assignment of proton, nitrogen and carbon chemical shifts of proteins and protein complexes with molecular weights in the range of 25 - 30 kDa (Fogh R.H. et al., 1994; Remerowski M.L. et al., 1994; Kwan A.H. et al., 2011), allowing thereby the determination of their solution structure (Clore G.M. and Gronenborn A.M., 1991; Weber C. et al., 1991; Fesik S.W. et al., 1991; Ikura M. et al., 1992, Garrett D.S. et al., 1997). Figure III.4 shows in a simplified way, the meaning of adding a new dimension in an NMR experiment.

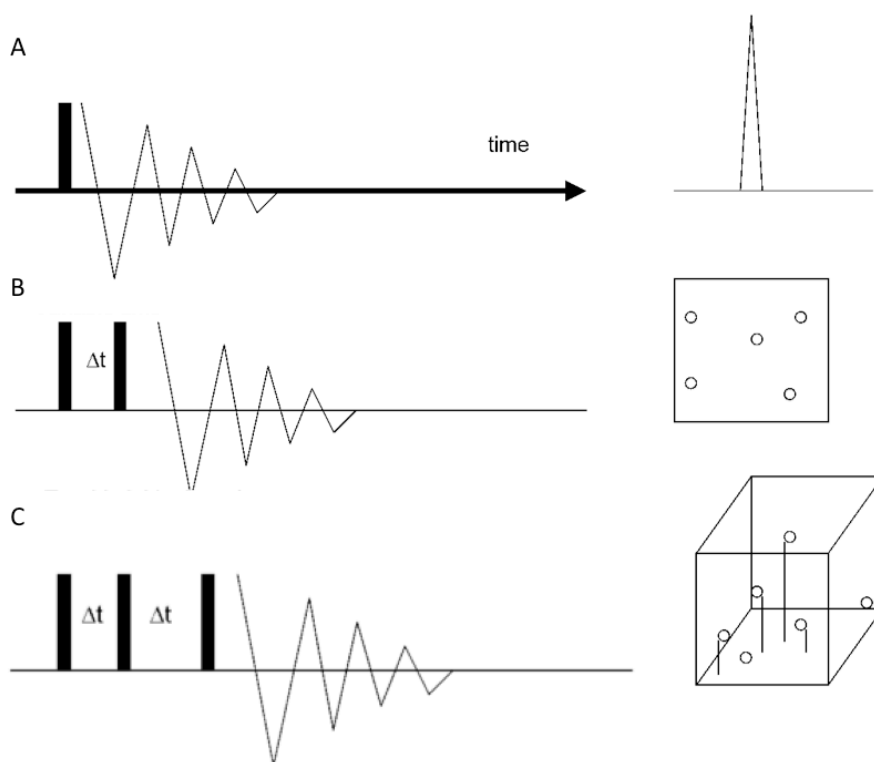


Figure III.4. Schematic representation of the FID and the result of the Fourier transform, of an NMR spectra in one-dimension (A), two-dimensions (B) and three dimensions (C). The introduction of a new dimension in the NMR spectrum is established by introducing an additional pulse and systematically increasing the time between the pulses. This is combined with the observation of one, or more nuclei (i.e. ^{13}C , ^{15}N). Figure reprinted from Poulsen F.M., 2002.

3.3.2.1 The ^1H - ^{15}N and ^1H - ^{13}C correlation spectra

In an isotopically enriched protein, every amino acid residue, except for proline residues, contains a backbone amide proton that is directly coupled to a ^{15}N atom. This one-bond ^1H - ^{15}N coupling is an important starting point for the heteronuclear NMR analysis of proteins. This coupling is recorded using an experiment called ^1H - ^{15}N HSQC (heteronuclear single quantum correlation). The HSQC is a 2D-NMR experiment that can also be applicable to other heteronuclei like a ^1H - ^{13}C system. Despite that the gyromagnetic ratios of ^{15}N and ^{13}C are markedly lower than of ^1H [gyromagnetic ratios ($10^7 \text{ rad T}^{-1} \text{ s}^{-1}$) of ^1H , ^{13}C and ^{15}N are 26.75, 6.72 and -2.71, respectively] they can be detected indirectly by polarization transfer through the ^1H during the HSQC experiments (Fig. III.5). This enhancement of sensitivity is possible because the HSQC employs two INEPT (insensitive nuclei enhanced by polarization transfer) steps.

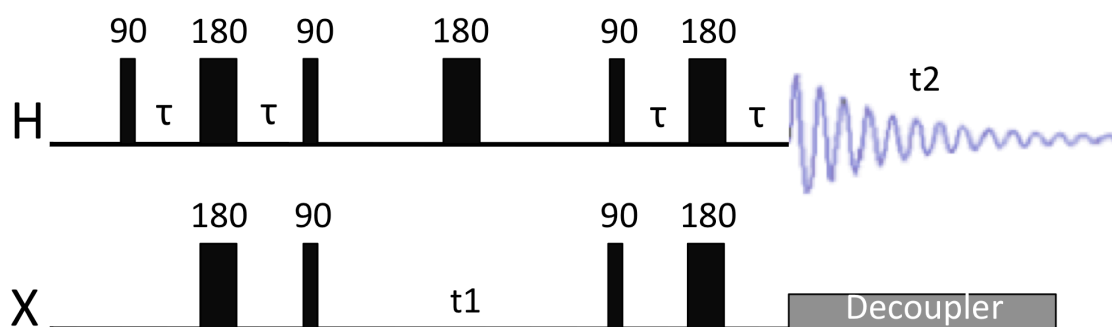


Figure III.5. Pulse sequence of an HSQC experiment involving ^1H and a heteronuclei ($X = ^{15}\text{N}$ or ^{13}C). $t1$ is chemical shift evolution of X (^{15}N or ^{13}C), and $t2$ chemical shift evolution of the one-bond coupled ^1H . Figure adapted from <http://www.chem.queensu.ca/facilities/nmr>.

The first 90° radiofrequency pulse excites the ^1H -NMR transitions and leaves them in a state of resonance (coherence). For the ^{15}N (or ^{13}C)-coupled protons, this coherence is subsequently transferred to the ^{15}N (or ^{13}C) resonance by the first INEPT step. This process of coherence transfer may be linked to a resonance energy transfer, here mediated through the particular combination of scalar coupling and pulses. The coherence is kept on the ^{15}N or ^{13}C spins during a time delay ($t1$) and is subsequently transferred back to the proton via a reverse INEPT step, after which the ^1H is finally detected (during a detection time, $t2$) and stored. A two-step phase cycle is needed to filter out the undesirable magnetization coming from ^1H attached to non-isotopically label heteronuclei (Mandal P.K. and Majumdar A., 2004). Only those ^1H resonances that have experienced a coherence transfer through the ^{15}N (or ^{13}C) resonance can be detected. Thus, in a mixture of an isotopically labeled and unlabeled protein, only the amide resonances of the labeled protein are observed.

An example of an HSQC spectrum of ^{15}N -labeled CIN85-SH3A domain is shown in figure III.6. In this spectrum the X-axis correlates to the ^1H chemical shifts ($\delta\text{H}^{\text{N}}$) and the Y-axis to ^{15}N chemical shifts (δN). The signals thus report the coupling between the H^{N} and N of each non-proline residue in CIN85-SH3A and appear at the corresponding $\delta\text{H}^{\text{N}}$, δN . The ^1H - ^{15}N HSQC spectrum is seen as a protein's "fingerprint" since it uniquely reflects its amino acid composition and structure. An analogous experiment (^1H - ^{13}C -HSQC) can be performed for detecting ^1H - ^{13}C correlations.

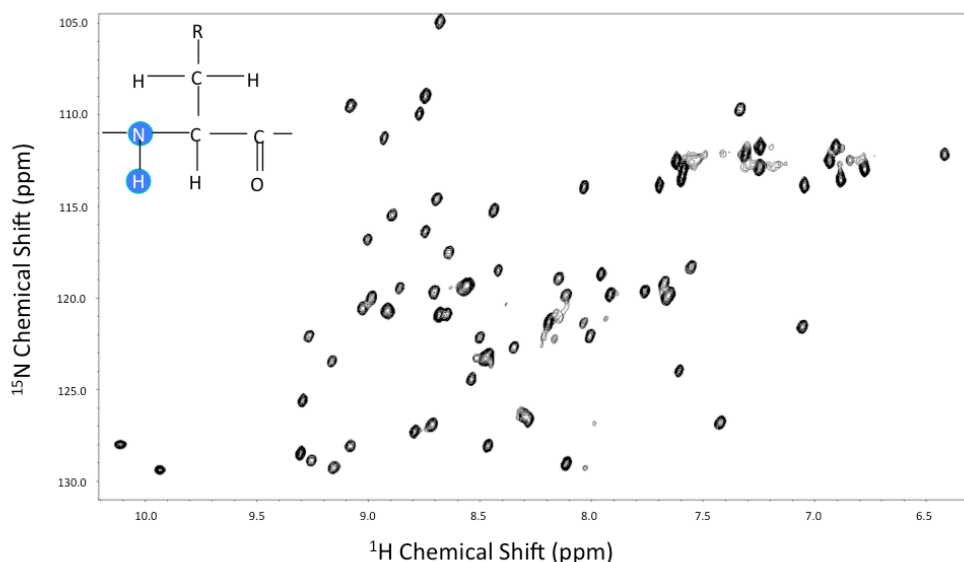


Figure III.6. ^1H - ^{15}N -HSQC spectrum of the SH3A domain of CIN85 and schematic representation of a single amino acid with the N and H atoms that are coupled in this experiment. Different to 1D spectra (Fig. III.3), in this 2D spectrum the peaks appear spread all over the window, and due to the small size of this globular protein (7 kDa), there is almost no signal-overlap, which strongly facilitates the chemical shift assignment process of the protein.

3.3.2.2 Chemical exchange by NMR

Chemical exchange refers to a dynamic process that exposes an NMR probe to at least two distinct chemical environments, or states, in a time-dependent manner. This is what happens when a protein is titrated with a ligand.

The simplest model of a dynamic process is the interconversion between two states (Fig. III.7), which can reflect a variety of dynamic processes, including exchange between the ligand-free and ligand-bound states of a protein, monomer and dimer or in general, conformations A and B. The two different states have NMR resonance frequencies ν_A and ν_B and the chemical shift difference is $\Delta\nu = |\nu_A - \nu_B|$. Importantly, the appearance of the NMR spectrum depends on the **populations** of each of the states, P_A and P_B , the relative values of the **exchange rate** $k_{\text{ex}} = k_A + k_B$ (Fig. III.7) and the **chemical shift difference** ($\Delta\nu$). k_{ex} quantifies the average number of stochastic exchange events per unit of time and is therefore expressed in s^{-1} . This allows a direct comparison of k_{ex} and $\Delta\nu$, since they have the same units.

NMR spectra affected by chemical exchange are divided into three distinct exchange regimes:

- Slow exchange ($k_{\text{ex}} \ll |\Delta\nu|$), characteristic of tight binding.
- Intermediate exchange ($k_{\text{ex}} \approx |\Delta\nu|$), characterized by signal broadening.
- Fast exchange ($k_{\text{ex}} \gg |\Delta\nu|$), characteristic of weak binding.

In the slow exchange regime (Fig. III.7), signals from both states are observed reflecting their distinct chemical shifts, intensities and linewidths. This is because there is not significant A and B interconversion during the detection period of the NMR experiment and the intensity of each peak directly reports on the population of that species. For either state (Fig. III.7), if the departure rate k is greater than $\approx 10 \text{ s}^{-1}$ the signal linewidth

will increase due to the depopulation occurring during the ≈ 100 ms detection period. This process directly contributes to the exponential decay of the FID intensity and can be quantified by an increase in transverse relaxation rate via $R_2^{\text{obs}} = R_2^0 + k$, where R_2^0 is the relaxation rate in the absence of exchange.

At the opposite, in the fast exchange regime, one signal is observed corresponding to the population-weighted chemical shift $\delta^{\text{obs}} = P_A\delta_A + P_B\delta_B$ and relaxation rate $R_2^{\text{obs}} = P_AR_{2A} + P_BR_{2B}$. This is because there is a rapid A and B interconversion and hence averaging during the detection period of the NMR experiment. The trajectories of the shifting resonances in fast exchange are informative. Indeed, linear two-dimensional trajectories indicate a single binding event, while curved trajectories of the resonances point out towards the implication of more than one binding event (Zuiderweg E. R., 2002).

In intermediate exchange, or coalescence, one signal is observed at a chemical shift between δ_A and δ_B . Importantly, the linewidth of this peak is broadened due to interference from A and B interconversion during the detection period. This additional broadening is denoted R_{ex} such that $R_2^{\text{obs}} = R_2^0 + R_{\text{ex}}$, where R_2^0 is the relaxation rate in the absence of exchange. If the lines become broad enough, the resonances may disappear from the NMR spectrum (Fig. III.7).

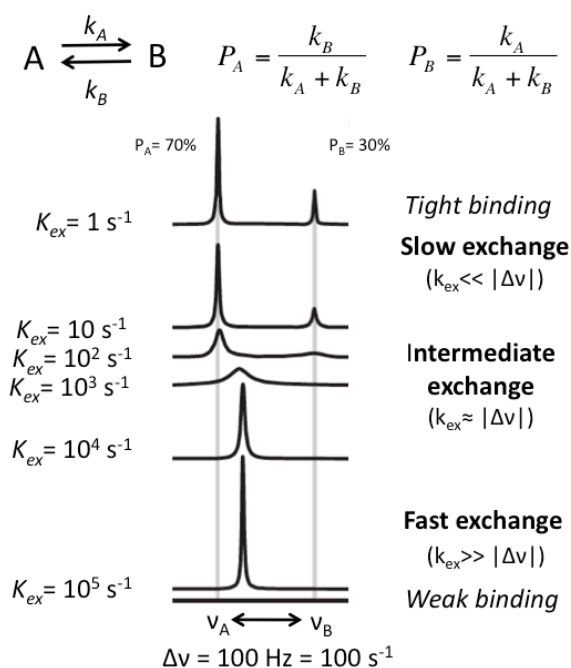


Figure III.7. Exchange between two states A and B can be described by their rates of departure k_A and k_B and chemical shifts ν_A and ν_B . k_{ex} is the total exchange rate and P_A and P_B are the population fractions of each state. The figure represents the effect of varying k_{ex} on NMR spectra with $P_A = 70\%$ and $\Delta\nu = 100$ Hz. Note that the time regime of chemical exchange is defined by the relative magnitudes of k_{ex} and $\Delta\nu$. In the slow exchange regime, signals from both states are observed reflecting their distinct chemical shifts, intensities and linewidths. In fast exchange regime, only one signal is observed reflecting the population-weighted averages of chemical shift, intensity and linewidth. In the intermediate exchange regime, only one signal is observed with intermediate chemical shift. Importantly, the linewidth is increased via exchange broadening, which sometimes can render this signal undetectable. (Adapted from Mittermaier A.K. and Kay L. E., 2009 and Kleckner I.R. and Foster M.P., 2011).

3.3.2.3 Cross-relaxation and the nuclear Overhauser effect

Cross relaxation is a result of dipole-dipole interaction between proximate nuclear spins (see dipolar-coupling, section 3.3.1). It is possible by NMR pulse techniques to reverse the direction of the magnetic dipole of a nuclear spin or to turn it off and measure how this affects any other spin. The rate by which this effect is transmitted to the other nuclear spin is the cross relaxation rate, and this is inversely proportional to the sixth power of the distance between the two nuclear spins. It is because of that the nuclear Overhauser effect (NOE) can be measured between nuclei, which are less than 5 Å apart.

The cross relaxation rate and the related nuclear Overhauser effect can be used to estimate distances between two nuclei in a protein molecule. The effect is not only depending on the distance between the nuclei, but also depends on the overall molecular and internal motions in the protein, which affect the inter-nuclei distance with time. In general, small molecules exhibit positive NOEs, whereas large molecules exhibit negative ones. The experiment used to determine the nuclear Overhauser effect is called NOESY (Nuclear Overhauser Effect Spectroscopy), the most common experiment is a homo-nuclear two dimensional NMR experiment (2D-NOESY).

3.3.3 Chemical shift assignment and structure calculation from NMR data

3.3.3.1 Sequential assignment using heteronuclear scalar coupling

Sequential assignment is a process by which a particular amino acid spin system identified in the spectrum is assigned to a particular residue in the amino acid sequence. In a protein, which is ^{13}C and ^{15}N labeled, the peptide backbone is one long series of scalar coupled nuclei, which allow the complete assignment of the protein.

During the last 20 years, triple resonance pulse sequences have been introduced for structural protein determination by NMR (Sattler M. et al., 1999). These methods rely on ^1H detection and use multiple magnetization transfers through well-resolved one-bond J couplings.

To carry out the chemical shift mapping of the interaction of CIN85-SH3A with the atypical proline-rich targets by NMR (Chapter IV, section 4.2.4), it was necessary to assign the backbone of the protein by correlating the peaks of the ^1H - ^{15}N -HSQC spectrum with the residues in the protein sequence. Standard triple resonance experiments (HNCACB, CBCA(CO)NH, HNCO) were acquired at 25 °C on a Varian NMR Direct-Drive Systems 800 MHz spectrometer equipped with a cryoprobe, on a 0.5 mM of ^{15}N - ^{13}C -labelled CIN85-SH3A domain sample in 50 mM NaPi buffer at pH 6.0 and containing 10%(v/v) D_2O .

The HNCO experiment is designed to correlate the amide $^1\text{H}^{\text{N}}$ and ^{15}N shift of one residue with the carbonyl (^{13}CO) resonance of the preceding residue. HNCACB correlates the same amide $^1\text{H}^{\text{N}}$, ^{15}N pair with its $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ resonances and the $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ of the previous residue. The CBCA(CO)NH correlates the same amide $^1\text{H}^{\text{N}}$, ^{15}N pair with only the $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ resonances of the previous residue. The combined

analysis of the latter two spectra enables the identification of the $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ in the same residue and the preceding residue at individual $^1\text{H}^{\text{N}}\text{-}^{15}\text{N}$ peaks.

For the backbone and side chain assignment of the UBZ1+2 domains of TAX1BP1, NMR experiments were run on a 0.7 mM ^{15}N - ^{13}C -labelled UBZ1+2 sample in 50 mM NaPi buffer at pH 6.0 containing 1 mM TCEP, 10%(v/v) D_2O and 30 μM ZnCl_2 . All the NMR spectra were acquired at 25 °C on a Varian NMR Direct-Drive Systems 600 or 800 MHz spectrometer, the latter equipped with a cryoprobe.

For the complete assignment of UBZ1+2, 2D heteronuclear single quantum correlation (^1H , ^{15}N -HSQC and ^1H , ^{13}C -HSQC), 3D ^{15}N - and ^{13}C -NOESY-HSQC (mixing times of 125 ms) and triple resonance spectra CBCA(CO)NH, HNCACB, HNCO, HBHA(CO)NH and CCH-TOCSY (Sattler M. et al., 1999) and aromatic (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE (Yamazaki T. et al., 1993) experiments were recorded on the double labelled sample.

The HBHA(CO)NH experiment is designed to correlate the ^1H and ^{15}N amide resonances of one residue with both $^1\text{H}^\alpha$ and $^1\text{H}^\beta$ resonances of its preceding residue via the intervening ^{13}CO , $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ spins. The CCH-TOCSY experiment is designed to correlate side chain aliphatic proton and ^{13}C resonances, through the $^1\text{J}(\text{CC})$ and $^1\text{J}(\text{CH})$ coupling constants. It provides an almost complete assignment of the aliphatic ^1H and ^{13}C resonances, except for residues with long aliphatic side chains (like Lys and Arg) where there is a substantial overlap. On the other hand, the (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE experiments, allow the correlation between the side chain C^β chemical shift with the H^δ and H^ϵ aromatic protons, respectively.

All NMR data were processed using NMRPipe (Delaglio F. et al., 1995) and analyzed in CCPNMR (Vranken W. et al., 2005).

3.3.3.2 Structure calculation from NMR data

The structure determination by NMR depends critically on the extent and quality of the assignments. Indeed, the NMR signals contain information about the structural geometry around every observable nuclei in the protein: the distance to the nearby nuclei in the structure can be determined using NOE spectroscopy and the dihedral angles can be derived from the coupling constants and from the chemical shifts, using the program TALOS+ (Shen Y. et al., 2009).

Different programs can be used for the structure calculation based on NMR data: XPLOR-NIH, CNS, ARIA, CYANA and Chesire, among others.

For the structure calculation of the free form of the UBZ1+2 tandem of TAX1BP1 we have used the program CYANA version 2.1 (Guntert P. et al., 1997; Gunter P., 1997; Hermann T. et al., 2002) in combination with RECOORD (Nederveen F.H. et al., 2005) for the water refinement. The experimental input needed to run CYANA is: amino acid sequence, chemical shift list from sequence-specific resonance assignment, list of NOESY cross-peak positions and volumes, and, optionally, conformational constraints from other sources, like for example, dihedral angles or residual dipolar couplings (RDCs). The program also contains information about the covalent structure geometry of all the amino acids of the protein regarding bond lengths and angles as well as

potential functions to ensure that the bonds are maintained within realistic limits. However, based on these inputs, CYANA cannot provide a unique structure output for all the atoms in the protein. Therefore, the structure calculation based on NMR derived structure constraints takes its origin in the generation of a random protein starting structures. The structure determination is repeated several times with new and different starting points. Typically one hundred structures are calculated, and those structures, which comply best to the NMR input data and are energetically most favorable, are selected as group of structures, the NMR ensemble, which is the result of a structure determination by NMR. The superimposition of all members of the NMR ensemble reveals the reproducibility of the structure calculation, reflected by the root mean square deviation (r.m.s.d) of every atom in the structure. The r.m.s.d for the peptide backbone is the sum of the r.m.s.d for all the atoms in the peptide backbone divided by the number of atoms.

For the determination of the solution structure of UBZ1+2, NOE cross peaks were obtained by manual peak picking in each strip of the 800 MHz 3D ^{15}N - and ^{13}C -NOESY-HSQC spectra taken at the assigned ^1H , ^{15}N and ^1H , ^{13}C frequencies, respectively, and by automatic peak picking of the 2D NOESY in H_2O followed by removal of diagonal peaks and peaks arising from artifacts (e.g. residual water). NOEs were assigned using the automated NOE assignment procedure of CYANA version 2.1 (Güntert P. et al., 1997; Güntert P., 1997; Hermann T. et al., 2002). *Phi* and *psi* torsion angle restraints were included based on the analysis of the $^1\text{H}^{\text{N}}$, ^{15}N , $^1\text{H}^{\alpha}$, $^{13}\text{C}^{\alpha}$, ^{13}CO and $^{13}\text{C}^{\beta}$ chemical shifts using the program TALOS+ (Shen Y. et al., 2009). The standard protocol was used with seven cycles of combined automated NOE assignment and structure calculation of 100 conformers in each cycle. Restraints that were unambiguously assigned were used for a final structure calculation run using CNS (Brünger A.T. et al., 1998) in explicit solvent using the RECOORD protocol (Nederveen F.H. et al., 2005). The twenty water-refined structures of lowest energy were used for the final analysis. The quality of the structures was checked with PROCHECK-NMR (Laskowski R.A. et al., 1996) and WHATIF (Vriend G., 1990). MOLMOL (Koradi R. et al., 1996) and PYMOL (<http://www.pymol.org/>) were used for visualization.

3.3.3.3 Chemical shifts to map protein-protein interactions

NMR is a very powerful technique to map the interaction site of a protein upon binding with its partner. The technique is based on analyzing ^1H - ^{15}N -HSQC spectra. As it has been previously indicated, the HSQC spectrum is like the "fingerprint" of a protein because each peak corresponds to an NH group of each residue (except for the first amino acid and prolines). Side chain NH groups (from Asn, Gln, His and Trp) can also give a peak in the spectrum. The chemical shift in both nitrogen and proton dimension is sensitive to the chemical environment of the two nuclei. To map the interaction of a protein with its partner, first it is necessary to record an HSQC of the ^{15}N -labeled protein alone and then HSQC spectra of the protein containing increasing amounts of its unlabeled partner. The peaks showing changes in chemical shifts experience different chemical environments upon binding. If we overlay all the HSQC spectra recorded during the titration, and if the free and bound states are in fast to intermediate exchange (see section 3.3.2.2), the trajectories of the shifting resonances will provide information on the residues involved and/or affected by the binding event.

The magnitude of the chemical shift perturbation ($\Delta\delta$) is calculated by:

$$\Delta\delta = \sqrt{(\Delta\delta_H)^2 + \left(\frac{\Delta\delta_N}{6.51}\right)^2} \quad \text{Eq.18}$$

where $\Delta\delta$ is the difference between the bound and free forms combined chemical shifts (Houben K. et al., 2007).

Dissociation constants are obtained by monitoring the chemical shift changes of the backbone amide as a function of the ligand concentration (or the molar ratio macromolecule/ligand).

The ^{15}N -labeled SH3A domain of CD2AP and CIN85, the ^{15}N -labeled UBZ1+2 domains of Tax1BP1, and ^{15}N -labeled ubiquitin titration experiments were analyzed by plotting the chemical shift deviations upon titration ($\Delta\delta$) as a function of ligand concentration. These data were fitted using the following equations (Eq.19 and Eq.20) that take into account the protein and ligand concentrations at each step of the titration.

For a single binding-site model (Auguin D. et al., 2004):

$$\Delta\delta_{\text{ppm}} = B_{\text{max}} \frac{[P_T] + [L_T] + K_d - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]}}{2[P_T]} \quad \text{Eq.19}$$

where $[P_T]$, $[L_T]$, K_d and B_{max} , are the total protein concentration, ligand concentration, dissociation constant and chemical shift deviation at saturation, respectively.

This equation can be improved by considering possible cooperative binding effect. In that case, the equation includes the Hill coefficient α (Montaville P. et al., 2008):

$$\Delta\delta_{\text{ppm}} = B_{\text{max}} \frac{\left([L_T] - \left(\frac{[P_T] + [L_T] + K_d - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]}}{2} \right) \right)^\alpha}{K_d^\alpha + \left([L_T] - \left(\frac{[P_T] + [L_T] + K_d - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]}}{2} \right) \right)^\alpha} \quad \text{Eq.20}$$

where $[P_T]$, $[L_T]$, K_d and B_{max} , are as in Eq. 19, and α is the Hill coefficient. In case of a Hill coefficient $\alpha = 1$, this equation is simplified to a single binding model (Auguin D. et al., 2004).

The analysis of the chemical shift perturbations and their mapping on the protein structure provides *i.* information of the binding interface and conformational changes of the protein, and *ii.* enables estimation of the affinity, stoichiometry and specificity of the binding.

3.3.4 Paramagnetic Relaxation Enhancement (PRE)

Paramagnetic relaxation enhancement (PRE) provides long-range distance information, which can complement short-range ($\leq 5\text{\AA}$) NOE restraints. The PRE is induced by paramagnetic spin labels (Clare M.G. and Iwahara J., 2009). The spin labels contain unpaired electrons, which affect the magnetic properties of nuclei in their vicinity and cause a distance-dependent line broadening in the NMR spectra (Clare M.G. and Iwahara J., 2009). The protein nuclei close to the paramagnetic centre experience very large PREs, which permits to detect lowly-populated states of macromolecules and their complexes under equilibrium conditions (Clare G.M. and Iwahara J., 2009). This exquisite sensitivity of PRE NMR makes possible the study of transient intermediates in biomolecular interactions. The ranges of distances that can be probed depend of the nature of the particular paramagnetic group (Battiste J.L. and Wagner G., 2000; Iwahara J. et al., 2003). In general, long-range distance information in the 10 - 35 \AA range can be provided through the use of PREs (Bertini I. et al., 1996; Bertini I. et al., 2001; Iwahara J. et al., 2004; Clare G.M. et al., 2007; Clare G.M., 2013). The effects corresponding to the longer distances are going to be smaller than what can be measured in the NMR experiment, and the distances effects shorter than 10 \AA could hardly be estimated because the paramagnetic effects would lead to a substantial decrease in the signal intensity, well below the noise levels of the HSQC spectra.

One of the most common methods to attach a paramagnetic probe to a protein is to conjugate it to a solvent-exposed cysteine residue introduced by site-direct mutagenesis (if the protein lacks it) (Battiste J. L. and Wagner G., 2000; Clare M.G. and Iwahara J., 2009; Volkov A.N. et al., 2010). Basically there are two types of probes designed for Cys modification (Keizers P.H.J. and Ubbink M., 2011): nitroside stable radical and metal chelators (like pEDTA, EDTA, DTPA and metal binding peptides). It is important to remark that the presence of multiple Cys residues in a protein makes difficult the selective incorporation of a paramagnetic probe. It is also necessary to be careful not to alter the interaction energy, structure or kinetics of the interacting molecules by the probe attachment.

For quantitative PRE investigation of protein-protein interactions, measurement of PRE rate Γ_2 is required. To determine the Γ_2 rate most of the examples in the literature make use of a single time-point measurement in which two regular 2D ^1H - ^{15}N -HSQC spectra are recorded, one with a paramagnetic (we used Mn^{2+}) and other with a diamagnetic sample (we used Zn^{2+} as control) at the same concentration.

The use of Mn^{2+} ($3d^5$) as a paramagnetic probe is very common (Tang C. et al., 2006; Suh J.Y. et al., 2007; Tang C. et al., 2008). Mn^{2+} is a paramagnetic ion with the maximum number of unpaired electrons, which presents a characteristic long electron spin relaxation time and provides a highly efficient relaxation enhancement (Sharp R., 2008). Due to the large magnetic moment of an unpaired electron, the paramagnetic effects can extend to distances as long as 35 \AA (Clare G.M., 2013). Paramagnetic metals as Mn^{2+} are characterized by an isotropic g-tensor that produces larger PREs in which dipolar relaxation is predominant and Curie-spin relaxation is insignificant (Clare G.M. et al., 2007). In the absence of ZFS (zero field splitting) isotropic g-tensors only give rise to PRE effects, while anisotropic tensors also results in pseudo-contact shifts (Clare G.M., 2013).

3.3.4.1 PRE experiments

For the PRE NMR study of the interaction of the UBZ1+2 tandem with ubiquitin, we designed three Cys mutants: ubiquitin D39C and E51C, close to the UBZ1+2 binding interface (based on the chemical shift perturbations); and D32C, far from the binding site, and use it as a negative control. These mutants were expressed and purified as explained in Chapter V (section 5.2.2). The paramagnetic probe, a EDTA-based metal chelator N-[S-(2-pyridylthio)cysteaminy]ethylendiamine-N,N,N,N'-tetraacetic acid monoamine (pEDTA) was purchased from Toronto Research Chemicals (P996250, North York, ON, Canada). pEDTA binds divalent metal ions, like paramagnetic Mn^{2+} ($MnCl_2$ solution 1M from Sigma-Aldrich, grade: for molecular biology) and diamagnetic Zn^{2+} ($ZnSO_4$ from Sigma-Aldrich, 99.999% trace metal basis) with very high affinity ($K_a \approx 10^{13} - 10^{16} M^{-1}$). The protocol to covalently attach the probe to ubiquitin's Cys-mutants, based on a thiol-specific reaction, is described in the Chapter V (section 5.2.2.2). The proper attachment of the probe to the three ubiquitin mutants was determined using the DTDP (4,4'-dithiopyrdine, Aldrithiol-4 form Sigma Aldrich) protocol (Piener C.K. et al., 2002) that measure the number of free Cys by spectrophotometric determination of protein shulfydryls, but also a mass spectrometry analysis of every sample was performed (De Duve Institute, platform for proteomic and protein analysis MASSPROT, Brussels, Belgium).

For each observed amide proton of UBZ1+2 tandem, the PRE values were calculated using the equation (Battiste J.L. and Wagner G., 2000):

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2,dia} e^{-\Gamma_2 t}}{R_{2,dia} + \Gamma_2} \quad \text{Eq.21}$$

where I_{para} and I_{dia} are measured intensities of the HSQC peaks of the UBZ1+2 tandem in complex with ubiquitin-pEDTA- Mn^{2+} and ubiquitin-pEDTA- Zn^{2+} respectively; $R_{2,dia}$ is the transverse relaxation rate in the diamagnetic sample; Γ_2 is the paramagnetic contribution to the relaxation rate (PRE); and t is the total polarization transfer time of the HSQC (Volkov A.N. et al., 2006). For each amide peak, $R_{2,dia}$ was estimated from the peak width at half-height ($\Delta\nu_{1/2}$) in proton dimension by using $R_2 = \pi\Delta\nu_{1/2}$. In the case of the residues whose resonances disappear in the paramagnetic spectrum, I_{para} was estimated from noise level of the spectrum. However, this simplified approach underestimates the true value of Γ_2 unless a very long repetition delay is employed (we used 3 s delay, instead of the 1 s delay of a typical HSQC experiment). This is because at short repetition delays, the recovery levels for the paramagnetic sample are always higher than those of the corresponding diamagnetic sample (Otting G., 2010).

3.3.4.2 Ensemble refinement against intermolecular PREs

The ubiquitin coordinates were taken from the published X-ray structure (PDB code 1UBQ), and the UBZ1+2 coordinates obtained from the NMR structure determined in this thesis (Chapter III, section 3.3.3 and Chapter V, section 5.3.1).

Surface Cys mutations of ubiquitin were introduced *in silico*, followed by addition of the pEDTA- Mn^{2+} (spin label, SL) probe and energy minimization of the labeled protein (Volkov A.N. et al., 2010). To take into account the mobility of the pEDTA- Mn^{2+} attached to the surface of ubiquitin, we performed r^{-6} ensemble averaging of the

intermolecular distance restraints with different pEDTA-Mn²⁺ orientations (45 different orientations for every mutant).

Using the Γ_2 dataset obtained from all pEDTA-Mn²⁺ conjugation sites, the rigid-body simulated annealing refinement of the UBZ12-ubiquitin complex was carried out using Xplor-NIH (Schwieters C.D. et al., 2003, 2006) following the published procedure (Iwahara J. et al., 2004; Tang C. et al., 2006; Volkov A.N. et al., 2010). Basically, the position of the UBZ1+2 was fixed, and multiple copies of ubiquitin molecules representing ensembles with $N = 1 - 20$ were docked to minimize the energy function consisting of the PRE target term, van der Waals (vdW) repulsion term to prevent atomic overlap between ubiquitin and UBZ12, and a weak radius-of-gyration restraint used to encourage intermolecular UBZ12-ubiquitin contacts (Tang C. et al., 2006). It is important to note that this procedure allows for the atomic overlap among ubiquitin molecules constituting an ensemble. For every N , 100 independent refinement runs were performed, and for every run expected PRE (Γ_2) values were back-calculated.

To evaluate the agreement between the observed PREs and the PREs back-calculated from ubiquitin ensembles generated in each run, we computed the Q factor (Q_f):

$$Q_f = \sqrt{\frac{\sum_j \sum_i (\Gamma_{2,ij}^{obs} - \Gamma_{2,ij}^{calc})^2}{\sum_j \sum_i (\Gamma_{2,ij}^{obs})^2}} \quad \text{Eq.22}$$

where $j = 1 - 3$ runs over three positions showing paramagnetic effects (D32C, D39C and E51C), and $\Gamma_{2,ij}^{calc}$ is given by the equation (Volkov A.N. et al., 2010):

$$\Gamma_{2,ij}^{calc} = \frac{p_2}{N} \sum_{k=1}^N \Gamma_{2,ijk}^{low} + (1 - p_1) \Gamma_{2,ij}^{high} \quad \text{Eq.23}$$

where p_1 is the percentage (100%) of bound ubiquitin in the high affinity site (UBZ2) and p_2 is the percentage (65.5%) of bound ubiquitin in the low affinity site (UBZ1) at the ratio UBZ1+2/ubiquitin 1/3; N is the size of the ubiquitin ensemble, $\Gamma_{2,ijk}^{low}$ is the PRE from SL(j) back-calculated for the residue (i) of the ubiquitin ensemble member (k), and $\Gamma_{2,ij}^{high}$ is the PRE backcalculated from SL(j) for the residue (i) of ubiquitin in the high affinity site of the complex.

3.3.5 Using relaxation to study ¹⁵N backbone dynamics

Relaxation is the process by which the nuclear magnetization in a non-equilibrium state returns to its equilibrium distribution. There are two basic types of relaxation:

- **T1**, the spin-lattice relaxation: the magnetization remaining along the z-axis relaxes back to its equilibrium. This relaxation happens by the loss of energy from the excited nuclear spins to the surrounding molecular lattice (enthalpic).
- **T2**, the spin-spin relaxation: the transverse component (x-y plane) of the magnetization vector, exponentially decays coming back to equilibrium

values. This relaxation is due to the excited spins exchanging energy between each others, and represent the loss of phase coherence among nuclei (entropic).

T2 is the same or faster than T1 relaxation, because return of magnetization to equilibrium in the z-axis inherently causes loss of magnetization in the x-y plane. In general, for small molecules $T1 \approx T2$, but for large molecules $T1 \gg T2$.

During the backbone dynamics experiments by NMR, we commonly measure three ^{15}N heteronuclear relaxation rates: the spin-lattice relaxation rate, also known as longitudinal relaxation ($R_1 = 1/T_1$), the spin-spin relaxation rate, also known as transverse relaxation ($R_2 = 1/T_2$) and the ^1H - ^{15}N -heteronuclear NOE (^1H - ^{15}N steady-state NOE). These parameters are sensitive to the internal motions of the protein on the subnanosecond timescale. Molecular motion causes a field fluctuation that induces the relaxation process. There are several relaxation mechanisms: dipolar and quadrupolar couplings, paramagnetic, scalar coupling, chemical shift anisotropy (CSA) and spin rotation.

For the nuclear relaxation of an amide nitrogen in proteins, two relaxation mechanisms should be taken into account, the dipole-dipole interaction between a nitrogen and the hydrogen directly bound to it and the CSA of nitrogen nucleus. The relaxation rates R_1 and R_2 depend on the spectral density function $J(\omega)$ (Eq. 24 and Eq.25), which is the probability function of finding motions at a given angular frequency ω (Korzhnev D.M. et al., 2001).

$$R_1 = \frac{1}{4}d^2 \{J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\} + c^2J(\omega_N) \quad \text{Eq.24}$$

$$R_2 = \frac{1}{8}d^2 \{4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H - \omega_N)\} + \frac{c^2}{6} \{4J(0) + 3J(\omega_N)\} + R_{ex} \quad \text{Eq.25}$$

$$\text{NOE} = \frac{d^2}{4R_1} \frac{\gamma_H}{\gamma_N} \{6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\} + 1 \quad \text{Eq.26}$$

where

$$d = \frac{\mu_0}{4\pi} \gamma_H \gamma_N \frac{h}{2\pi} (r_{NH}^{-3}) \quad \text{Eq.27}$$

$$c = \frac{\omega_N (\sigma_{\parallel} - \sigma_{\perp})}{\sqrt{3}} \quad \text{Eq.28}$$

where μ_0 is the permeability of the free space, γ_H and γ_N are the gyromagnetic ratios of ^1H and ^{15}N ; ω_H and ω_N are the Larmor frequencies of ^1H and ^{15}N , r_{NH} is the N-H bond length (assumed to average of 1.01Å), and $J(\omega_i)$ are the spectral densities at the angular frequency ω_i . σ_{\parallel} and σ_{\perp} are the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor (approximated to -160ppm, Tjandra N. et al., 1996).

R_1 , R_2 and the heteronuclear NOE can be described in terms of the order parameter (S^2) and the correlation times, using a software like TENSOR2 (Dosset P. et al., 2000), which carries out an analysis of internal and global motion using the model-free formalism of Lipari and Szabo (1982) extended by Clore G.M. et al., (1990). The order parameter S^2 specifies the degree of spatial restriction of the N-H bond: $S^2 = 1$ when the bond is rigid, relaxation is described only by global motion, and $S^2 = 0$ when internal motion is isotropic, and relaxation is described only by internal motion. The overall rotational correlation time (τ_c) characterizes the tumbling of the molecule. Molecules that have a fast random rotation (small molecules) present short correlation times and have a broad spectral density function, and because of that, can sample a wide range of frequencies. However, molecules that tumble slowly (big molecules) have longer correlation times, and sample only lower frequencies. The correlation time is mainly determined by the molecular weight, but can be influenced by parameters like temperature, solvent viscosity, aggregation in solution, and so on.

For the NMR experiments, samples of the free form of UBZ1+2 domains at 0.5 mM and the complex formed with ubiquitin at different ratios (UBZ1+2/ubiquitin: 1/0.5; 1/1.1: 1/5) were prepared in 50 mM NaPi pH 6.0, containing 9% D₂O. The relaxation parameters ^{15}N R_1 , R_2 and ^1H - ^{15}N steady-state NOE were measured at 600 MHz and 25 °C. Relaxation values were obtained from series of 2D experiments with coherence selection achieved by pulse field gradients according to standard methods (Farrow N.A. et al., 1994). The ^1H - ^{15}N heteronuclear NOEs were determined from the ratio of peak intensities ($I_{\text{sat}}/I_{\text{eq}}$) with and without the saturation of the amide protons for 3 s. ^{15}N R_1 and ^{15}N R_2 relaxation rates were measured from spectra with different relaxation delays: 100, 200, 300, 400, 500, 700, 900, 1200 and 1500 ms for R_1 and 10, 30, 50, 70, 90, 110, 150, 190 and 230 ms for R_2 measurements. The relaxation parameters and their corresponding errors were extracted with the program NMRview (Johnson B.A., 2004) and CCpNMR (Vranken W. et al., 2005). R_1 and R_2 values for the free form of UBZ1+2 were extracted by fitting the peak intensities to a single exponential decay function in NMRview. T_1 and T_2 values of the complex UBZ1+2 with ubiquitin at different ratios, were extracted by fitting the peak intensities to a single exponential decay function in CCpNMR, and converted to R_1 and R_2 values ($R_1 = 1/T_1$ and $R_2 = 1/T_2$). Peak height uncertainties were determined by recording duplicate spectra at a particular time points (Jarymowycz V.A. and Stone M.J., 2006).

We have used the program TENSOR2 (Dosset P. et al., 2000) to perform a model-free analysis (Lipari and Szabo) of the local internal mobility affecting the backbone amides in the presence of an anisotropic tensor, to calculate the apparent rotational correlation time (τ_c) and the order parameter (S^2) of the free UBZ1+2 domains of TAX1BP1. To obtain a reliable estimation of the overall rotational diffusion tensor of a protein, it is necessary to only consider for the analysis of the relaxation parameters, the residues having fast internal motions and for which there is no significant contribution to R_2 from conformational exchange. The residues not satisfying these criteria are identified by their low NOE values (in the case of the fast internal motion) and elevated R_2/R_1 ratios (in the case of residues experiencing conformational exchanges on a slow time scale).

Once that these residues have been excluded, the rotational diffusion tensor of the protein can be obtained by analysis of the R_2/R_1 ratios of the remaining residues using the program TENSOR2. It is worth to notice that the overall rotational diffusion

correlation time can be determined using the R2/R1 ratio as described (Kay L.E. et al., 1989). It is necessary to independently determine isotropic, axially symmetric and fully anisotropic rotational diffusion tensor, to be able to select the most appropriate diffusion tensor to use for analysis of internal dynamics, based on the fitting statistics. In the program TENSOR2, uncertainties are determined based on Monte Carlo simulations (Metropolis N. and Ulam S., 1949). Once that the rotational diffusion tensor has been determined, and also by using the program TENSOR2, relaxation data can be analyzed in the context of the selected diffusion tensor using any of the five versions (model 1-5) of the model free formalism (Jarymowycz V.A. and Stone M.J., 2006). The program TENSOR2 (based on the existence or not of conformational exchange or the existence or not of two different time scales of internal motions) fits each residue independently to each mathematical model to yield the corresponding internal dynamics parameters.

3.4 Small Angle X-ray Scattering (SAXS)

Small Angle X-ray Scattering provides information about the shape and conformation of a macromolecule in solution at low resolution (to a maximum of about 10 Å resolution) (Putnam C.D. et al., 2007). To obtain good SAXS datasets, it is important that the samples are homogeneous, monodisperse and lacking long-range interactions in solution (Koch M.H. et al., 2003).

In SAXS, the scattering signal is derived from the difference in the average electron density of particle $\rho(r)$, and bulk solvent ρ_s .

$$\Delta\rho(r) = \rho(r) - \rho_s \quad \text{Eq.30}$$

The scattering curve resulting from this subtraction, $I(q)$, is isotropic due to the randomly distribution of the orientations of the protein in solution. $I(q)$ is a function of the momentum transfer (q):

$$q = \frac{(4\pi\sin 2\theta)}{\lambda} \quad \text{Eq. 31}$$

where 2θ is the scattering angle and λ is the wavelength of the incident X-ray beam (in Å⁻¹ or nm⁻¹). We can obtain the scattering curve of a homogeneous sample from the electron distribution of the protein [$P(r)$].

$$I(q) = 4\pi \int_0^{D_{\max}} P(r) \frac{\sin(qr)}{qr} dr \quad \text{Eq.32}$$

where D_{\max} is the maximum distance present in the scattering particle (the protein or protein complex).

The low q region of the SAXS curve is dictated by the shape and radius of gyration (R_g) of the scattering particle. The R_g is the square root of the average squared distance between points in the particle (taking mass into account). R_g is a good indicator of how the mass of a particle is distributed around its center of mass. The Guinier approximation (extrapolation of the intensity at $q = 0$) allows the calculation of the R_g :

$$I(q) = I(0) \cdot e^{-\frac{(q^2 R_g^2)}{3}} \quad \text{Eq.33}$$

From the slope of the plot of $I(q)$ against q^2 (*Guinier* plot) we can extract the values of R_g and $I(0)$ (Guinier A. and Fournet F., 1955). Higher q values contain details regarding molecular shape. The intensity of the scattering of globular particles (systems with a clear difference between particle and surrounding solvent) obeys the Porod's law (Porod G., 1951): $I(q) \propto q^{-4}$. Complex mixtures or highly conformationally heterogeneous samples such as unfolded proteins do not follow this rule.

To evaluate the folding of the sample we can use the Kratky plot [representation of $q^2I(q)$ vs q]. For folded domains, the Kratky plot yield a peak roughly shaped like a parabola, and the position of the peak provides information about its overall size. However, random coil of unstructured samples, lack this characteristic shape and in the large q -region, are linear with respect to q (Brulet A. et al., 1996).

From one-dimensional SAXS curve it is possible to derive a three-dimensional reconstruction shape of the protein (*ab initio* shape). Several programs have been created for this purpose: DALAI_GA (Chacon P. et al., 1998, 2000), DAMMIN (Svergun D.I., 1999), SASHA (Svergun D.I. et al., 1996) and so on. In general, these programs assume that the *ab initio* shape is a continuous object, which substantially reduces the search space. *Ab initio* shape determination can be applied to reconstruct the low-resolution envelopes of the protein. The final resolution of the envelopes varies depending on the data quality, the particle size, the shape, and flexibility. The combination of external constraints and multiple runs is crucial for the low-resolution envelope reconstruction of many systems (Volkov V.V. and Svergun D.I., 2003; Petoukhov M.V. and Svergun D.I., 2005; Pascal J.M. et al., 2006).

The quality of the fit between the experimental scattering data and those predicted from the model is described by the χ^2 statistics (Grishaev A. et al., 2005). The χ^2 statistics varies greatly with the size of the protein, since the scattering intensity $I(q)$ grows quadratically with the number of atoms.

During this thesis, SAXS has been used as a complementary tool to confirm the results obtained by other solution techniques like ITC and NMR. For the study of the interaction of the SH3A domain of CD2AP and CIN85 with atypical proline-rich targets (Chapter IV, section 4.3.6), SAXS has been key to differentiate between the presence of hetero-dimers and hetero-trimers in solution. For the study of the interaction of the TAX1BP1 UBZ1+2 tandem with ubiquitin (Chapter V, section 5.3.6), SAXS was used to validate the high-resolution structure of the free UBZ1+2 obtained by NMR, and to confirm the complex that ubiquitin forms with UBZ2 domain determined by PRE NMR. Specific experimental conditions of the SAXS measurement carried out with the SH3A domains of CD2AP and CIN85 and the UBZ1+2 domains of TAX1BP1 are indicated in Chapter IV (section 4.2.5) and Chapter V (section 5.2.6), respectively.

The general protocol followed to analyze the experimental data is indicated below. The data were averaged, background-subtracted, and merged to generate the scattering curve with PRIMUS (Konarev P.V. et al., 2003). The radius of gyration (R_g) was calculated from the *Guinier* analysis as implemented in PRIMUS and from the indirect Fourier transform package GNOM (Konarev P.V. et al., 2003; Konarev P.V. et al., 2006). CRY SOL (Svergun D.I. et al., 1995) was used to generate theoretical scattering curves based on the PDB, to calculate the theoretical R_g from atomic coordinates and to compare experimental and theoretical scattering curves.

3.5 Circular Dichroism (CD)

Circular dichroism (CD) spectroscopy measures the wavelength dependence of the differential absorption of the left and right-handed circularly polarized light. Both components (left and right) of a polarized beam of light interact differently with the chiral centres of an optically active molecule. This interaction causes a decrease in the velocity of wave propagation and its absorption (which depends of the wavelength, Eq.34):

$$\Delta\text{Abs}(\lambda) = \Delta\varepsilon \cdot c \cdot p = \text{Abs}_L(\lambda) - \text{Abs}_R(\lambda) = [\varepsilon_L(\lambda) - \varepsilon_R(\lambda)] \cdot c \cdot p \quad \text{Eq. 34}$$

where $\text{Abs}_L(\lambda)$ and $\text{Abs}_R(\lambda)$ are the absorbances of the sample of the left-and right-handed circularly polarized light by the sample, respectively; c is the sample concentration in $\text{mol}\cdot\text{L}^{-1}$; p is the light path length in cm; and $\varepsilon_L(\lambda)$ and $\varepsilon_R(\lambda)$ the molar extinction coefficients for each of the two light components in $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$. The result is a difference in both amplitude and phase between the two components of the polarized light; the sum vector of these two components rotates following an elliptical helical path, which is proportional to the $\Delta\text{Abs}(\lambda)$, measured in ellipticity units (θ) (degrees). The normalized CD spectra are expressed in molar ellipticity per residue in $\text{deg}\cdot\text{dmol}^{-1}\cdot\text{cm}^2$.

The CD bands observed in proteins are situated in two regions of the spectrum: the far-UV (170 - 250 nm), where most contributions to the absorption bands come from the peptide bonds and is very sensitive to changes in the different secondary structure elements (α -helix, β -sheet and random coil) in proteins (Compton L.A. and Johnson W.C., 1986); and the near-UV (250 - 320 nm) where the bands arise from those amino acids in the sequence of the protein containing aromatic side chains, and gives information about changes in the tertiary structure of the protein.

CD experiments have been used in this thesis to check the proper folding of the single UBZ1 and UBZ2 domains and the tandem UBZ1+2 in presence and absence of Zn^{2+} . Samples were prepared in 50 mM NaPi buffer at pH 6.0 containing 1 mM TCEP. Measurements in the Far-UV CD spectra (250-195 nm) were performed at 25 °C on a Jasco J-715 (Tokyo, Japan) spectropolarimeter equipped with a Peltier temperature control system and using a 0.1 mm path length quartz cuvette. Each resulting spectrum was the average of eight scans at $100 \text{ nm}\cdot\text{min}^{-1}$, 1 nm step resolution, 1 s response and 1 nm bandwidth. A nitrogen flow of $10 \text{ L}\cdot\text{min}^{-1}$ was used to purge and refrigerate the system.

For comparison, the results were normalized and expressed in molar ellipticity ($\text{deg}\cdot\text{dmol}^{-1}\cdot\text{cm}^2$) per residue, according to Eq.35

$$[\theta] = \frac{\theta}{c \cdot p \cdot 10 \cdot n} \quad \text{Eq. 35}$$

where, θ is the experimental ellipticity in millidegrees; p is the cuvette path length in cm; c is the sample concentration in $\text{mol}\cdot\text{L}^{-1}$; 10 is a factor accounting for the conversion of the concentration units from $\text{mol}\cdot\text{L}^{-1}$ to $\text{dmol}\cdot\text{cm}^{-3}$ and ellipticity units from millidegrees to degrees; and n is the number of residues of the protein.

3.6 Bio-Layer Interferometry (BLI)

Bio-layer interferometry (BLI) is an optical analytical technique based on the use of a biosensor tip, where a protein (ligand) can be immobilized. The sensor is dipped into a sample in a microtiter plate, where the binding partner (analyte) is located. BLI is able to measure the interference pattern of white light reflected from two surfaces: the layer on the biosensor tip and an internal reference layer (Fig. III.9A, reflected beams). This interference is captured by a spectrometer as a pattern of intensity variation by wavelength ($\Delta\lambda$) with a characteristic profile of peaks and troughs. Any change in the number of molecules bound to the biosensor tip causes a shift in the interference pattern that can be measured in real-time, providing the ability to monitor binding specificity, the association rate constant (k_{on}), dissociation rate constant (k_{off}), and equilibrium dissociation constant (K_d).

We used this technique to study the interaction of tandem of UBZ1+2 domains (analyte) of TAX1BP1 with different polyubiquitin chains (ligand): di- and tetra- linear, Lys48- and Lys63-polyubiquitin chains. The experiments were conducted on an Octet RED96 bio-layer interferometry instrument (Fortebio Inc.) operating at 25 °C and with eight channels for simultaneous detection.

3.6.1 Experimental setups

The polyubiquitin chains were coupled to the biosensor tips, and dipped in solutions containing different concentrations of UBZ1+2. Typical concentration for the immobilized ligand was 25 $\mu\text{g/mL}$, while the concentrations of UBZ1+2 varied from 0 to 300 μM , based on the affinity intervals of the system. All ubiquitin and polyubiquitin samples were purchased from Boston Biochem (US). Biotinylated N-terminal monoubiquitin, Lys48- and Lys63-tetraubiquitin chains were directly immobilized in streptavidin-coated (SA) biosensors (Fortebio). The linear tetraubiquitin and di- (linear, Lys48- and Lys63-) polyubiquitins were immobilized onto amine reactive (AR2G) biosensors (Fortebio) by chemical modification of surface lysines present in ubiquitin. The AR2G sensors allow the covalent coupling of proteins via EDC and sulfo-NHS (Sigma-Aldrich) mediated amide bond formation.

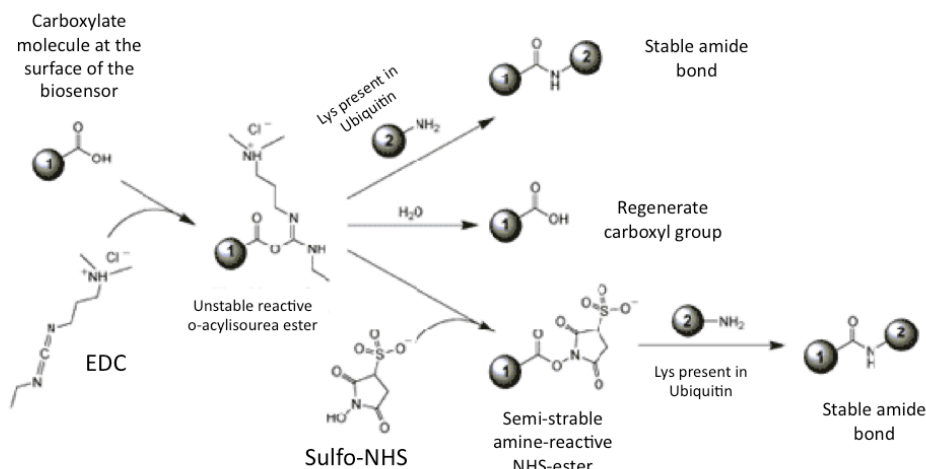


Figure III.8. Schematic representation of the chemical reactions involved in the activation process of the amine reactive sensors. EDC reacts with the carboxylic groups of the biosensor and form an amine-reactive O-acylisourea intermediate, which may react immediately with an

amine group present on ubiquitin (Lys), yielding a conjugate of the ubiquitin and the sensor by a stable amide bond. However, the intermediate is also susceptible of hydrolysis, making it unstable in water solutions. The addition of sulfo-NHS stabilizes the amine-reactive intermediate by converting it to an amine-reactive sulfo-NHS ester, increasing thereby the efficiency of EDC-mediated coupling reactions. Figure adapted from Thermo-Scientific website (www.piercenet.com).

The amine reactive biosensor surface contains many available carboxylic acid groups. To activate this surface toward nucleophilic attack (Fig.III.8) was necessary to treat the surface with an EDC/NHS (Sigma-Aldrich) mixture at 0.4 M and 0.1 M respectively. After that, the exposure of a protein (containing at least one Lys) at a pH below its pI resulted in an amide bond with the carboxylate group of the biosensor surface. In this way, the di-ubiquitin chains were covalently bound to the biosensor. The surface of the biosensor was subsequently treated with ethanolamine (1 M, pH 8.5) to block the unused activated carboxylate groups, minimizing non-specific binding on the amine reactive biosensor surface (quench).

The main steps of the BLI experiments are described as follows (Fig.III.9). Like starting point, a linkage-specific polyubiquitin chain was immobilized on the sensor surface with the appropriate coupling chemistry (SA sensor or AR2G sensor), for that, eight sensor tips (Octed RED96 contains eight channels for simultaneous detection) were introduced in eight wells of the microtiter plate containing buffer [Buffer A: 50 mM Tris-HCl at pH 7.5, 150 mM NaCl, 0.5 mM DTT and 0.005% tween-20 for the streptavidin (SA) sensors, and Buffer B: 50 mM acetic acid, pH 4.0, for the amine reactive sensors (AR2G)], this process had a variable duration, around 100 - 300 s. The sensors were then introduced in eight wells containing the polyubiquitin chains to be coupled to the sensor (loading process). In the case of amine reactive sensors (AR2G) was also necessary to carry out an activation step (EDC/NHS), followed by quench (ethanolamide) of the surface as explained previously (Fig. III.8). A 300 s washing step (baseline) was applied prior to the analysis of the association of the ligand (polyubiquitin chains) on the biosensor to the analyte (UBZ1+2) in buffer A (0 to 300 μ M) for 600 - 900 s. The association process was preceded by a dissociation step for 600 - 900 s. After that, it was necessary to include a step of regeneration (100 - 300 s), inserting the biosensors in buffer A with 500 mM NaCl. The cycle of baseline, association, dissociation and regeneration (Fig. III.9) was repeated at least two times for every studied interaction between N-terminal mono, di- and tetra- (linear, Lys48- and Lys63-) polyubiquitin chains.

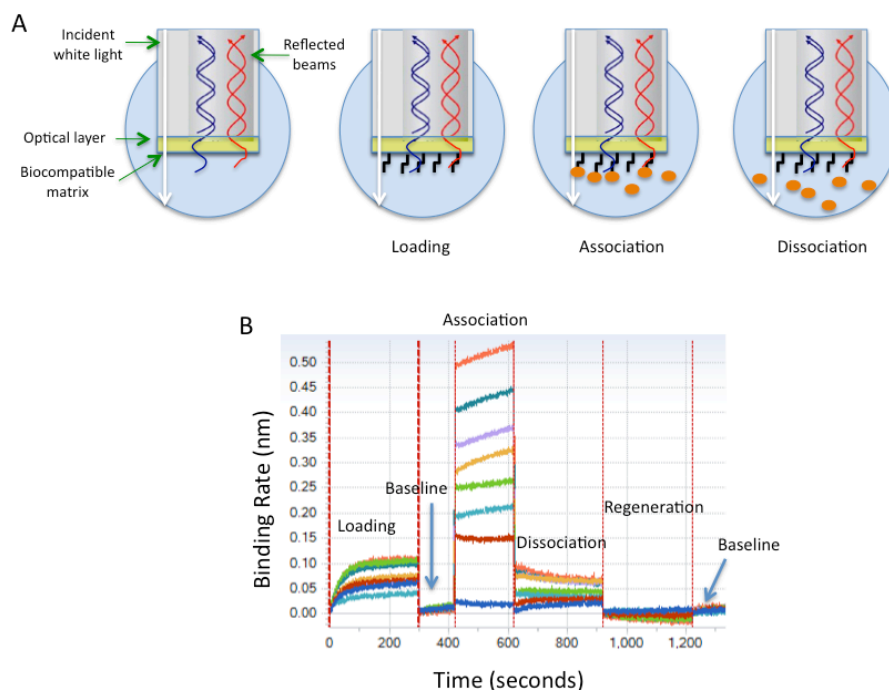


Figure III.9. *A. Representation of the tip surface of the biosensors used during the measurements performed on the BLI Octet RED96 (fortebio). The main parts of the tip are indicated by arrows. The principal events on the biosensor surface are represented: loading of the polyubiquitin chains (black sticks), association of the UBZ1+2 domains (orange circles) and dissociation. B. Typical binding cycle observed with the Octet RED interferometer. Step 1: loading of the polyubiquitin chains, that are coupled to the biocompatible matrix by the proper chemistry; step 2: baseline; step 3: association kinetics of the UBZ1+2 domains; step 4: dissociation kinetics; step 5: regeneration; step 6: baseline. The eight different colors correspond to different concentrations of UBZ1+2 domains measured simultaneously.*

3.6.2 Data analysis

Data Analysis and curve fitting were done using the Octet software version 7.0. Steady state analysis were done for every data set to calculate the apparent K_d using the estimate response at equilibrium for each analyte concentration rather than the K_{on} and K_{off} . All experiments were repeated at least twice and the results obtained were similar.

To determine the overall equilibrium dissociation constants, binding responses in the steady-state region of the sensorgrams were plotted against analyte (UBZ1+2) concentration. The data were subject to non-linear regression according to Eq.36.

$$R_{esp} = C \cdot \frac{R_{max}}{C + K_d} \quad \text{Eq. 36}$$

where R_{esp} represents the changes in refractive index, C is the concentration of analyte in solution, R_{max} is the maximum specific binding (in the same units as R_{esp}), which is the specific binding extrapolated to very high concentrations of analyte, and so its value is almost always higher than any specific binding measured in the experiment. And K_d is the apparent equilibrium binding constant (accounting for affinity and avidity effects), in the same units as C , needed to achieve a half-maximum binding at equilibrium.

CHAPTER IV

MOLECULAR BASIS OF THE VERSATILE INTERACTIONS OF SH3A DOMAINS OF CD2AP AND CIN85 WITH THE ATYPICAL PROLINE-RICH SEQUENCE PRESENT IN CD2 AND Cbl-b*

* SH3A DOMAINS OF CIN85/CD2AP

Part of the work presented in this Chapter has been published with the title:

Multimeric and differential binding of CIN85/CD2AP with two atypical proline-rich sequences from CD2 and Cbl-b

In FEBS Journal, May 2013

Ceregido M.A., Garcia-Pino A., Ortega J.L., Casares S., Mayorga O., Bravo J., van Nuland N.A.J. & Azuaga A.I.

Thus, the task is, not so much to see what no one has yet seen; but to think what nobody has yet thought, about that which everybody sees. Erwin Schrödinger.

Chapter IV. MOLECULAR BASIS OF THE VERSATILE INTERACTIONS OF SH3A DOMAINS OF CD2AP AND CIN85 WITH THE ATYPICAL PROLINE-RICH SEQUENCES PRESENT IN CD2 AND Cbl-b

4.1 Introduction

Adaptor molecules are non-catalytic polypeptides that contain one or more domains able to bind to other protein or non-protein complexes that transmit intracellular signals involved in regulation of cell growth, differentiation, migration and survival (Pawson T. et al., 2002). Formation of higher-order oligomers is required to mediate efficient transduction of signals in many cascades such as internalization of epidermal growth factor receptors induced by Receptor Tyrosine Kinases (RTKs). The CIN85/CMS (CD2AP is the mouse ortholog of CMS) family of adaptor proteins is involved in orchestrating multiple steps in RTK signaling and endocytosis (Dikic I., 2002) and share a similar domain organization with a high degree of sequence conservation (Chapter I, section 1.3).

CIN85 and CD2AP are implicated in Cbl-mediated down-regulation of diverse tyrosine kinases (Kowanetz K. et al., 2003). Growth factor-induced tyrosine phosphorylation of Cbl, which enhances the binding of the atypical proline-rich sequence present in Cbl/Cbl-b to CIN85 or CD2AP via its SH3 domains (Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003; Jozic D. et al., 2005), is crucial for targeting RTKs to clathrin-mediated endocytosis. In the cell the interaction of the SH3 domains of CIN85/CD2AP family is involved in the internalization of signal receptors by ligand stimulation (Dikic I., 2002). In this sense, the association of CD2AP and CIN85 to CD2 receptor is also induced upon T-cell activation (Dustin M.L. et al., 1998; Dikic I., 2002). It has been proposed that central clustering of CD2 has an essential role in antigen receptor engagement and signaling (Dustin M.L. et al., 1998). Direct molecular interaction of CD2AP with CD2 is crucial for CD2 clustering as well as T-cell polarization and recruitment of CD2 during the formation of an immunological synapse (Kurakin A.V. et al., 2003; Dustin M.L. et al., 1998). CIN85 has also been proposed to participate in this process since it is highly expressed in T-cells and can bind CD2 (Dikic I., 2002; Kurakin A.V. et al., 2003).

The composition and stoichiometry of protein clusters involving such adaptor molecules is crucial for receptor downregulation and signaling, however, it is still not fully established. CIN85 and CD2AP facilitate the formation of large multimeric complexes required for an efficient internalization of cell surface receptors by means of three N-terminal SH3 domains (Dikic I., 2002). Most SH3 domains bind to proline-rich sequences containing a PxxP motif (Chapter I, section 1.4). They are recognized in either type I orientation, when the SH3 hydrophilic pocket accommodates basic residues N-terminal to the PxxP motif, or type II when the pocket is occupied by basic residues C-terminal to the motif. However, the N-terminal SH3 domains (SH3A) of CIN85 and CD2AP were found to recognize an atypical polyproline-arginine binding motif

(PxxxPR) present in the ubiquitin ligases Cbl-b and c-Cbl (Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003). Crystallographic studies showed that a Cbl-b derived peptide covering the proline-rich region from residue 902 to 912 interacts with two SH3 domains in a pseudo-symmetrical fashion forming a hetero-trimeric complex (Jozic D. et al., 2005). The peptide is simultaneously recognized by two SH3A domains; one SH3A engages the peptide resembling a type II orientation, whereas the second SH3A binds on the opposite side of the ligand resembling a type I orientation. Moreover, ITC data shows a marked deviation from the 1:1 stoichiometry expected for a single site interaction, which suggests the formation of higher oligomers such as the trimers observed in the crystal structure (Jozic D. et al., 2005). In support of this, hetero-trimerization was also shown *in vivo* (Jozic D. et al., 2005).

Similarly to Cbl-b and other target proteins, CD2 is also recognized via its atypical polyproline-arginine sequence PxxxPR (Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003; Zarrinpar A. et al., 2003). The crystal structure of CMS N-terminal SH3A domain (that shares 100% of sequence identity with the SH3A domain of CD2AP) in complex with a CD2 derived peptide also suggested the formation of hetero-trimers providing a conserved feature in the molecular mechanism of cluster formation involving the CIN85/CMS family of adaptor proteins (Moncalian G. et al., 2006). However, these findings were later disputed since apparently no evidences for two binding sites were found for CIN85-Cbl-b under solution conditions (Kowanetz K. et al., 2003; Ababou A. et al., 2008).

In our attempt to understand the oligomerisation mechanisms that involves atypical proline-rich targets, we carried out an in-dept thermodynamic characterization by ITC and structural characterization using NMR and SAXS, to clarify the composition and stoichiometry of the complexes that the SH3A domains of the adaptor proteins CD2AP and CIN85 form in solution with their natural targets CD2 and Cbl-b.

4.2 Experimental procedures

4.2.1 Purification of ^{15}N - ^{13}C -labeled SH3A domain of CIN85

Unlabeled, ^{15}N -labeled and ^{15}N - ^{13}C -labeled SH3A domain of CD2AP and unlabeled CIN85-SH3A were overexpressed and purified as described in Chapter III (section 3.1). Protein concentrations were measured in an UV-VIS spectrophotometer (Cary 100Bio, Agilent) using an $\epsilon_{280} = 9,970 \text{ M}^{-1}\text{cm}^{-1}$, and $12,490 \text{ M}^{-1}\text{cm}^{-1}$ for the SH3A domain of CD2AP and CIN85, respectively, that were determined using the ProtParam algorithm (www.expasy.ch).

For the overexpression and purification of the ^{15}N and ^{15}N - ^{13}C isotopically labeled CIN85-SH3A, a new protocol was developed (Chapter III, section 3.1.1 and 3.1.2).

We achieved optimal overexpression of the SH3A domain of CIN85 (Fig. IV.1A) in BL21 (DE3) *E. coli* strains incubated in 2xTY medium at 37 °C and changing to enrich minimal media M9 (proportion: cells growth in 4 L of 2xTY were centrifuged and resuspended in 1L of minimal media) before induction with IPTG (1 mM), followed by incubation at 37 °C during the optimized period of 18 hours (Fig. IV.1A, lane 4). The

expression decreased upon longer incubation (Fig. IV.1A, lane 5, 24 hours incubation). Figure IV.1 shows the different steps of the purification protocol of the isotope enriched samples of CIN85-SH3A₂₋₅₈ His-tag.

Since the protein was expressed in inclusion bodies, 8 M urea was added to the lysis buffer to break them.

After extraction of inclusion bodies (as described in Chapter III, section 3.1.2) the protein was soluble in the lysis buffer, which also contained imidazole in low concentration (see buffer composition in Chapter III, section 3.1.2) to avoid competition of contaminants in binding to the nickel (Ni) charged sepharose resin (GE Healthcare). An imidazole gradient (from buffer A with 20 mM imidazole to buffer B with 500 mM imidazole) was performed in presence of 8 M urea on a Ni-chelating column (GE Healthcare), where the unfolded His-tagged protein was selectively eluted (See Fig. IV.1B) as a single peak with the maximum at 18.6% of buffer B (see gel in Fig. IV.1C). Afterwards the SH3A domain of CIN85 with the His-Tag was refolded by successive dialysis steps (G2 dialysis cassettes molecular weight cut-off 3.5 kDa, Thermo Scientific) as described in Chapter III, section 3.1.2.

To remove the His-tag, the sample was incubated with the enzyme protease preScission (GE Healthcare) as described in Chapter III, section 3.1.2. As can be seen in the SDS-PAGE in figure IV.1E, the cleavage process worked properly. After the cleavage, the protein was loaded again on a Ni-chelating column (GE Healthcare) to separate the cleaved SH3A domain of CIN85 from the uncleaved one. As shown in the figure IV.1D, most of the protein (around 98%) was properly cleaved, and eluted in buffer A without imidazole, in a single peak. After a final purification step on a gel filtration column Superdex-75 16/60 (GE Healthcare), maximum purity of was achieved, and the folded protein was recovered at an elution volume of 83 mL (Fig. IV.1F), corresponding to a monomeric state (7.0 kD) according to the column calibration curve.

The protein concentration was measured in an UV-VIS spectrophotometer (Cary 100Bio, Agilent) using an $\epsilon_{280} = 12,490 \text{ M}^{-1}\text{cm}^{-1}$ and the sample was stored at -20 °C after fast freezing in liquid nitrogen.

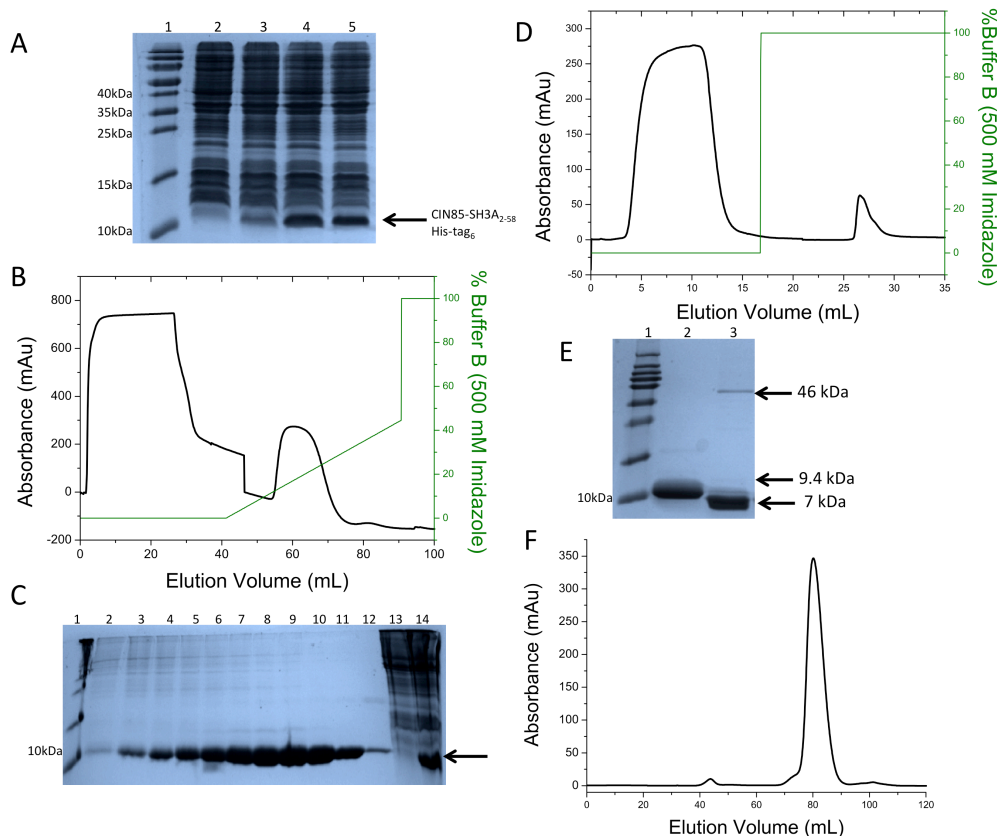


Figure IV.1. Summary of chromatograms and SDS-PAGE obtained during the overexpression and purification of the ^{15}N - ^{13}C -labeled SH3A domain of CIN85. **A.** SDS-PAGE of the best overexpression-test conditions: lane 1, marker (Bio-Rad Life Science); lane 2, cell culture before induction with IPTG; lane 3, cell culture 6 hours after induction; lane 4, cell culture 18 hours after induction; and lane 5, cell culture 24 hours after induction. The arrow indicates the position of CIN85-SH3A₂₋₅₈His-tag₆ (9.4 kDa). **B.** First Ni-column step, including the gradient from 20 to 500 mM Imidazole, CIN85-SH3A₂₋₅₈His-tag₆ was eluted in a single peak with a maximum at 18.6% of buffer B. **C.** SDS-PAGE of the fractions of the peak eluted in the first Ni-column step. Lane 1, marker (Bio-Rad Life Science); Lanes 2 to 12, fractions eluted in the peak that appears in figure IV.1B with maximum at 18.6% of buffer B; lane 13, supernatant of the cell lysate after centrifugation (BECKMAN COULTER, JA-20 rotor); lane 14, pellet after the centrifugation of the cellular lysate (showing that all the protein was in inclusion bodies). The arrow indicates the position of CIN85-SH3A₂₋₅₈His-tag₆ (9.4 kDa). **D.** Second Ni-column step after the cleavage of the His-tag with the enzyme PreScission protease to separate the cleaved SH3A domain (7.06 kDa) from the uncleaved one (9.4 kDa). **E.** SDS-PAGE of the His-tag cleavage using PreScission protease: lane 1, marker (Bio-Rad Life Science); lane 2, fraction eluted from the Ni-column (lane 11 in the SDS-gel IV.1C) before His-tag cleavage (9.4 kDa); lane 3, the same fraction of the SH3A domain of CIN85 after His-tag cleavage (7.06 kDa), the line at 46 kDa correspond to the PreScission protease. **F.** Superdex-75 16/60 chromatogram, showing two peaks, one around 42 mL corresponding to the PreScission protease and another at 83 mL corresponding to the SH3A domain of CIN85 (7.06 kDa).

4.2.2 Atypical proline-rich peptides

Peptides containing stretches of 13 residues including the atypical PxxxPR recognition sequence for CD2 (residue 322 to 334) and Cbl-b (residues 901 to 913) and mutants of them (see Table IV.1) at more than 98% purity were purchased from ChinaPeptides Co., Ltd (Shanghai, China) and SynBioSci Co. (San Francisco, US). A tyrosine residue was added at the N-terminus to ensure proper determination of the concentration spectrophotometrically ($\epsilon_{280} = 1,490 \text{ M}^{-1} \text{ cm}^{-1}$). All peptides were N-terminal acetylated and C-terminal amidated. The purity of the peptides was confirmed by mass spectrometry.

Before usage in the different biophysics experiments, the proper amount of lyophilized peptide was weighted using an analytical balance (precision: $\pm 0.1 \text{ mg}$) and dissolved in the proper buffer, with more than 24 hours in advance to ensure possible conformational changes as sometimes observed for polyproline peptides. When necessary, peptides were dialyzed using mini dialysis membranes with a 500 Da cut-off (GE Healthcare).

Peptide	Sequence
CD2	Y ₋₃₂₂ QQK GPPLPRPRV Q ₃₃₄
CD2R332A	Y ₋₃₂₂ QQK GPPLPRP<u>A</u> VQ ₃₃₄
Cbl-b	Y ₋₉₀₁ APAR PPKPRPRRT ₉₁₃
Cbl-bR904G	Y ₋₉₀₁ APAG PPKPRPRRT ₉₁₃
Cbl-bR911A	Y ₋₉₀₁ APAR PPKPRP<u>A</u> RT ₉₁₃
Cbl-bK907L	Y ₋₉₀₁ APAR PP<u>L</u> PRPRRT ₉₁₃

Table IV.1. Amino acid sequences of CD2 and Cbl-b peptides used in this study. The prolines and arginines of the atypical consensus PxxxPR motif are emphasized in bold, and residues mutated are underlined. The N-terminal and C-terminal residues of the sequence are indicated as subscripts. An N-terminal tyrosine (Y) was added to all peptides for concentration measurement purposes.

4.2.3 Isothermal Titration Calorimetry

ITC experiments (Chapter III, section 3.2.1) were carried out on a VP-ITC instrument (MicroCal Inc., GE Healthcare). During most experiments, the SH3A domain of CD2AP and CIN85 (40 - 90 μM) were titrated with the different peptides (0.6 - 1.9 mM) while during the inverse experiments, the different peptides (15 - 50 μM) were titrated with the SH3A domains (1 mM), in 50 mM cacodylate buffer pH 6.0 at 20 °C. Additionally, SH3A of CD2AP (50 μM) was titrated with CD2 peptide at 1 mM in different buffers (50 mM MES and 50 mM cacodylate) at pH 6.0, and at different temperatures (15, 20, 25, 30 °C). Most titrations consisted of a preliminary 3 μL injection (neglected in data analysis), followed by 30 subsequent 6 μL injections. In some cases we used a variable volume profile titration of 30 injections based on simulations (Chapter III, section 3.2.1). Heats of dilution, determined by titrating the peptides (and the SH3A domains, in the case of the inverse titrations) into buffer alone, were subtracted from the raw titration data before analysis. Data were fitted by least-

squares procedures assuming a n identical and independent binding site model (Chapter III, section 3.2.2) using MicroCal Origin version 7.0 (MicroCal, GE Healthcare). All the ITC experiments were repeated three times and errors were calculated as the standard deviation from the mean value.

4.2.4 Nuclear Magnetic Resonance Spectroscopy

^{15}N -labeled and ^{15}N - ^{13}C -labeled samples of the SH3A domains of CD2AP and CIN85 were prepared for NMR experiments in a range of concentration between 0.3 and 1 mM in 50 mM NaPi at pH 6.0 containing 10% (v/v) D_2O . All NMR spectra were recorded at 25 °C on a Varian NMR Direct-Drive Systems 600 or 800 MHz spectrometer, the latter equipped with a salt tolerance triple-resonance PFG-Z cold probe. The backbone resonances of CIN85-SH3A were assigned as described (Chapter III, section 3.3.3). CD2AP-SH3A was previously assigned (Roldan J.L. et al., 2011).

^{15}N -labeled SH3A domains (of CD2AP and CIN85 respectively) were titrated with increasing amounts of the different peptides (stock solution of 6.3 mM) (see Table IV.1) at pH 6.0 and at 25 °C. The progresses of the titrations were monitored by recording 2D ^1H - ^{15}N -HSQC spectra. The magnitude of the chemical shift perturbations ($\Delta\delta$) was calculated as previously described (Chapter III, section 3.3.3.3).

To evaluate the titration experiments, the ^1H - ^{15}N chemical shifts of the indole moiety from Trp36 of CIN85-SH3A and Trp37 of CD2AP-SH3A were plotted against the ligand concentration and fitted to an equation taking into account the Hill coefficient for possible cooperative effects (Chapter III, section 3.3.3.3), and from which it is possible to calculate the apparent K_d of the interaction.

The K_d s were used to calculate the fraction of free and bound SH3A domain at each titration point considering that the protein can recognize the two different binding sites of the ligand (Cbl-b, type I and type II interaction) forming hetero-dimers in an independent way, and taking into account that the most favorable interaction (smaller K_d) occurs first (type II).

$$K_{dII} = \frac{[\text{SH3A}][\text{Cblb}]}{[\text{SH3A-Cbl-b}_{II}]} = \frac{(\text{SH3A} - x)(\text{Cblb} - x)}{x} \rightarrow$$

$$\rightarrow x = \frac{(\text{K}_{dII} + [\text{SH3A}] + [\text{Cblb}]) - \sqrt{(\text{K}_{dII} + [\text{SH3A}] + [\text{Cblb}])^2 - 4 \cdot [\text{SH3A}][\text{Cblb}]}}{2}$$

$$K_{dI} = \frac{([\text{SH3A}] - x)[\text{Cblb}]}{[\text{SH3A-Cbl-b}_I]} = \frac{(\text{SH3A} - x - y)(\text{Cblb} - y)}{y} \rightarrow$$

$$\rightarrow y = \frac{(\text{K}_{dI} + [\text{SH3A}] - x + [\text{Cblb}]) - \sqrt{(\text{K}_{dI} + [\text{SH3A}] - x + [\text{Cblb}])^2 - 4 \cdot ([\text{SH3A}] - x)[\text{Cblb}]}}{2}$$

where

$$x = [\text{SH3A-Cbl-b}_{II}]$$

$$y = [\text{SH3A-Cbl-b}_I]$$

$$[\text{SH3A}] = [\text{SH3A-Cbl-b}_{II}] + [\text{SH3A-Cbl-b}_I] + [\text{SH3A}]_{\text{free}}$$

This analysis was used to simulate the chemical shift perturbations of the Trp indole H ϵ 1 of the SH3A domain of CD2AP and CIN85 during the NMR titration with Cbl-b peptide and the mutants Cbl-bR904G (isolated type II interaction) and Cbl-bR911A (isolated type I interaction).

2D-NOESY spectra (mixing times of 150 ms) were recorded at the end of every NMR titration to check the intermolecular NOE pattern of the indole proton of Trp36 of CIN85-SH3A and of Trp37 of CD2AP-SH3A. For this purpose the CD2 proton resonances were assigned from a combination of 2D ^1H - ^1H TOCSY (mixing time of 20 and 70 ms), 2D ^1H - ^1H ROESY (mixing time of 200 ms) and 2D ^1H - ^1H NOESY (mixing time of 200 ms) experiments. A 0.8 mM CD2 solution was titrated with increasing amounts of unlabeled CD2AP-SH3A domain (stock solution of 5.8 mM). The progress of the titrations was monitored by recording 2D ^1H - ^1H TOCSY (mixing time of 70 ms) and 2D ^1H - ^1H NOESY (mixing time of 200 ms) spectra. Proton chemical shifts for CD2 were assigned at each titration point.

All NMR data were processed using NMRPipe (Delaglio F. et al., 1995) and analyzed in CCPNMR (Vranken W.F. et al., 2005) or NMRview (Johnson B.A., 2004).

4.2.5 Small Angle X-Ray Scattering

SAXS data for characterizing the different complexes formed between SH3A domains of CD2AP and CIN85 with CD2 and Cbl-b peptides, were collected at the beamline X33 of the DESY synchrotron (Hamburg, Germany). The camera length was 2.7 m and the wavelength 1.5 Å, with 2 min of exposure time for data collection. The processing and analysis of the data were done as described (Chapter III. Section 3.4). Samples of CD2AP-SH3A were prepared in absence and presence of the peptides CD2 and Cbl-b (1:1 and 1:1.6 ratios for both peptides, one at each time). Samples of CIN85-SH3A were also prepared in absence and presence of CD2 and Cbl-b (1:0.5 and 1:1.6 ratios for both peptides, one at each time), and in presence of Cbl-bK907L in 1:0.5 ratio. All the samples were measured at four different concentrations: 7, 5, 3 and 1 mg/mL in 50 mM NaPi buffer at pH 6.0.

4.3 Results and discussion

4.3.1 Thermodynamic characterization of the interaction of CD2AP-SH3A with CD2

To properly characterize the interaction of the SH3A domain of CD2AP with its natural target CD2 from a thermodynamic point of view, we carried out a complete set of ITC experiments to calculate the stoichiometry (n), the dissociation constant (K_d), Gibbs energy (ΔG_b), entropy (ΔS_b) and enthalpy (ΔH_b) of the protein-peptide interaction. The experiments were performed in different buffers (MES and cacodylate) and at different temperatures (15, 20, 25, 30 °C), to calculate the contribution from the de-solvation to the binding process (Murphy K.P. and Freire E., 1992; Gomez J. et al., 1995) and the change of heat capacity (ΔC_p), respectively (Freire E. et al., 1990).

The mathematic analysis of the thermograms obtained during the titrations, including the definition of the baseline, subtraction of the dilution heats, integration of the thermogram peaks and fitting of the net heats, were done with the software Origin 7.0 of MicroCal (Chapter III, section 3.2).

The ITC thermograms performed at different temperatures and with different buffers are shown in figure IV.2.

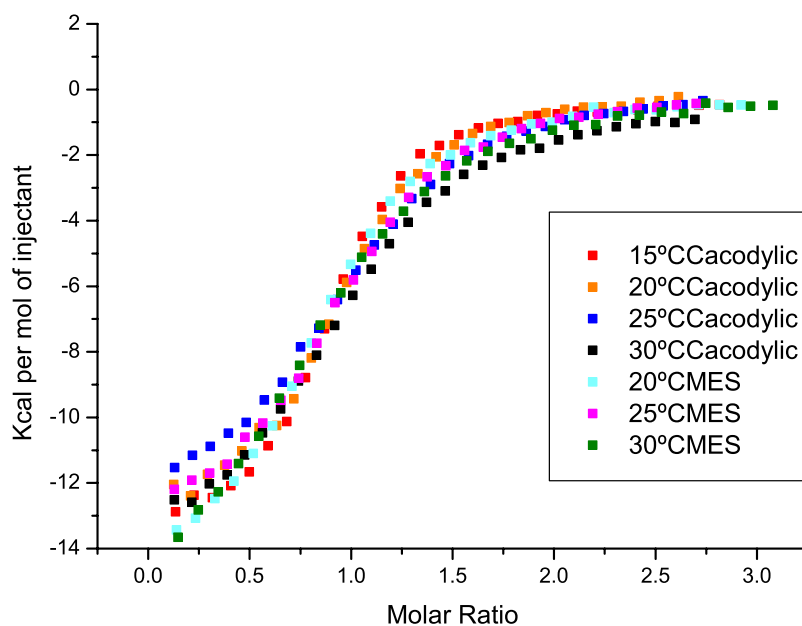


Figure IV.2. Representation of the individual net heat per mol of injectant versus the molar ratio, obtained during the ITC experiments of the SH3A domain of CD2AP (50 μ M) with the peptide CD2 (1 mM) at different temperatures (from 15 to 30 °C) and in different buffers (50 mM cacodylate and 50 mM MES at pH 6.0).

Buffers	T(°C)	n	$K_d(\mu M)$	ΔH_b (kcal/ mol)	ΔG_b (kcal/ mol)	$-T \cdot \Delta S_b$ (kcal/mol)
Cacodylate Buffer 50 mM, pH6.0	15	1.0 ± 0.1	2.7 ± 0.2	-13.9 ± 0.3	-7.3 ± 0.2	6.6 ± 0.3
	20	1.0 ± 0.1	3.1 ± 0.2	-13.2 ± 0.4	-7.3 ± 0.1	5.9 ± 0.4
	25	1.0 ± 0.1	4.6 ± 0.3	-13.7 ± 0.3	-7.3 ± 0.2	6.3 ± 0.4
	30	1.0 ± 0.1	8.0 ± 0.4	-13.1 ± 0.5	-7.1 ± 0.4	6.1 ± 0.5
MES Buffer 50 mM, pH6.0	20	1.0 ± 0.1	4.2 ± 0.3	-13.0 ± 0.4	-7.2 ± 0.4	5.9 ± 0.4
	25	1.1 ± 0.1	7.0 ± 0.4	-13.0 ± 0.5	-6.5 ± 0.3	6.6 ± 0.5
	30	1.0 ± 0.1	7.3 ± 0.4	-12.5 ± 0.5	-7.1 ± 0.3	6.8 ± 0.5

Table IV.2. Thermodynamic parameters obtained from the fitting of the isotherms (Fig. IV.2) acquired during the titration of the SH3A domain of CD2AP (50 μM) with the peptide CD2 (1 mM) at different temperatures and in different buffers. The reported values are averaged over three experiments and the errors are calculated as the standard deviation from the mean values.

The analysis of the parameters obtained from the fitting of the experimental data (Table IV.2), indicates that the stoichiometry of the interaction is $n = 1$ independently of the buffer and the temperature in which the analysis were performed. The data shows that the binding enthalpy (ΔH_b) is negative, indicative of an exothermic process.

In order to fully resolve the binding thermodynamics and to check the suitability of the binding model used to fit the data, we also carried out the inverse experiments. For 1:1 biomolecular reactions, like is the case here, it is expected that the measured thermodynamic parameters are invariant when changing the orientation of the experiment (Velazquez-campoy A. and Freire E., 2006).

The inverse titration experiments were done in 50 mM cacodylate buffer at pH 6.0, and at different temperatures (from 15 to 30 °C). Data are shown in figure IV.3.

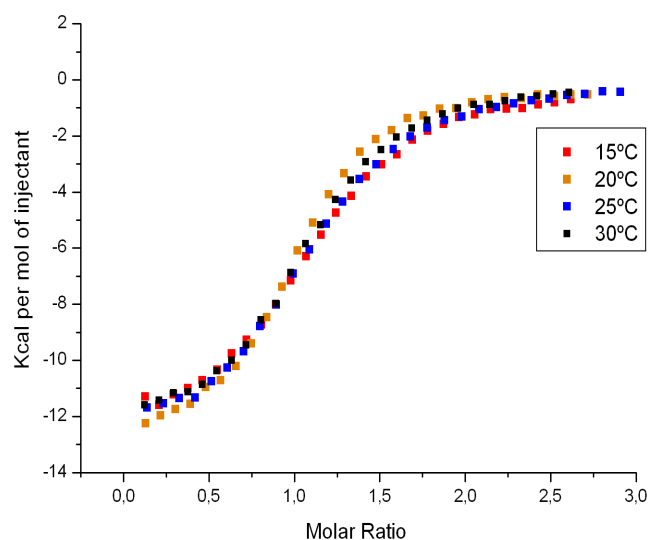


Figure IV.3. Representation of the individual net heat per mol of injectant versus the molar ratio, obtained during the inverse ITC experiments of the peptide CD2 (50 μM) with the SH3A domain of CD2AP (1 mM) at different temperatures (from 15 to 30 °C) in 50 mM cacodylate buffer at pH 6.0.

Cacodylate buffer 50 mM, pH 6.0					
T(°C)	n	K _d (μM)	ΔH _b (kcal/mol)	ΔG _b (kcal/mol)	-T·ΔS _b (kcal/K·mol)
15	1.0 ± 0.1	2.8 ± 0.2	-13.1 ± 0.3	-7.3 ± 0.3	5.8 ± 0.3
20	1.0 ± 0.1	3.4 ± 0.3	-13.2 ± 0.4	-7.3 ± 0.2	5.9 ± 0.4
25	1.1 ± 0.1	4.4 ± 0.4	-13.1 ± 0.4	-7.3 ± 0.3	5.7 ± 0.4
30	1.0 ± 0.1	5.9 ± 0.5	-13.0 ± 0.4	-7.2 ± 0.2	5.8 ± 0.4

Table IV.3. Thermodynamic parameters obtained from the fitting of the isotherms (Fig. IV.3) acquired during the inverse titration of the peptide CD2 (50 μM) with the SH3A domain of CD2AP (1 mM) at different temperatures (from 15 to 30 °C) in 50 mM cacodylate buffer at pH 6.0. The reported values are averaged over three experiments and the errors are calculated as the standard deviation from the mean values.

The thermodynamic parameters (Table IV.2 and Table IV.3) obtained from the direct and the inverse titrations fit perfectly with a 1:1 stoichiometry characteristic of a single binding event. The results indicate that, independently of the concentration of protein and peptide, temperature and protonation heat of the buffer, the SH3A of CD2AP interacts with the peptide CD2 forming hetero-dimers, in which one molecule of CD2AP-SH3A domain interacts with one molecule of CD2 peptide. In both type of experiments, the binding enthalpy (ΔH_b) is negative ($\langle \Delta H_b \rangle = -13.3 \pm 0.7$ kcal/mol), indicative of an exothermic process and it is partially compensated by an unfavorable entropic (-TΔS_b) positive term ($\langle -T \cdot \Delta S_b \rangle = 6.0 \pm 0.8$ kcal/K·mol), showing that the binding interaction is enthalpically driven. Binding enthalpy reflects the strength of the interactions with the ligand (van der Waals, hydrogen bonds, and so on) relative to those existing with the solvent (Levitt S. and Freire E., 2001).

On the other hand, the entropic change upon binding ($\langle -T \cdot \Delta S_b \rangle = 6.0 \pm 0.8$ kcal/K·mol) is difficult to interpretate, because the total entropic contributions associated with binding are the sum of three terms: the solvation entropy, that describes the change in entropy resulting from solvent release upon binding; the conformational entropy, that reflects the reduction of rotational degrees of freedom around torsion angles of protein and ligand; and the roto-translational entropy, that entails the loss of translational and rotational degrees of freedom when a complex is formed from two molecules free in solution. The most important contribution to entropy change arises from the solvation term, primarily due to burial of apolar surface area. The solvation entropy used to be positive because it represents the gain of freedom degrees of the water molecules that previously were on the surface of the interacting molecules, and that are released upon binding. However, the conformational and roto-translational entropies, used to be negative due to the loss of degrees of freedom of the interacting molecules after binding.

Previous studies on the binding of SH3 domains to proline-rich ligands have shown similar compensations between the enthalpic and the entropic terms (Wittekind M. et al., 1994; Renzoni D.A. et al., 1996; Ferreon J.C. and Hilser V.J., 2003 and 2004), presenting K_ds in the same order of magnitude and enthalpy values from -5 to -12 kcal/mol, comparable to the values obtained with the SH3A domain of CD2AP. The association is enthalpy-driven contrarily to what is usually expected for a highly hydrophobic interaction, as is the case for SH3-peptide binding. Several studies have been reported to explain this discrepancy (Palencia A. et al., 2004; Ferreon J.C. and

Hilser V.J., 2004). Existence of water molecules buried at the binding interface or the propagation of binding interaction throughout the SH3 domain must be considered (Palencia A. et al., 2004; Casares S. et al., 2007). Furthermore, other authors pointed out that the folding of the peptide into the binding competent conformation has a quantitative impact on the observed energetics (Ferreon J.C. and Hilser V.J., 2004). All this indicates that SH3-peptide interaction is not a simple binding process from a thermodynamic point of view.

Using the values of ΔH_b obtained at different temperatures, we have calculated (Chapter III, section 3.2) the change in heat capacity (ΔC_p) by the slope of the linear regression analysis of ΔH_b plotted vs the temperature (Fig. IV.4).

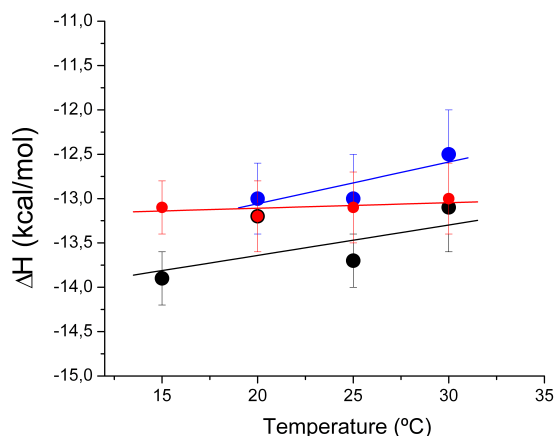


Figure IV.4. Temperature dependence of the binding enthalpies of CD2AP-SH3A with CD2 (direct titration in buffer cacodylate in black, direct titration in buffer MES in blue and inverse titration in buffer cacodylate in red). Values of binding enthalpies are reported in Table IV.2 and IV.3. The data were fitted to straight lines with the slopes corresponding to ΔC_p of binding.

For the direct experiments in cacodylate and in MES buffer, $\Delta C_p = 0.04 \pm 0.03$ kcal/mol·K and $\Delta C_p = 0.05 \pm 0.03$ kcal/mol·K were estimated, respectively. In both cases, the values are close to zero. This is also the case for the inverse experiments, where a $\Delta C_p = 0.006 \pm 0.003$ kcal/mol·K (in cacodylate buffer) was estimated. In general, there is a strong correlation between ΔC_p and the surface area buried upon complex formation (Murphy K.P. and Freire E., 1992; Gomez J. et al., 1995; Jelesarov I. and Bosshard H.R., 1999). It has been shown that the removal of protein surface area from the contact with solvent results in a large negative ΔC_p , higher than 1 kcal/mol·K in absolute value (Velazquez-Campoy A. et al., 2004). Through this correlation of ΔC_p and burial of surface area, the heat capacity provides a link between thermodynamic data and structural information. The fact that for the interaction of CD2AP-SH3A with CD2, the ΔC_p is close to zero, is indicating that the net balance of buried surface area of the SH3 domain is almost negligible.

There are often linked protonation-deprotonation reactions accompanying binding. In order to determine whether such protonation reactions are linked to the binding reaction, ITC experiments should be carried out in several buffers with different protonation enthalpies. If a net change in protonation occurs, a corresponding change in buffer protonation will add to the observed enthalpy (Leavitt S. and Freire E., 2001). We have performed the binding experiments in two different buffers at pH 6.0 (enthalpy of deprotonation, ΔH_i at 25 °C in brackets): MES ($\Delta H_i = 3.71$ kcal/mol) and cacodylate ($\Delta H_i = -0.47$ kcal/mol). The observed binding constants and enthalpies, were found to be buffer independent (within the error of the experiment), indicating that no protonation-deprotonation processes are coupled to the binding, as the binding enthalpy is independent of the ionization enthalpy of the buffer in which the reaction takes place (Table IV.2). Other SH3 domains, for example, the SH3 domains of Abl-tyrosin kinase

and of Spectrin (Spc), neither show influence of coupled ionization equilibria upon binding to polyproline peptides (Palencia A. et al., 2004; Casares S. et al., 2007).

The values of the dissociation constant (K_d) during the direct and inverse ITC experiments are in the micromolar range, in agreement with the typical K_d values reported for other peptides interacting with SH3 domains, which are in the range of 1 - 100 μ M (Kay B.K. et al., 2000). In table IV.2 we can observe that when the temperature increases, there is a slightly increase in the K_d values. This is due to the fact that the K_d is a function of the temperature.

4.3.2 NMR titration of CD2AP-SH3A with CD2

While ITC is likely the best method to obtain the thermodynamic parameters of a binding event, analysis of the NMR chemical shift perturbations is a widely used method to map protein interfaces. The ^1H - ^{15}N -HSQC spectrum of a protein is monitored upon titration with the unlabeled interaction partner. The interaction causes environmental changes on the protein interfaces that affect the chemical shifts of the nuclei in this area (Chapter III, section 3.3.3.3).

The NMR titration of ^{15}N -labeled CD2AP-SH3A with the unlabeled peptide CD2, followed by ^1H - ^{15}N -HSQC experiments (Fig. IV.14), shows that the most significant changes in the chemical shift are observed in the RT loop (residues 13, 16, 18), the n-Src loop (residues 30, 32, 34, 35, 36), the β -III (residues 39, 40), β -IV (residues 48, 49) and in the 3_{10} helix (residues 51 - 54). Specially affected is the indole $\text{N}\epsilon_1\text{-H}\epsilon_1$ of Trp37, one of the most conserved residues in SH3 domains (Fernandez-Ballester G. et al., 2003), which shows, like the rest of the affected residues, a linear shift change. Linear shift changes are indicative of a simple two-state binding event (Zuiderweg E.R., 2002), in agreement with the 1:1 stoichiometry found by ITC.

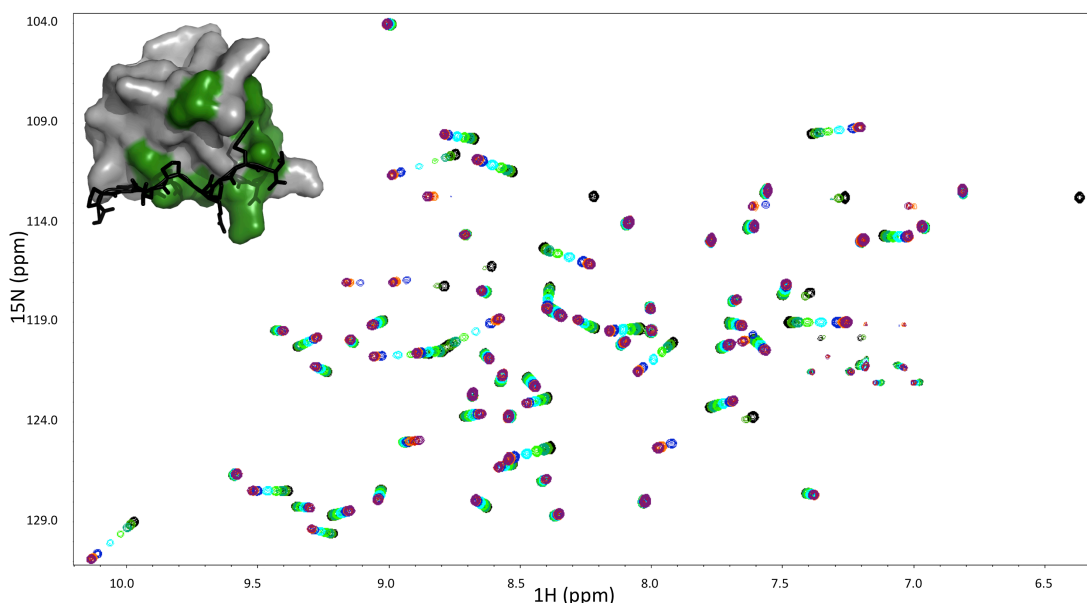


Figure IV.5. Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the NMR titration of the ^{15}N labeled SH3A domain of CD2AP, with the unlabeled CD2 peptide. Most of the affected residues are in intermediate to fast exchange, and show linear trajectories. Spectra of free SH3A in black, CD2-bound SH3A in violet (final titration point corresponds to a 2.5-fold excess of

CD2). The surface representation of CD2AP-SH3A engaging the CD2 peptide (black sticks) in a type II orientation (PDB code 2J7I). The most affected residues ($\Delta\delta$ /average more than 1) of CD2AP-SH3A appear in green (see Fig. IV.14).

Figure IV.5 shows an overlap of the ^1H - ^{15}N -HSQC spectra recorded during the titration of labeled SH3A domain of CD2AP with the peptide CD2. The ^1H - ^{15}N -HSQC spectra is a fingerprint of the protein, and the chemical shift changes of the residues during the NMR titration are indicative of environmental changes of the residues of SH3A involved in binding with the peptide CD2. The profile of the NMR titration shows intermediate to fast exchange, indicative of intermediate to weak binding. This is in agreement with the micromolar range of K_d values obtained by ITC, and in concordance with the reported affinities of other SH3 domains interacting with proline-rich peptides (Kay B.K. et al., 2000).

Identifying the type II interaction of CD2AP-SH3A with CD2

The results from the ITC and NMR titrations of CD2AP-SH3A with CD2 show that they interact forming a hetero-dimeric complex, with intermediate-weak affinity. However, since the SH3 domains can interact with proline motifs in two different modes, type I and type II (Chapter I, section 1.4), we used NMR to determine if the dimers that CD2AP-SH3A forms with CD2 were type I or type II.

2D NOESY spectra of free CD2AP-SH3A and the CD2AP-SH3A:CD2 complex were compared (Fig. IV.6), showing that in the complex NOESY additional NOEs involving the indole H ϵ 1 proton of Trp37 can be assigned to different protons in the CD2 peptide belonging to Arg330, Pro331 and Arg332. These intermolecular NOEs can only be explained in the crystal structure with the SH3 domain binding CD2 in a type II mode where all the protons are within 4.5 Å distance of the Trp37 H ϵ 1 proton. In the case of a type I binding, most of these protons, except those of Arg330, are as far as 6.5 to 9.5 Å and thus would not be observable in a NOESY spectrum.

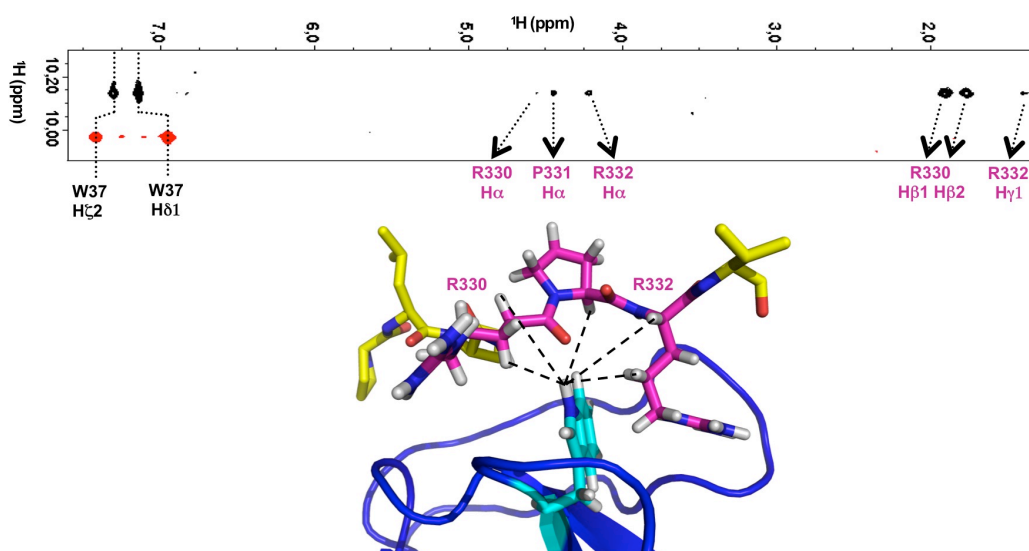


Figure IV.6. Intermolecular NOEs between CD2 peptide and CD2AP-SH3A (cyan). Selected region from a 2D NOESY spectrum of a 1:1 ratio CD2:SH3A sample taken around the ^1H frequency of the H ϵ 1 proton of Trp37 in SH3A. For comparison the same region in the 2D NOESY spectrum of free SH3A is plotted in red contours. Intra-residue NOEs to the H δ 1 and

$H\zeta$ protons are found in both the complex and the free form spectra. Only the NOESY of the complex shows additional NOEs that can be assigned to different protons in the CD2 peptide indicated by arrows and labeled according to their assignments. These intermolecular NOEs can only be rationalized in the crystal structure (lower panel, PDB entry 2J6O) when the SH3 domain binds CD2 in a type II orientation.

Moreover, eliminating the arginine of CD2 responsible for the type II interaction by mutating it into an alanine (mutant CD2-R332A, see Table IV.1) leads to a very small linear downfield shift change of all the residues (also of the Trp37 $N\epsilon$ - $H\epsilon$) upon addition of a large excess of peptide (the last point of the titration corresponds to 7-fold excess of peptide) indicating very weak residual type II binding. A K_d of more than 2 mM was estimated from these chemical shift changes of the most affected residue Asn52 (see Fig. IV.7).

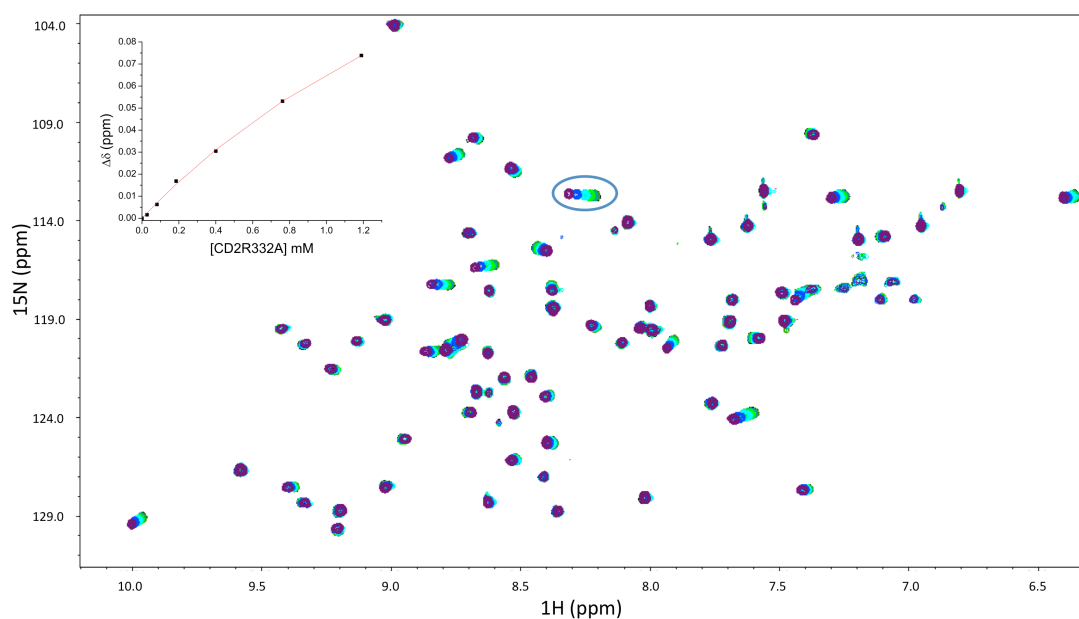


Figure IV.7. Overlay of the 1H - ^{15}N -HSQC spectra recorded during the NMR titration of CD2AP-SH3A with the mutant CD2-R332A peptide. Titration was done in 50 mM NaPi at pH 6.0 and 25 °C, on an 800 MHz spectrometer. The insert shows the determination of the binding constant by fitting the shift changes ($\Delta\delta$) of the most affected residue Asn52 (surrounded by a blue circle) against the ligand concentration, using a quadratic equation (Chapter III, section 3.3.3.3). Final titration point corresponds to a 7-fold excess of peptide. Spectra of free SH3A in black and CD2R332A-bound SH3A in violet.

These results, together with the intermolecular NOEs, confirm the importance of the Arg332 of the peptide CD2 in the type II interaction with the SH3A domain of CD2AP and excluded the formation of a type I dimer.

4.3.3 Characterization of the interaction of CD2AP-SH3A with Cbl-b and mutant peptides using ITC and NMR

It was previously reported that CD2AP also recognizes the atypical proline-rich sequence present in Cbl-b (Kowanetz K. et al., 2003), crucial for Cbl-directed RTK downregulation (Dikic I., 2002) (see section 4.1). In order to complement the study of the recognition mode of atypical-proline sequences by the SH3A domain of CD2AP, we also carried out the thermodynamic and structural characterization of the interaction with Cbl-b, using ITC and NMR.

The titration of CD2AP-SH3A domain with the Cbl-b peptide followed by ITC (Fig. IV.9A) represents an exothermic process enthalpically driven, as in the case of the interaction with the CD2 peptide. The fitting of the experimental isotherms, using a model of n identical and independent sites, gave a stoichiometry of 1:1, with a K_d of 0.19 μM , one order of magnitude lower than for the interaction with CD2 (Table IV.4), which indicates the formation of a hetero-dimeric complex with Cbl-b tighter than the hetero-dimeric complex with CD2.

The interaction of CD2AP-SH3A with both peptides is driven by a large and favourable enthalpy change (-13.2 ± 0.4 kcal/mol and -10.4 ± 0.2 kcal/mol for CD2 and Cbl-b, respectively), partially compensated by an unfavorable entropy change (5.9 ± 0.4 kcal/K·mol and 1.5 ± 0.2 kcal/K·mol for CD2 and Cbl-b, respectively). However, the entropic penalty in the interaction with Cbl-b is significantly smaller (Table IV.4), pointing to a more dynamic interaction or different conformations of the ligand in the bound state (Chang C.E et al., 2007). The small entropic term during the interaction with Cbl-b results in a gain of free energy (-7.3 ± 0.1 kcal/mol and -8.9 ± 0.1 kcal/mol for CD2 and Cbl-b, respectively) in agreement with the stronger interaction of Cbl-b with CD2AP-SH3A.

During the NMR titrations of ^{15}N -labeled CD2AP-SH3A with increasing amounts of Cbl-b peptide (Fig. IV.8), selective shift changes of the $^1\text{H}^{\text{N}}$ and ^{15}N resonances were observed for the same SH3A residues that were affected during the titration with CD2 (RT-loop, n-Src loop, β -III, β -IV and 3_{10} helix) (Fig. IV.14.D). These residues correspond to the typical proline binding area characteristic of SH3 domains (Larson S.M. and Davidson A.R., 2000; Cesareni G. et al., 2002; Ferreon J.C. and Hilser V.J., 2004). Again, most of the residues in the binding site were in a fast-to-intermediate exchange regime. In the figure IV.8 we can observe that most chemical shift changes display a linear pattern. Interestingly, the NMR titration shows an unusual curvature of the Ne1-He1 moiety of Trp37.

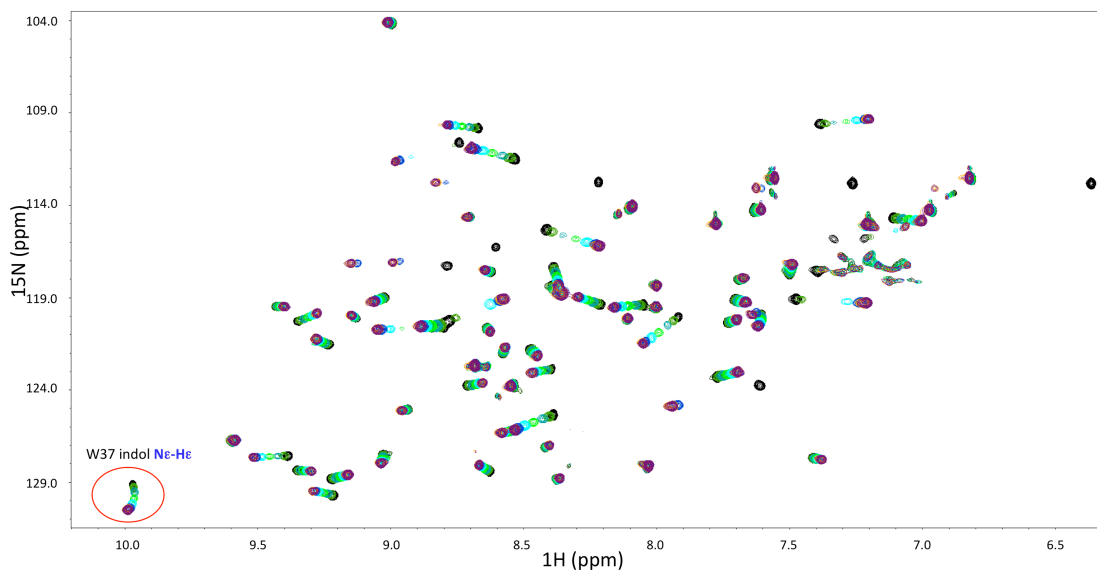


Figure IV.8. Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the NMR titration of the SH3A domain of CD2AP with the peptide Cbl-b. Titration was performed in 50 mM NaPi at pH 6.0 and 25 °C on an 800 MHz spectrophotometer. Spectra of free SH3A in black and Cbl-b-bound SH3A in violet (final titration point corresponds to a 2-fold excess of Cbl-b). The curve trajectory of the amide indole proton $\text{N}\epsilon\text{-H}\epsilon$ of the Trp37 is surrounded by a red circle.

The presence of curvature in HSQC spectra during titration (Fig. IV.8) is indicative of two or more binding events (Zuiderweg E.R., 2002). However, we did not observe extensive line-broadening during the titration, suggesting the absence of extensive conformational exchange phenomena, in agreement with the ITC titrations that showed a 1:1 stoichiometry. All these features indicate that, in solution, the predominant complexes are dimers formed by the interaction of one molecule of CD2AP-SH3A and one molecule of Cbl-b likely arranged in either type I or type II orientations. To confirm our hypothesis we looked for a possibility to knock out the interaction type I and type II separately that should result in linear titration patterns with opposite directions in the HSQC spectra.

The crystal structure of the SH3A domain of CIN85 with the peptide Cbl-b (Jozic D. et al., 2005) indicates that arginines 904 and 911 in Cbl-b are determinant for the type I and type II interactions, respectively. Based on that, we designed two mutants (see Table IV.1): Cbl-bR904G to eliminate type I interaction (considering that there is a glycine at the equivalent position in CD2) and Cbl-bR911A to eliminate type II interaction. If the curvature is indeed due to the formation of two different dimers (type I and II), elimination of one of the arginine crucial for the interactions (Arg904 or Arg911) should result in a linear shift change indicative of a simple one-site interaction, type I or type II binding (Zuiderweg E.R., 2002).

We characterized the interaction between CD2AP-SH3A with the mutants Cbl-bR904G and Cbl-bR911A using ITC and NMR (see Fig. IV.9). Panel E in figure IV.9 shows that for both mutants the curvature disappears: the isolated type II interaction (green contours) shows a linear downfield change for the $\text{N}\epsilon\text{1-H}\epsilon\text{1}$ of Trp37 and the isolated type I (blue contours) a linear upfield change for the same ^{15}N - ^1H shift.

Based on the comparison with published crystal structures, we noticed that the directionality of the shift change (downfield or upfield) is related to the ability of Trp37 N ϵ 1-H ϵ 1 of CD2AP-SH3A to form or not hydrogen bonds (H-bond) with the ligands. In the interaction type II with Cbl-bR904G (downfield change) Trp37 N ϵ 1-H ϵ 1 indeed makes a H-bond with Arg908-CO, which is absent (PDB entry **2J6F**) in the type I orientation (upfield change) with Cbl-bR911A. The downfield change observed for the Cbl-bR904G mutant is in agreement with the downfield change of the H ϵ 1 shift of Trp37 observed during the NMR titration with the peptide CD2 (Fig. IV.9, black contours). CD2 interacts with CD2AP-SH3A exclusively in a type II orientation (see section 4.3.1 and 4.3.2) forming a H-bond between the Trp37 N ϵ 1-H ϵ 1 of CD2AP-SH3A and the R330-CO of CD2 (PDB entry **2J7I**).

We also followed the titrations of CD2AP-SH3A with the mutants Cbl-bR904G and Cbl-bR911A by ITC, to confirm the 1:1 stoichiometry of the isolated type I and type II interaction and to thermodynamically characterize them. Figure IV.9 shows the thermograms obtained during the titration of CD2AP-SH3A with the different peptides, CD2 (**A**), Cbl-bR904G (**B**), Cbl-b (**C**) and Cbl-bR911A (**D**), fitted with a model of n identical and independent sites (red lines), and the region of the NMR spectra around the Trp37 indole N ϵ -H ϵ moiety (**E**), colored according to the color-code assigned to every peptide. The thermodynamic parameters derived from these fittings are summarized in Table IV.4. The ITC data show one-to-one stoichiometries for both mutants with a K_d of 0.71 μ M for the type II interaction and a K_d of 46.5 μ M for type I. It is the first time that the K_d s of the isolated type I and type II interaction of Cbl-b with the SH3A domain of CD2AP have been determined, showing that both types of binding are possible between the SH3A domain of CD2AP and Cbl-b, in contrast to the only type II binding with CD2. These novel results also highlight the crucial role of the Arg904 of Cbl-b in the type I interaction with the SH3A domain of CD2AP, never reported before.

Isolated type I and type II interactions of Cbl-b with CD2AP-SH3A are both enthalpically driven, with entropic penalties smaller than the one observed during the interaction with CD2. The value of the enthalpic term (Table IV.4) determined for the isolated type II interaction (mutant Cbl-bR904G, -11.9 ± 0.3 kcal/mol) is double than the value for the isolated type I (mutant Cbl-bR911A, -6.1 ± 0.5 kcal/mol). This is directly related with the different strength of the interactions, which is stronger in the case of the type II orientation, where the formation of H-bond between the Trp37 N ϵ 1-H ϵ 1 with the R908-CO of Cbl-b (absent in the type I interaction) contributes to the enhancement of the enthalpic value. The entropic term of the isolated type I interaction is much smaller than the one in the type II interaction (Table IV.4), so the principal contribution to the free energy is coming from the enthalpic term. The smaller value of the free energy for the type I interaction compared with type II is in agreement with the weaker affinity of the type I interaction.

PEPTIDES	n	K_d (μ M)	ΔH_b (kcal/ mol)	ΔG_b (kcal/ mol)	$-T \cdot \Delta S_b$ (kcal/mol)
CD2	1.0 \pm 0.1	3.1 \pm 0.2	-13.2 \pm 0.4	-7.3 \pm 0.1	5.9 \pm 0.4
Cbl-b	1.0 \pm 0.1	0.19 \pm 0.02	-10.4 \pm 0.2	-8.9 \pm 0.1	1.5 \pm 0.2
Cbl-bR904G	1.0 \pm 0.1	0.71 \pm 0.03	-11.9 \pm 0.3	-8.2 \pm 0.1	3.7 \pm 0.3
Cbl-bR911A	1.0 \pm 0.1	46.5 \pm 1.7	-6.1 \pm 0.5	-5.8 \pm 0.2	0.3 \pm 0.4
Cbl-bK907L	1.0 \pm 0.1	0.78 \pm 0.03	-13.1 \pm 0.4	-8.1 \pm 0.1	5.0 \pm 0.4

Table IV. 4. Thermodynamic parameters calculated from the fitting of the isotherms acquired during the titration of the SH3A domain of CD2AP (50 - 90 μ M) with the different peptides (1 - 1.9 mM), performed in 50 mM cacodylate buffer pH 6.0, at 20 °C in a VP-ITC of MicroCal. Values calculated using the model of n identical and independent sites. The reported values are averaged over three experiments and the errors are calculated as the standard deviation from the mean values.

We have determined that the interactions of CD2AP-SH3A with the peptides CD2 and Cbl-b are enthalpically driven (see section 4.3.1 and 4.3.3), and present a 1:1 stoichiometry. Two main differences have been found between both interactions: first, the affinity of CD2AP-SH3A for Cbl-b is one order of magnitude stronger than for CD2; and second, the entropic penalty of the interaction Cbl-b is significantly smaller than the one obtained for CD2. This small entropic value might be related with the dynamic nature of the interaction, since Cbl-b can interact with CD2AP-SH3A domain forming dimers type I and dimers type II, while the interaction with CD2 peptide only give rise to type II dimers.

There are some differences between the type I and type II dimers formed during the interaction of CD2AP-SH3A with Cbl-b. The K_d for the type I interaction is two orders of magnitude higher than for the type II, indicating a stronger affinity of the SH3A domain of CD2AP for the type II binding site of Cbl-b. This is in agreement with the double value of the enthalpy for the interaction with Cbl-bR904G in comparison with Cbl-bR911A, indicating the existence of more favorable interactions between Cbl-bR904G and the SH3A domain (like the H-bond between the Trp37 N ϵ 1-H ϵ 1 with the R908-CO of Cbl-b). The other important difference between type I and type II binding modes of Cbl-b can be observed in the NMR spectra (Fig. IV.9E). For the isolated type II interaction, the indole proton of the Trp37 presents a downfield shift change (related with the formation of H-bonds), but for the isolated type I interaction, the indole proton Trp37 presents an upfield shift change (related with the lack of H-bonds formation). We, therefore, propose that the linear downfield change is indicative of the formation of type II dimers and the linear upfield is indicative of the formation of type I dimers. In this sense, the curved pattern observed for the chemical shift changes of the Trp37 indole N ϵ -H ϵ moiety of SH3A during the interaction with Cbl-b wild type (WT) is likely caused by the type I (upfield change) and type II (downfield change) contribution.

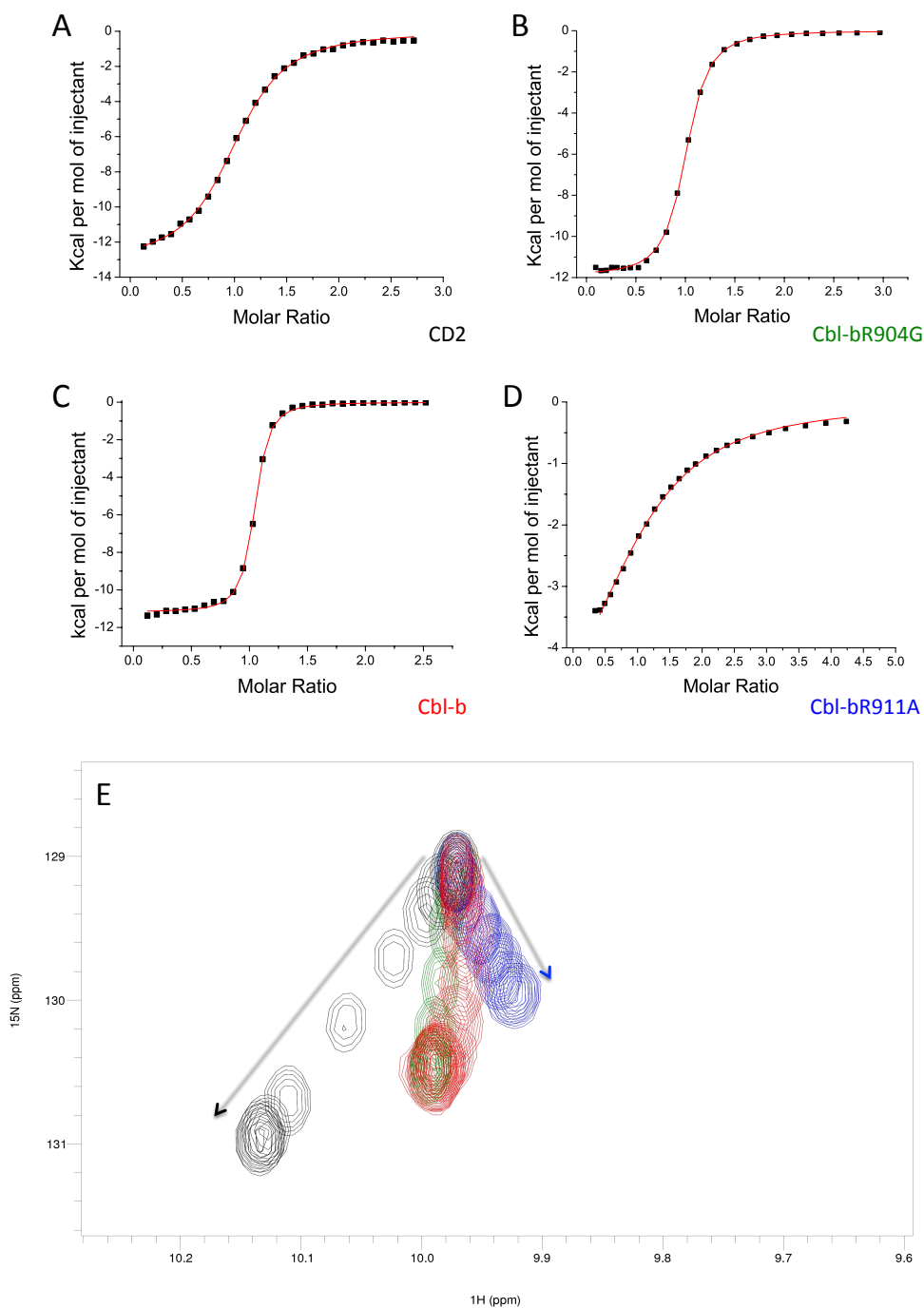


Figure IV.9. ITC and NMR titration of CD2AP-SH3A with CD2 and Cbl-b variants. **A-D.** ITC titration curves corresponding to titration with CD2 (**A**), Cbl-bR904G (**B**), Cbl-b (**C**) and Cbl-bR911A (**D**). Solid lines in the ITC show the best fit using the model of n identical and independent sites. For clarity each ITC titration curve is labeled with the peptide used in the titration and boxes of panels **A-D** are colored according to the colors used in the NMR titrations. **E.** Selected region of NMR spectra during titrations showing the chemical shift changes of the Trp37 indole N ϵ -H ϵ moiety. All spectra including those of the Cbl-b titration (red) are plotted at the same contour level. Colors used in the NMR spectra are black for CD2, red for Cbl-b, green for Cbl-bR904G and blue for Cbl-bR911A. All titrations start at the same chemical shift of the Trp37 N ϵ 1-H ϵ 1 in the free form. For clarity, the black and blue arrows indicate the direction of the chemical shift changes during the titration with CD2 peptide and mutant Cbl-bR911A, respectively.

To further characterize the curvature of the Trp37 indole proton observed during the titration with Cbl-b, we tried to simulate it by using the K_{ds} resulting from the fitting of chemical shift changes of the Trp indole $N\epsilon$ - $H\epsilon$ from the NMR titration of the isolated type I (with Cbl-bR911A) and type II (with Cbl-bR904G) interactions. The K_{ds} were used to determine the fraction of free SH3 and of bound type I and type II at every point of the titration assuming an independent binding-site model with preferential type II interaction (see section 4.2.4). These fractions were used to calculate the expected chemical shift changes of the Trp indole $H\epsilon 1$ of the SH3A domain of CD2AP during the titration with Cbl-b and mutants. As a proof-of-principle the experimental data of the isolated type II and type I interactions can be perfectly reproduced (Fig. IV.10, type II green line and type I blue line). Using this mathematical approach we have been able to simulate the curvature observed during the titration of the SH3A domain of CD2AP with Cbl-b WT (Fig. IV.10, red line), which confirms the existence of two independent types of hetero-dimers (type I and type II) in solution. This mathematical simulation supports the non-artifactual nature of the unusual curvature, since its existence can be predicted.

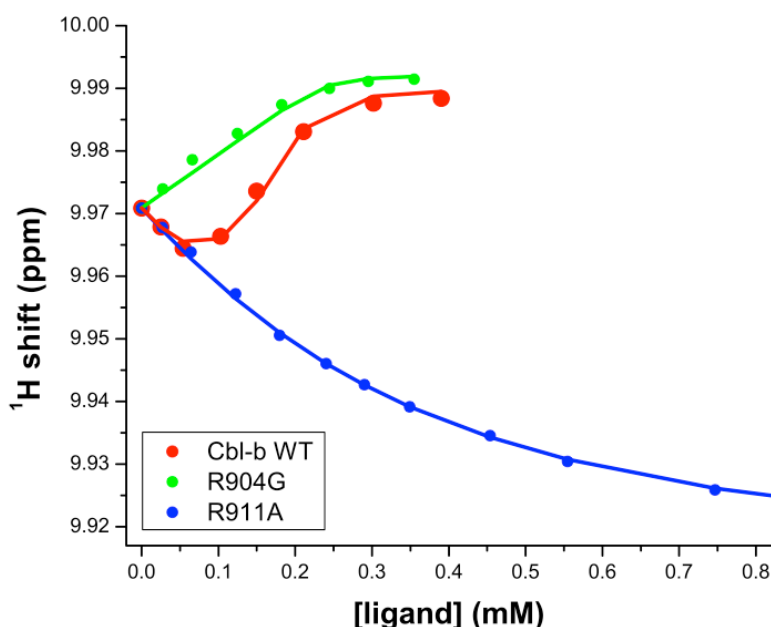


Figure IV.10. Representation of the 1H chemical shifts of the $H\epsilon 1$ from Trp37 of CD2AP-SH3A during the titration with Cbl-b (red), Cbl-bR911A (blue) and Cbl-bR904G (green). The chemical shifts of the $H\epsilon 1$ (dots) at each ligand concentration were simulated (solid lines) as described in the text. The changes in the chemical shift of the $H\epsilon 1$ during the titration with Cbl-b (red) can be rationalized by a combination of different contributions of the chemical shifts of free SH3A and the isolated type I and type II interactions represented by the Cbl-bR911A (blue) and Cbl-bR904G (green) mutants.

The graphic represented in figure IV.10 is very informative: at the beginning of the titration there is an important statistical contribution from both, type I and type II dimers, however at the end of the titration, Cbl-b basically follows the pattern of the isolated type II interaction (in green), which is going to be the predominant species in solution under saturation condition. This is in agreement with the K_{ds} obtained during the ITC experiments for the isolated type I and type II binding (Table IV.4) showing that the type II interaction is stronger than the type I. The high degree of similarity

observed between the intermolecular NOE patterns (Fig. IV.11) at the end of the NMR titration between the SH3A domain of CD2AP with the peptides CD2 and Cbl-b, also confirms that at the end of the titration, in both cases, the predominant species are dimers type II.

The NMR titration of CD2AP-SH3A with Cbl-b shows that the residues more affected (yellow bars in Fig. IV.14D) are exactly the same as those during the titration with CD2 (green bars in Fig. IV.14D). This suggests that at the end of the titration in both cases, a dimeric type II complex is the predominant form. These results are in concordance with the same NOESY patterns found at the end of every titration (Fig. IV.11).

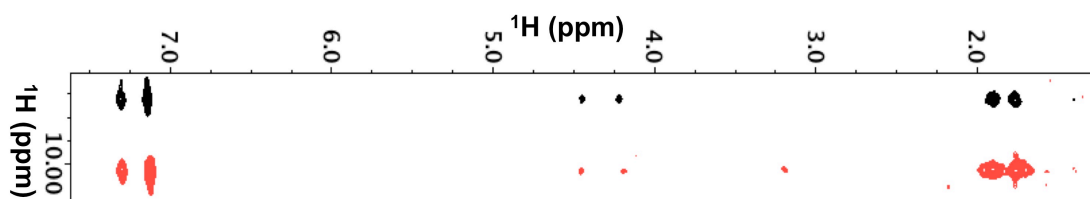


Figure IV.11. Intermolecular NOEs between CD2 and Cbl-b and CD2AP-SH3A. Selected region from a 2D NOESY spectrum at about 1:1.5 ratio SH3A:peptide sample, taken around the ^1H frequency of the $\text{H}\epsilon 1$ proton of Trp37 in SH3A. The same regions in the 2D NOESY spectra of SH3A with CD2 (black) and Cbl-b (red) are overlaid.

The previous results indicate that the CD2AP-SH3A domain is able to interact with the atypical proline-rich peptides CD2 and Cbl-b, in two different ways:

- Forming dimers type I and type II with Cbl-b, which contains two arginines, one in the N-terminal (Arg904) and another in the C-terminal (Arg911) region of the atypical rich sequence
- Forming dimers type II with CD2, which contains only one arginine in the C-terminal region of the atypical rich sequence (Arg332).

The aforementioned results suggest that the targets of CD2AP-SH3A lacking a N-terminal positively charged residue, as is the case of CD2 peptide, should bind solely as type II. However, when there is a positive charge in that position, like Arg904 in Cbl-b, the type I interaction is also possible. Based on these results, it is crucial to determine if this versatility in the interaction of the SH3A domain of CD2AP with atypical proline-rich targets is exclusively for the CD2AP adaptor protein or if it is characteristic of the CIN85/CD2AP family of adaptor proteins. Therefore, we also studied the interaction between the SH3A domain of the CIN85 adaptor protein and the CD2 and Cbl-b peptides.

4.3.4 Characterization of the interaction of CIN85-SH3A with CD2 by ITC and NMR

The structural and thermodynamic characterization of the interaction between CIN85-SH3A and CD2 had not been reported before. Like for CD2AP, ITC titration of the SH3A domain of CIN85 with CD2 (Table IV.5) showed a 1:1 stoichiometry, with a K_d of 9.9 μM , indicative of a dimeric complex with a similar affinity than the dimer formed between CD2AP-SH3A and CD2. Also the interaction of CIN85-SH3A with CD2 is an

enthalpically driven process ($\langle \Delta H_b \rangle = -9.9 \pm 0.4$ kcal/mol). This observation is consistent with the notion that enthalpy generally drives the interaction of SH3 domain with ligands (Wittekind M. et al., 1994; Arold S. et al., 1998; Ferreon J.C. and Hilser V.J., 2004; McDonald C.B. et al., 2009), and implies that both electrostatic and hydrophobic interactions play a crucial role in the assembly of the complex. As it is common in enthalpically driven processes, the binding is accompanied by an entropic penalty ($\langle -T \cdot \Delta S_b \rangle = 3.2 \pm 0.4$ kcal/K·mol) most likely due to the loss of degrees of motion of the interacting molecules when they are free in solution, but becoming more compact upon binding (McDonald C.B. et al., 2009).

Meanwhile, the NMR titration of ^{15}N -labeled CIN85-SH3A was used to map the residues affected during the interaction upon binding with CD2. The most significant changes in the chemical shift of the HSQC cross-peaks of CIN85-SH3A are observed in the RT loop (residues 8, 9, 11, 12, 13, 16, 18), the n-Src loop (residues 32, 33, 35, 36), the β -III (residues 37, 39), the β -IV (residue 48) and at the 3_{10} helix (residues 50 - 53) (Fig. IV.12 and Fig. IV.14A). The main differences between the SH3A domain of CD2AP and CIN85 during the titration with CD2, are found at the beginning of the RT loop (residues 8, 9, 11, 12), that is more affected in the case of the CIN85-SH3A.

As in the case of CD2AP-SH3A, the NMR titration of CIN85-SH3A with CD2 shows a linear downfield shift change of the Trp36 $\text{N}\epsilon$ - $\text{H}\epsilon$ moiety (Fig. IV.12), indicating the formation of a dimeric complex in a type II orientation, driven by the Arg332 of the peptide CD2, which interacts with the specific pockets of the SH3A domains. The downfield change of the Trp36 $\text{N}\epsilon$ - $\text{H}\epsilon$ moiety in the NMR titration (Fig. IV.12) is likely related to the formation of an H-bond, most probably with Arg330-CO in the type II orientation, as observed in the X-ray structure of CD2AP-SH3A with CD2 (PDB entry 2J7I).

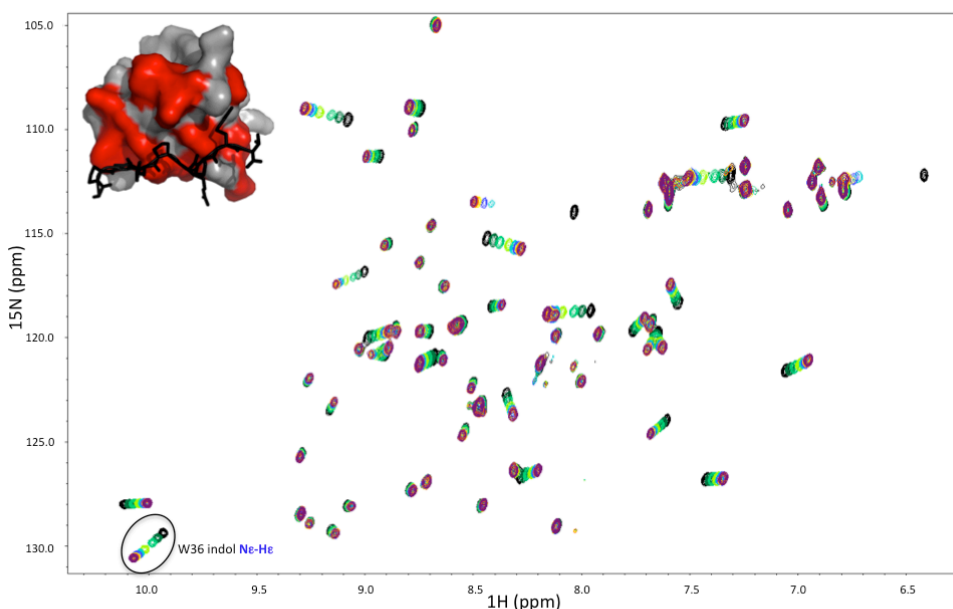


Figure IV.12. Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the NMR titration of the SH3A domain of CIN85 with the unlabeled CD2 peptide. Titration was done in 50 mM NaPi buffer at pH 6.0 and 25 °C, on an 800 MHz spectrophotometer. Most of the affected residues are in intermediate to fast exchange, and show linear trajectories. Inside of the spectra it is shown the surface representation of the SH3A domains of CIN85 engaging the CD2 peptide (black sticks) in a type II orientation (model based on the alignment of the X-ray structure of the

SH3A domain of CIN85 with the SH3A domain of CD2AP in complex with CD2, PDB code 2J7I). Residues more affected ($\Delta\delta$ /average more than 1) of CIN85-SH3A appear in red (see Fig. IV.14). Spectra of free SH3A in black and CD2-bound SH3A in violet (final titration point corresponds to a 2.5-fold excess of CD2). The trajectory of the amide indole proton N ϵ -H ϵ of the Trp36 is surrounded by a black circle.

This type II interaction was also confirmed using the mutant CD2-R332A in NMR titrations (as we did with CD2AP in section 4.3.2). Again small linear shift changes of the residues were observed upon the addition of 7-fold excess of CD2, indicating residual type II binding (with a K_d of more than 4 mM as estimated from the chemical shift changes). All these results show that both SH3A domains of CIN85 and CD2AP interact with CD2 (that lacks a positive residue in the N-terminal region) in the same manner, forming dimers predominately in type II orientation due to the presence of a C-terminal arginine in CD2 (Arg332).

4.3.5 Characterization of the interaction of CIN85-SH3A with Cbl-b and mutant peptides by ITC and NMR

To solve the discrepancy on the formation of hetero-trimers (described in section 4.1), we also carried out the thermodynamic and structural characterization of the interaction between CIN85-SH3A and Cbl-b and mutants.

ITC titration of CIN85-SH3A with Cbl-b shows an exothermic thermogram, and as shown before the interaction with atypical proline-rich targets is enthalpically driven. Thermodynamic parameters were obtained by fitting the experimental data using a model of n identical and independent sites (Table IV. 5). The affinity is higher for CD2AP-SH3A mainly due to larger enthalpy although the interaction in the case of CIN85-SH3A has less entropic penalty indicative of a more flexible interaction (Tables IV.4 and IV.5).

Importantly, the value of n obtained in the titration of CIN85-SH3A with Cbl-b was close to 0.5. In order to check the stoichiometry and the correctness of the binding model, we also carried out the inverse experiment, where Cbl-b was titrated with CIN85-SH3A, obtaining a value of n close to 2.0, thus in agreement with the direct experiment. These results point to a 2:1 stoichiometry where two molecules of CIN85-SH3A interact simultaneously with one molecule of Cbl-b, in accordance with previous ITC data, *in vivo* experiments and the crystal structure of the trimeric complex (Jozic D. et al., 2005) (Fig. IV.15).

PEPTIDES	n	K_d (μM)	ΔH_b (kcal/mol)	ΔG_b (kcal/mol)	$-T \cdot \Delta S_b$ (kcal/mol)
CD2	1.1 ± 0.1	9.9 ± 0.4	-9.9 ± 0.4	-6.7 ± 0.1	3.2 ± 0.4
Cbl-b	0.4 ± 0.1	1.8 ± 0.2	-7.6 ± 0.2	-7.7 ± 0.1	-0.1 ± 0.1
Cbl-bR904G	1.0 ± 0.1	2.0 ± 0.2	-10.0 ± 0.4	-7.6 ± 0.1	2.4 ± 0.4
Cbl-bR911A	1.0 ± 0.1	46.9 ± 1.8	-6.1 ± 0.5	-5.8 ± 0.2	0.3 ± 0.4
Cbl-bK907L	0.7 ± 0.1	3.3 ± 0.2	-11.7 ± 0.4	-7.3 ± 0.1	4.4 ± 0.4

Table IV.5. Thermodynamic parameters calculated from the fitting of the isotherms obtained during the titration of the SH3A domain of CIN85 (40 - 60 μM) with the different peptides (0.6 - 1.4 mM), performed in 50 mM cacodylate buffer pH 6.0 at 20 °C, on a VP-ITC of MicroCal. Values calculated using the model of n identical and independent sites. The reported values are averaged over three experiments and the errors are calculated as the standard deviation from the mean values.

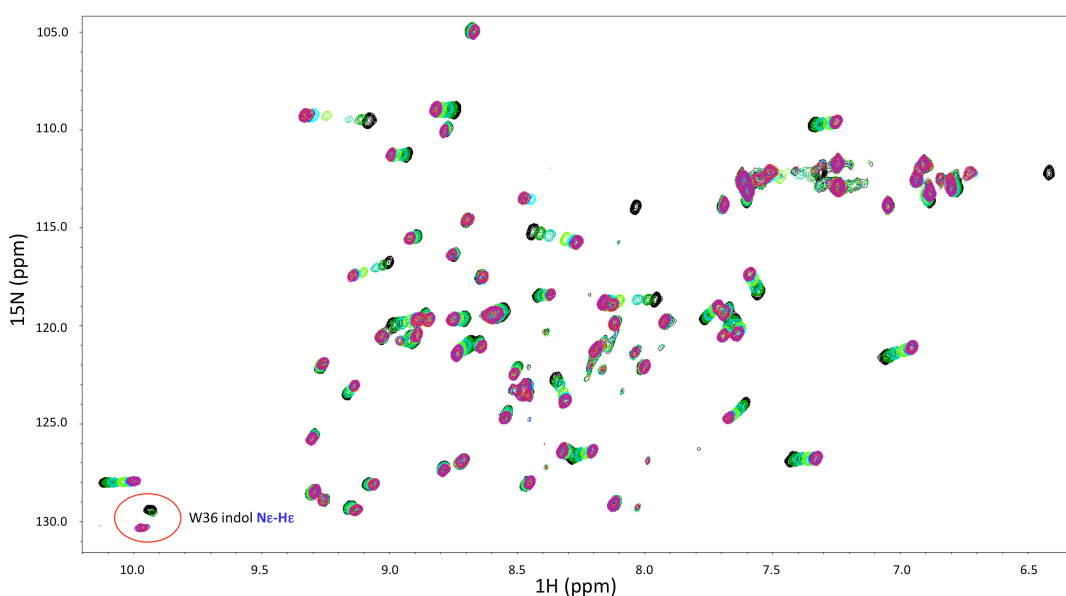


Figure IV.13. Overlay of the 1H - ^{15}N -HSQC spectra recorded during the NMR titration of the SH3A domain of CIN85 with the peptide Cbl-b. Titration was done in 50 mM NaPi buffer at pH 6.0 and 25 °C on an 800 MHz spectrophotometer. Spectra of free SH3A in black and Cbl-b-bound SH3A in violet (final titration point corresponds to a 2-fold excess of Cbl-b). The region of the amide indole proton $N\epsilon 1-H\epsilon 1$ of Trp36 is surrounded by a red circle.

To further investigate and confirm our ITC results, we followed the binding of CIN85-SH3A with Cbl-b by NMR. Titrations of ^{15}N -labeled CIN85-SH3A with increasing amounts of Cbl-b peptide at 25 °C caused a selective shift of amide proton and nitrogen resonances of several SH3A residues (Fig. IV.13), indicating a specific interaction between the protein and the peptide. Most of the residues in the binding site are in a fast-to-medium exchange regime. Like for the interaction with CD2, at the end of the titration, the residues mostly affected of CIN85-SH3A by binding Cbl-b are in the the RT loop (residues 8, 9, 11, 12, 13, 16, 18), the n-Src loop (residues 32, 33, 35, 36), the β -III (residues 37, 39), the β -IV (residue 48) and at the 3_{10} helix (residues 50-53) (blue bars in Fig. IV.14A,D).

Figure IV.14 summarizes and compares the chemical shift perturbations of all the NMR titrations that have been done between the SH3A domains of CD2AP and CIN85 and the atypical proline-rich peptides CD2 and Cbl-b. By comparing the binding interface of both adaptor proteins, we observe that despite the differences at the beginning of the RT loop, in both SH3A domains the acidic pocket and the two hydrophobic pockets are affected, which correspond to the canonical recognition surface of proline-rich peptides by SH3 domains (Larson S.M. and Davidson A.R., 2000; Cesareni G. et al., 2002; Ferreon J.C. and Hilser V.J., 2004).

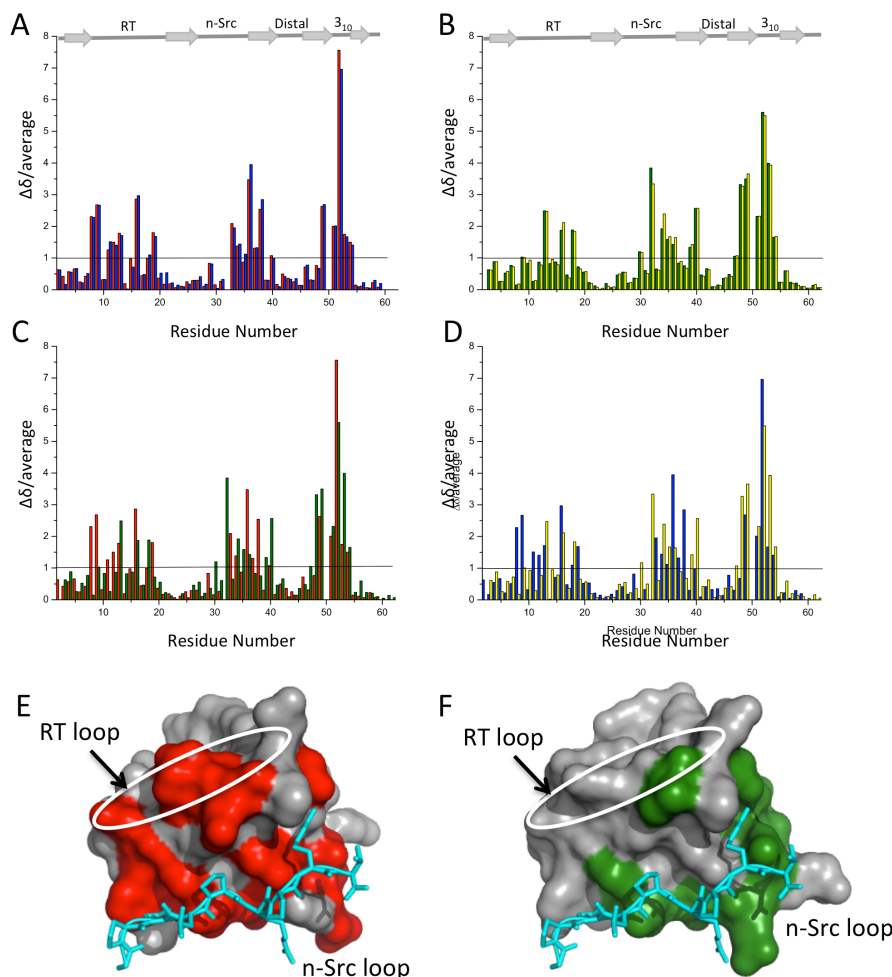


Figure IV.14. Summary of the NMR titrations of ^{15}N -labeled CIN85-SH3A and CD2AP-SH3A with CD2 and Cbl-b. **A-D.** Representation of the normalized chemical shift perturbations ($\Delta\delta/\text{average}$) of the backbone ^1H - ^{15}N pairs of residues of CIN85-SH3A and CD2AP-SH3A. For comparison, residue numbering is as in the SH3A domain of CD2AP. The mean perturbation is indicated by a black horizontal line. Secondary structure elements of the SH3A domains are represented above, the grey arrows correspond to the β -sheets and the positions of the loops and helix 3_{10} are indicated. **A.** Comparison of the titration of CIN85-SH3A with CD2 (red) and Cbl-b (blue). **B.** Comparison of the titration of CD2AP-SH3A with CD2 (green) and Cbl-b (yellow). **C.** Comparison of the titration of CIN85-SH3A with CD2 (red) and the titration of CD2AP-SH3A with CD2 (green). **D.** Comparison of the titration of CIN85-SH3A with Cbl-b (blue) and the titration of CD2AP-SH3A with Cbl-b (yellow). **E-F.** Surface representation of the SH3A domains of CIN85 (E) and CD2AP (F) engaging the CD2 peptide (cyan sticks) in a type II orientation. Residues more affected ($\Delta\delta/\text{average}$ more than 1, see panel C) in CIN85-SH3A in red and in CD2AP-SH3A in green. The location of the n-Src and RT loop are indicated, the latter by a white oval.

Interestingly, during the NMR titration of CIN85-SH3A and Cbl-b (Fig. IV.13), some residues (e.g. Trp36 and Phe52) show again the unusual curvature. Such curvature is most pronounced for the ^1H - ^{15}N shifts of the Trp36 indole N ϵ -H ϵ moiety (Fig. IV.16E, red contours), as observed during the titration of the SH3A domain of CD2AP with Cbl-b. This tryptophan is very well conserved in all SH3 domains as it plays a crucial role in binding polyproline stretches (Fernandez-Ballester G. et al., 2003). The indole moiety of this Trp is thus an excellent probe for monitoring changes that take place in the binding interface (Fig. IV.15).

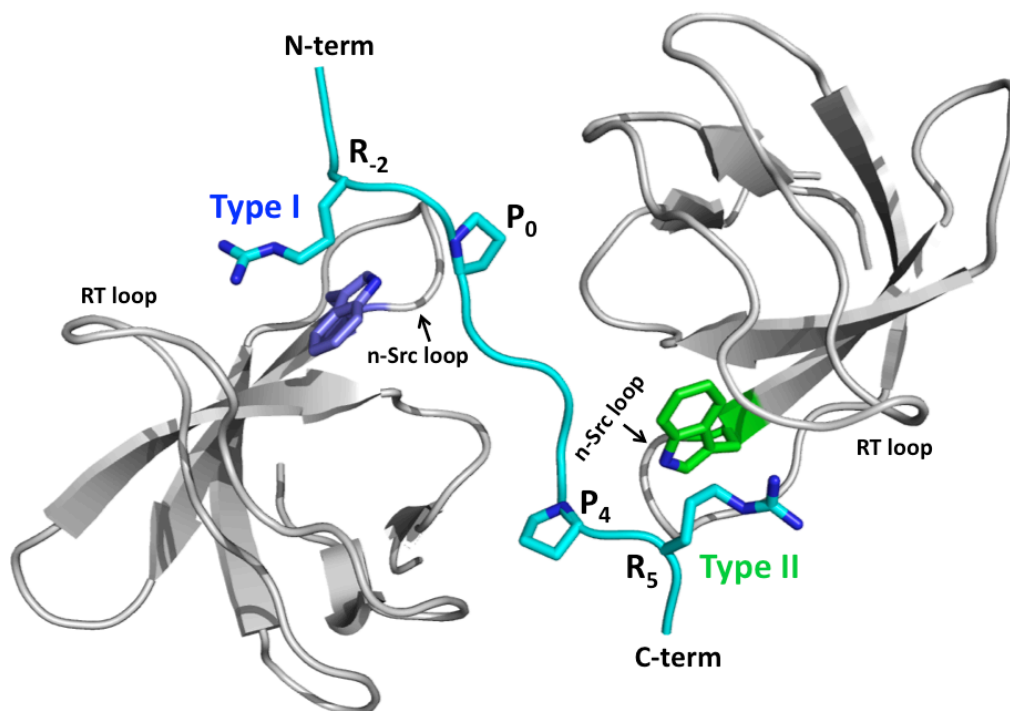


Figure IV.15. Structural presentation of the hetero-trimer formed by one Cbl-b sandwiched between two CIN85-SH3A molecules (PDB entry 2BZ8). The two SH3 molecules are colored grey, Cbl-b in cyan. The key residues of the atypical polyproline-arginine binding motif in Cbl-b are shown in sticks representation ($R_{-2}x_1P_0x_2x_3P_4R_5$). The conserved tryptophan (Trp36 in case of CIN85-SH3A) is also presented in sticks and colored in blue and green to indicate its involvement in type I and type II interactions, respectively. These same colors are used throughout the text to distinguish the two binding orientations.

The curvature of the Trp36 indole proton of CIN85-SH3A was also observed previously (Ababou A. et al., 2009), and was explained by the simultaneous formation of two dimers (type I and type II) instead of a trimeric complex, in disagreement with the previously determined crystal structure of the trimeric complex (Jozic D. et al., 2005). However, Ababou and coworkers (Ababou A. et al., 2009) did not provide experimental evidence to explain the relation between the directionality of the curvature and the type of interaction taking place in solution. They used a theoretical model without any experimental data to differentiate between dimers and trimers. In section 4.3.3 we have shown the simulation of the curvature of the Trp37 of CD2AP-SH3A using a mathematical model based on experimental K_{ds} (section 4.3.3), pointing to the existence of type I and type II dimers in solution. Moreover, the curvature observed in the Trp36 during the titration of CIN85-SH3A is different, being worth to notice that the Trp36

indole signal broadens out during the intermediate steps of the titration indicative of exchange between different conformational, possibly oligomeric states. Such exchange broadening was not observed for the Trp37 indole proton of CD2AP (Fig. IV.9E).

The crystal structure of the trimeric complex between the SH3A domain of CIN85 and Cbl-b (Jozic D. et al., 2005) indicates that arginines 904 and 911 in Cbl-b are determinant for the type I and type II interactions, respectively. So, as for the understanding of the curvature of the N ϵ 1-H ϵ 1 of Trp37 of CD2AP-SH3A during the titration with Cbl-b (section 4.3.3), we also titrated CIN85-SH3A with the two mutants of Cbl-b previously designed: Cbl-bR904G to eliminate type I interaction (considering that there is a glycine at the equivalent position in CD2) and Cbl-bR911A to eliminate the type II interaction.

The titrations of CIN85-SH3A with these mutant peptides were monitored by both ITC (Fig. IV.16B and 16D, respectively) and NMR. Figure IV.16E shows that no curvature of the Trp36 indole N ϵ -H ϵ peak is seen during the NMR titration of the SH3A domain of CIN85 with the two mutants; the isolated type II interaction (Cbl-bR904G, Fig. IV.16 green contours) shows a linear downfield shift change, while the isolated type I (Cbl-bR911A, Fig. IV.16 blue contours) shows a linear upfield shift change of the same Trp36 indole. As was already explained for CD2AP, the linearity of the shift change is related to a single binding event (Zuiderweg E. R., 2002), and our previous results already indicated that the directionality of the shift change is also related to the formation (downfield shift change for type II) or not (upfield change for type I) of a H-bond involving the Trp indole proton. Indeed, inspection of the trimeric CIN85-SH3A:Cbl-b structure (PDB entry 2BZ8) shows that the Trp36 N ϵ 1-H ϵ 1 is involved in a H-bond with the carboxyl group of Arg909 in Cbl-b in the type II orientation, while the distance between Trp36 N ϵ 1-H ϵ 1 and the carboxyl group of Arg905 is too long to form a H-bond in the type I orientation.

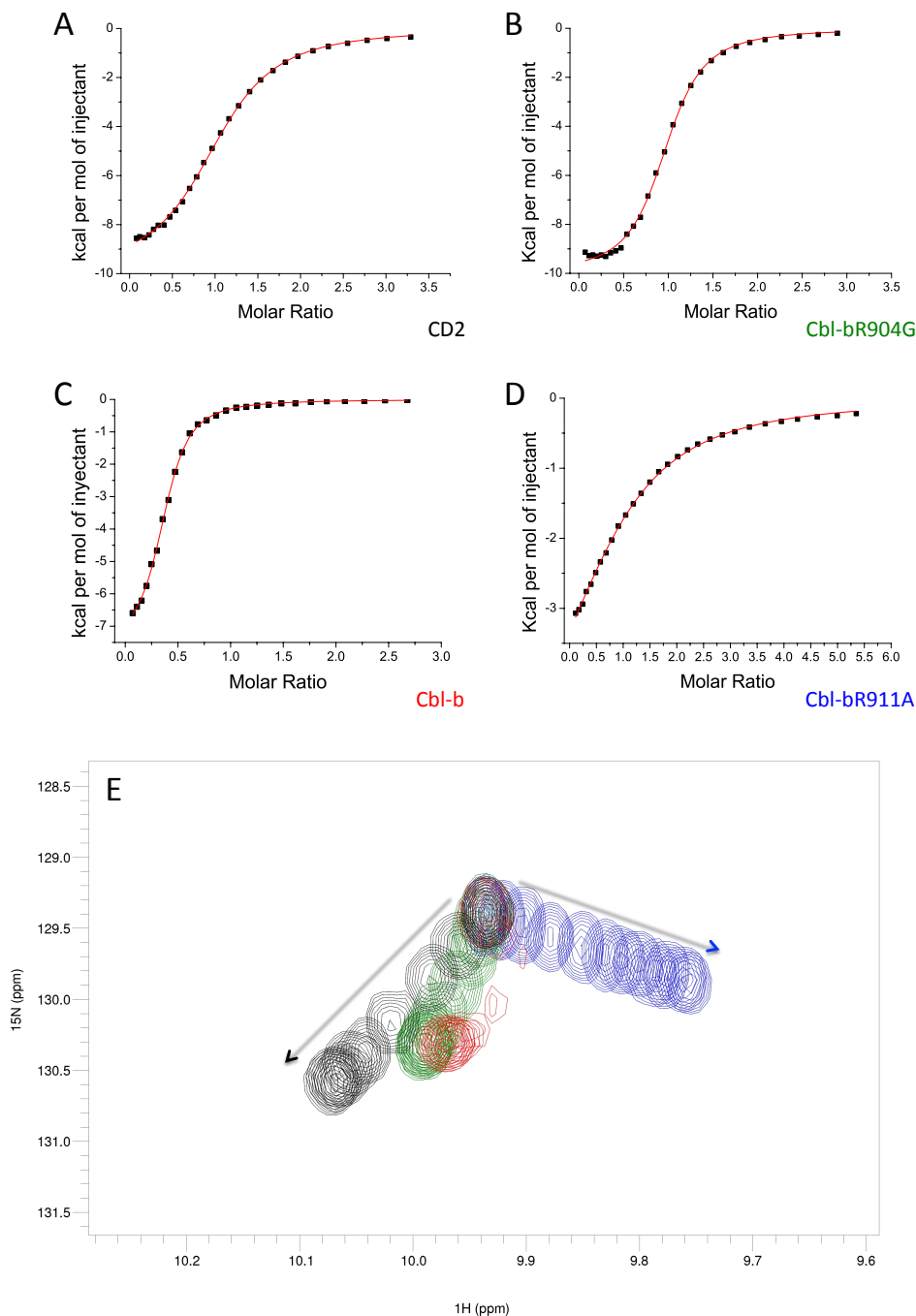


Figure IV.16. ITC and NMR titration of CIN85-SH3A with CD2 and Cbl-b variants. **A-D.** ITC titration curves corresponding to titration with CD2 (**A**), Cbl-bR904G (**B**), Cbl-b (**C**) and Cbl-bR911A (**D**). Solid lines in the ITC show the best fit using the model of n identical and independent sites. For clarity, each ITC titration curve is labeled with the peptide used in the titration and boxes of panels **A-D** are colored according to the colors used in the NMR titrations. **E.** Selected region of NMR spectra during titrations showing the chemical shift changes of the Trp36 indole N ϵ -H ϵ moiety. The spectra of intermediate titration points of the Cbl-b titration (red) are plotted at a lower contour level to enable to observation of the curvature. All other spectra are plotted at the same contour level. Colors used in the NMR spectra are black for CD2, red for Cbl-b, green for Cbl-bR904G and blue for Cbl-bR911A. All titrations start at the same chemical shift of the Trp36 N ϵ 1-H ϵ 1 in the free form. For clarity, the black and blue arrows indicate the direction of the chemical shift changes during the titration with CD2 peptide and mutant Cbl-bR911A, respectively.

The expected one-site binding was also confirmed by ITC (Figs. IV.16B,D), where for both mutants, a one-to-one stoichiometry was found (Table IV.5). The K_d of 2.0 μM for the type II interaction compared to 46.9 μM for the type I interaction indicates a preferential type II binding of the Cbl-b peptide by CIN85-SH3A, as in the case of the interaction of the SH3A domain of CD2AP with Cbl-b.

During the interaction of CIN85-SH3A with Cbl-b mutants, the enthalpy value of the interaction with Cbl-bR904G (isolated type II interaction) was twice the enthalpic term of the interaction with Cbl-bR911A (Table IV.5), like it was previously observed for their interaction with CD2AP-SH3A (section 4.3.3, Table IV.4). In both cases the interaction is enthalpically driven, with a small entropic contribution to the binding free energy, specially in the case of Cbl-bR911A. Despite that the SH3A domains present two hydrophobic pockets on their surfaces and one acidic pocket, the absolute value of the entropic term of the interaction with Cbl-bR911A is much smaller than the enthalpic term. Due to the hydrophobic nature of the binding pocket a higher entropic term would be expected, so the thermodynamic signature of the interaction indicates that other forces, mainly electrostatic (involving the interaction of Arg904 with the acidic pocket) are responsible for the isolated type I interaction.

It is worth to notice that the thermodynamic parameters of the interaction of CD2AP-SH3A (section 4.3.3) and CIN85-SH3A with the mutants Cbl-bR904G and Cbl-bR911A are very similar, which means that the isolated type I and type II interactions in both SH3A domains possess a similar nature, validating the role of Arg904 in the type I interaction with CD2AP-SH3A as observed in the previously reported X-ray structure of CIN85-SH3A in complex with Cbl-b (PDB code 2BZ8).

All the above described experiments indicate that for the interaction of the SH3A domains of CD2AP and CIN85 with Cbl-b, the Arg911 is crucial for the type II interaction, while the Arg904 is crucial for the type I; lack of one of the two arginines strongly weakens the binding in which they are implicated. However, Cbl-b contains another basic residue, Lys907, close to the Arg904, and it has been shown that this Lys is also involved in the type I interaction with the acidic pocket of the SH3A domain of CMS (Moncalian G. et al., 2006). To validate its importance in the type I binding and its contribution to the curvature of the indole proton of the conserved tryptophan, we mutated Lys907 to a leucine (as in CD2). The titrations of CIN85-SH3A with the Cbl-bK907L mutant followed by ITC and NMR showed a similar pattern as Cbl-b WT. In contrast with the mutant Cbl-bR904G, the ITC results with the mutant Cbl-bK907L does not show a clear 1:1 stoichiometry, but a $n = 0.7$, closer to the stoichiometry determined for the interaction of CIN85-SH3A with Cbl-b WT (Table IV.5). And the curvature of the $\text{N}\epsilon\text{-H}\epsilon$ indole moiety of Trp36 also appears during the NMR titration with this mutant (Fig. IV.17A), supporting a major role for Arg904 in the type I interaction. The chemical shift changes of the Trp36 side chain along the titration are slightly more downfield displaced than in the case of Cbl-b, indicating likely higher contribution of type II interactions. These small differences are probably due to the loss of the H-bond between Lys907 of Cbl-b and Tyr10 of CIN85-SH3A in the trimeric complex, thus destabilizing the type I interaction. As observed for CIN85, the NMR titration of CD2AP-SH3A with Cbl-bK907L shows again a slightly downfield displaced curvature in the $^1\text{H}\text{-}^{15}\text{N}\text{-HSQC}$ spectra (Fig. IV.17B). So we can conclude that Lys907 in Cbl-b contributes to the stabilization of the type I interaction with the SH3A domains of CD2AP and CIN85, however the absence of such a residue does not abolish type I

binding and does not remove the curvature, while the absence of the Arg904 removes the curvature and shows a lineal shift change by NMR together with a 1:1 stoichiometry by ITC.

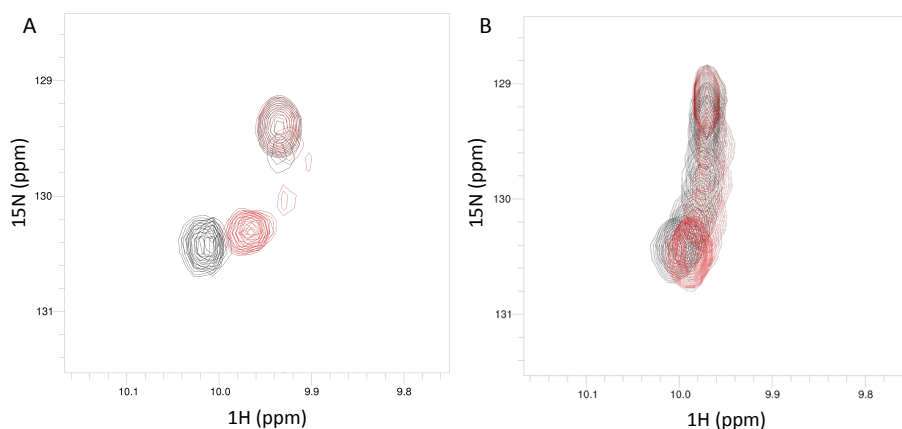


Figure IV.17. Superposition of ^1H - ^{15}N -HSQC spectra region of the indole $\text{N}\epsilon$ - $\text{H}\epsilon$ moiety of Trp36 in CIN85-SH3A (A) and the Trp37 in CD2AP-SH3A (B) during the titration with Cbl-b (red) and with the mutant Cbl-bK907L (black). In both NMR titrations the curvature is observed.

To further characterize the differences between the curvatures that we observed in the NMR titrations of the SH3A domains of CD2AP and CIN85 with Cbl-b, we also tried to simulate the curved trajectory followed by the CIN85-SH3A Trp36 indole $\text{N}\epsilon$ - $\text{H}\epsilon$, as we previously did with the Trp37 indole of CD2AP-SH3A. With that purpose, we used the K_{dS} resulting from the fitting of chemical shift changes of the Trp36 indole $\text{N}\epsilon$ - $\text{H}\epsilon$ from the NMR titration of the isolated type I (with Cbl-bR911A) and isolated type II (with Cbl-bR904G) interactions. The K_{dS} were used to calculate the percentage of type I, type II dimers and free SH3 domain at every point of the titration, assuming that only dimers could be formed in solution and that type II interaction, since it is tighter than type I, occurs first (section 4.2.4). These percentages were used to calculate the expected chemical shift changes of the Trp indole $\text{H}\epsilon 1$ of the SH3A domain of CIN85, as we previously did with the SH3A domain of CD2AP during the titration with Cbl-b and mutants. For both CIN85 and CD2AP SH3A domains, the experimental data of the isolated type II and type I interactions, can be perfectly simulated using this mathematical approach, which proves the adequacy of the simulation when there are only dimers in solution. Indeed the titration of the SH3A domain of CD2AP with Cbl-b (section 4.3.3) can be simulated perfectly as well because the main species in solution are independent dimers type I and type II (Fig. IV.10). However, as can be seen in the figure IV.18, in the case of the titration of CIN85-SH3A with Cbl-b WT we cannot simulate the experimental data using the mathematic approach that only considers the formation of dimers (type I and type II) in solution. This indicates that besides the formation of type I and type II hetero-dimers other multimeric states (e.g. trimer) exist in solution. These results are thus in agreement with the stoichiometries determined by ITC, and the strong broadening of the NMR signals in the HSQC spectra (Fig. IV.16C and E, respectively). Based on this simulation, we can conclude that at low Cbl-b concentration, the trimeric state is highly populated, and upon increasing Cbl-b concentration, the main species in solution are dimers type II, in agreement with the Trp36 indole chemical shift tendency to approach purely type II interaction at the end of the titration (Fig. IV.18). Thus under saturating ligand conditions, the main species are

type II hetero-dimers. This is in full agreement with the similar intermolecular NOE patterns involving the Trp36 indole H ϵ 1 proton and the higher affinity for the type II interaction (Table IV.5). However, since the Trp36 H ϵ 1 ^1H chemical shift does not completely reach the chemical shift of the isolated type II interaction, we expect that there is still a small fraction of trimeric complex present in solution at the end of the titration.

The fact that, at different concentration of peptide, the mechanisms of clustering are different for CIN85-SH3A (trimers at low concentration of peptide and dimers at high concentration) than for CD2AP-SH3A (where the main species in solution are dimers) might have important biological implications.

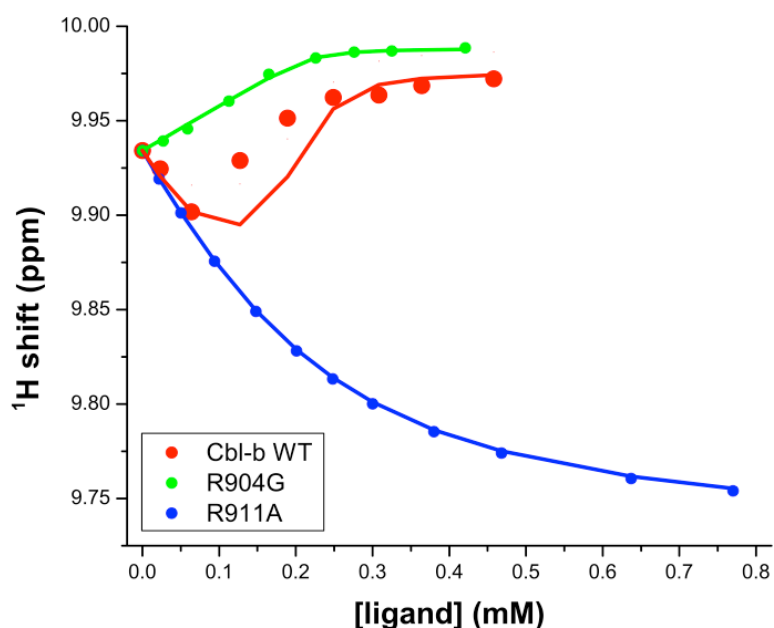


Figure IV.18. Representation of the ^1H chemical shift of the H ϵ 1 from Trp36 of CIN85-SH3A during the titration with Cbl-b (red), Cbl-bR911A (blue) and Cbl-bR904G (green). The chemical shift of the H ϵ 1 (dots) at each ligand concentration were simulated (solid lines) as described in the text.

4.3.6 Structural characterization of the interaction of CD2AP-SH3A and CIN85-SH3A with Cbl-b and CD2 using Small Angle X-ray Scattering (SAXS)

The results obtained from the combination of ITC and NMR data strongly support a scenario where CD2AP-SH3A forms only hetero-dimeric complexes with CD2 and Cbl-b atypical proline-rich peptides, while CIN85-SH3A can form both hetero-dimers (with CD2) as well as hetero-trimers (with Cbl-b). However, due to the controversy around the recognition mechanism of atypical proline-rich targets by the CIN85/CD2AP family of adaptor proteins, we aimed to confirm the results obtained by ITC and NMR, using another complementary technique in solution. The selected technique was SAXS, which can be used to study the composition of complex mixtures in solution (Chapter III, section 3.4).

We recorded scattering curves (Fig. IV.19A) of different SH3A:peptide complexes at different ratios to correlate the features observed in the NMR titrations with direct structural changes in the complexes (section 4.2.5). We assumed the presence of three major species present in detectable amounts at equilibrium throughout the titration. These species consist of monomers of SH3A, hetero-dimers (formed in type I or type II orientations) and hetero-trimeric complexes. We then used the program OLIGOMER (Konarev P.V. et al., 2003) that fits any experimental SAXS curve by a linear combination of a number of input scattering curves. This allows monitoring the evolution of the volume fractions of individual species present at different stages of the titration (see fractions of dimer and trimer at every ratio in table IV.6).

Species	Dimer/Trimer composition ($\pm 5\%$)	χ^2	R_g (Å)	D_{max} (Å)
CD2AP-SH3A:CD2 1:1	94.0 / 6.0	1.47	14.0	42.0
CD2AP-SH3A:CD2 1:1.6	95.5 / 4.5	1.85	13.8	43.5
CD2AP-SH3A:Cbl-b 1:1	99.6 / 0.4	1.08	12.9	40.1
CD2AP-SH3A:Cbl-b 1:1.6	99.0 / 1.0	2.1	14.3	43.0
CIN85-SH3A:CD2 1:0.5	91.1 / 8.7	1.01	14.3	42.0
CIN85-SH3A:CD2 1:1.6	87.5 / 12.5	0.98	14.1	45.0
CIN85-SH3A:Cbl-b 1:0.5	31.5 / 68.3	1.81	17.5	56.1
CIN85-SH3A:Cbl-b 1:1.6	70.5 / 29.5	1.01	14.6	48.0
CIN85-SH3A:Cbl-bK907L 1:0.5	34.0 / 66.0	1.01	17.2	55.5

Table IV.6. Computed solution scattering parameters for the complexes of CD2AP-SH3A and CIN85-SH3A domains with CD2 and Cbl-b peptides at indicated ratios. χ^2 fitting statistics were obtained from the program OLIGOMER (Konarev P.V. et al., 2003). R_g and D_{max} values are from figures IV.21 and IV.20, respectively.

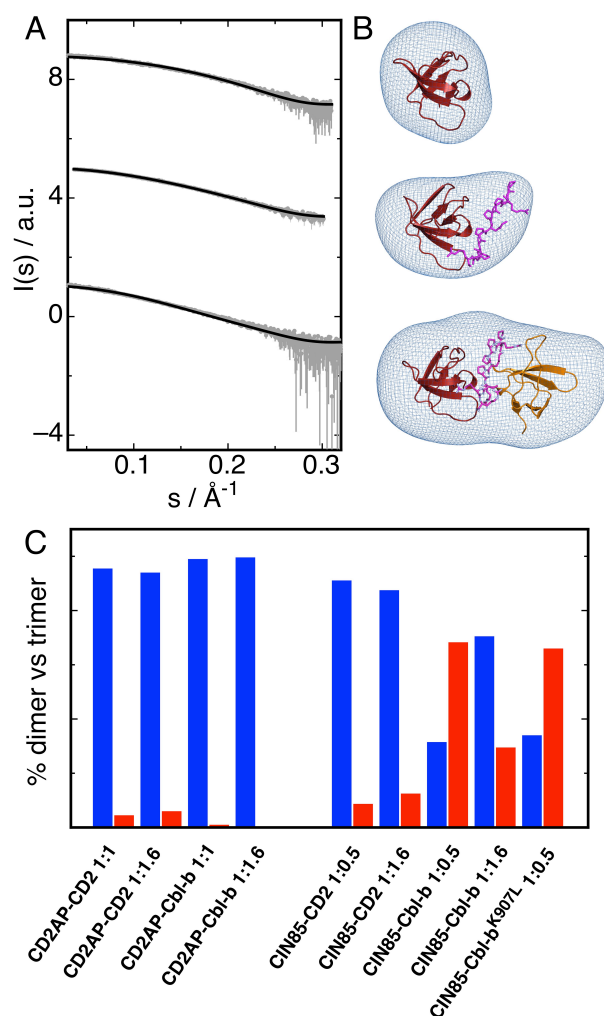


Figure IV.19. Solution structure of CIN85-SH3A, CIN85-SH3A:CD2 and CIN85-SH3A:Cbl-b:CIN85-SH3A. **A.** Experimental SAXS scattering curves (from top to bottom) for free CIN85-SH3A, the CIN85-SH3A:CD2 dimer and CIN85-SH3A:Cbl-b trimer/dimer mixture (~68%/32% ratio). The fitting of the different models to the experimental data is shown in every case as a solid line (see **Table IV.6** for further details). **B.** *Ab initio* reconstructed shapes for CIN85-SH3A, the CIN85-SH3A:CD2 dimer and CIN85-SH3A:Cbl-b trimer (from top to bottom) superposed onto the pseudo-atomic structure of each species. SH3 domains are colored in red and orange and the peptide is colored in magenta. **C.** The dimer/trimer (in blue/red respectively) composition of the complexes of CD2AP-SH3A and CIN85-SH3A with CD2 and Cbl-b (see **Table IV.6** for further details).

We selected the equivalent of 3 to 4 different stages of the NMR titration (the starting point with zero concentration of peptide, the middle point 1:0.5 ratio, the equimolar point, 1:1 and above saturation, 1:1.6) to assess the relative composition of the populations in solution. Figure VI.19 and **Table IV.6** show and summarize the analysis of the experimental SAXS curves. In the case of the interaction of CD2AP-SH3A with CD2 and Cbl-b, the SAXS data confirms the overwhelming presence of dimeric complexes as the main species in solution, with only traces of trimers, which percentages are within the error of the analysis.

On the other hand, CIN85-SH3A shows a different behavior. It interacts with CD2 to form mainly dimers in solution, however at half of the equimolar point and above,

trimers account for about 10% of the population (Fig. IV.19C and Table IV.6). With Cbl-b the situation is even more peculiar. The experimental scattering curves can only be explained by considering the presence of a trimeric complex in significantly large amounts in equilibrium with dimeric complexes. Figure IV.19C and figure IV.20 shows that the trimeric species (68%) dominate in solution at a ratio of 1:0.5 (SH3A:Cbl-b) and decreases to approximately half of the dimeric population above the equimolar point (29.5%). Similar behavior was found for the interaction of CIN85-SH3A with the Cbl-bK907L peptide, where at a molar ratio of 1:0.5, the percentage of hetero-trimeric complex in solution was basically the same than with the Cbl-b WT. The decrease in the fraction of hetero-trimers between CIN85-SH3A and Cbl-b upon increasing the amount of Cbl-b is thus in full agreement with the curvature observed in the NMR titration experiments (section 4.3.5).

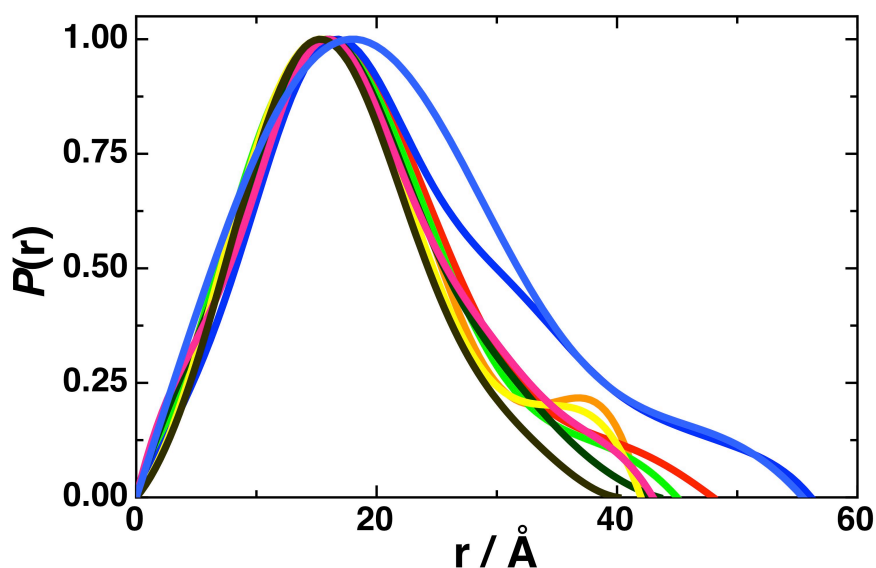


Figure IV.20. Pair-distance distribution functions, $P(r)$, obtained from the indirect Fourier transform of the experimental SAXS curves as implemented in the GNOM (Konarev P.V. et al., 2003 and 2006) package for CD2AP-SH3A:CD2 1:1 (in yellow); CD2AP-SH3A:CD2 1:1.6 (in dark green); CD2AP-SH3A:Cbl-b 1:1 (in brown); CD2AP-SH3A:Cbl-b 1:1.6 (in pink); CIN85-SH3A:CD2 1:0.5 (in orange); CIN85-SH3A:CD2 1:1.6 (in light green); CIN85-SH3A:Cbl-b 1:0.5 (in blue); CIN85-SH3A:Cbl-b 1:1.6 (in red); CIN85-SH3A:Cbl-bK907L 1:0.5 (in light blue). Note that the increase in D_{max} observed for the complexes of the CIN85-SH3A with Cbl-b (especially at 1:0.5 ratio) and the K907L mutant of Cbl-b, clearly reflects the presence of a trimeric complex populating in significantly large amounts the equilibrium with dimeric complexes.

The *ab initio* shape reconstruction of the different species provides additional support to our description of the populations in solution as obtained with OLIGOMER (Konarev P.V. et al., 2003). For modeling we used the program DAMMIF (Svergun D.I., 1999) without constraints imposed on the simulated annealing protocol. All the models predicted from the scattering curves, were very reproducible in independent runs using the program DAMMIF, with average normalized spatial discrepancy (NSD) values below 1.0. The models were highly consistent with the pseudo-atomic models of free SH3A, dimers and trimers used in our calculations and validated the presence of trimers in solution (Fig. IV.19 B,C).

Figure IV.20, shows the representation of the pair-distance distribution function corresponding to the different complexes. The $P(r)$ contains information about the distances between electrons in the scattering particles present in the sample. For globular samples, $P(r)$ is zero at $r = 0$ and at $r \geq D_{\max}$ (D_{\max} correspond to the maximum linear dimension in the scattering particle), and is an indicator of good quality data (Putnam C.D. et al., 2007). The pair distribution function shows that the samples with the highest percentage of hetero-trimer (ratio 1:0.5 for Cbl-b and the mutant Cbl-bK907L) have bigger D_{\max} .

The figure IV.21 summarizes all the *Guinier* plots from the scattering curves of the different samples, and as expected (see Chapter III, section 3.4), all of them are linear. As a rule, linearity of the *Guinier* plot is a sensitive indicator of the quality of the experimental data and deviations from linearity usually point to strong interference effects like polydispersity of the samples or improper background subtraction.

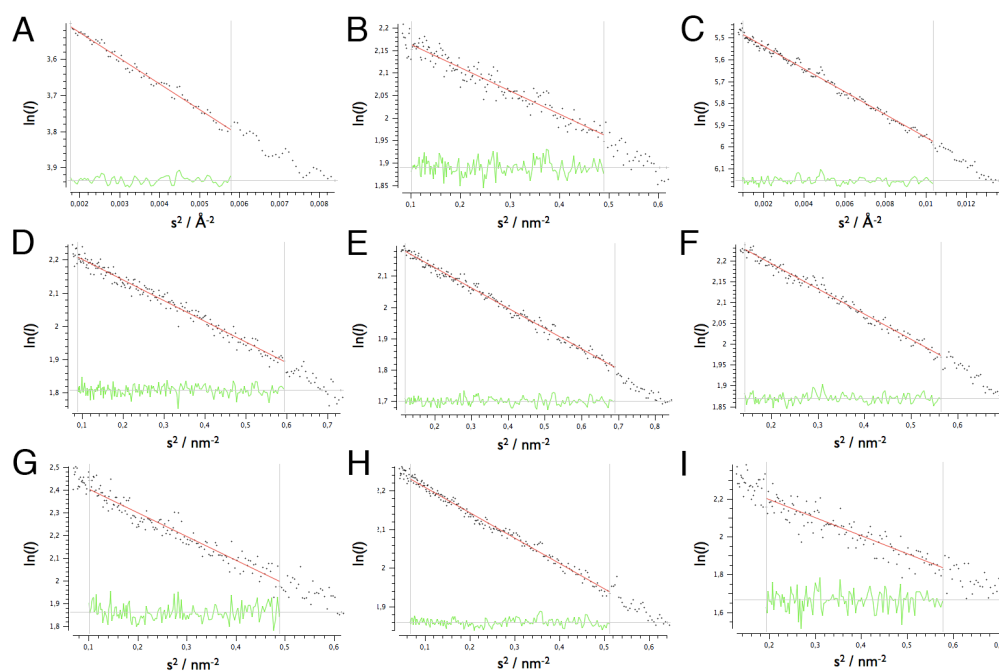


Figure IV.21. Determination of the R_g from the experimental scattering curves using the *Guinier* approximation. **A**, CD2AP-SH3A:CD2 1:1; **B**, CD2AP-SH3A:CD2 1:1.6; **C**, CD2AP-SH3A:Cbl-b 1:1; **D**, CD2AP-SH3A:Cbl-b 1:1.6; **E**, CIN85-SH3A:CD2 1:0.5; **F**, CIN85-SH3A:CD2 1:1.6; **G**, CIN85-SH3A:Cbl-b 1:0.5; **H**, CIN85-SH3A:Cbl-b 1:1.6; **I** CIN85-SH3A:Cbl-bK907L 1:0.5. The grey lines demark the range used for fitting of the *Guinier* region. Residuals are shown in green underneath the experimental points.

The *Guinier* approximation (Fig. IV.21) has been used to determine the radius of gyration (R_g) and the extrapolated intensity at zero scattering angle $I(0)$. R_g describes the mass distribution of a protein around its center of gravity. This parameter is really useful especially when the protein is undergoing conformational changes during ligand binding. For the mainly dimeric complexes formed by the SH3A domain of CD2AP with Cbl-b and CD2 and by CIN85-SH3A with CD2, all the R_g present a similar value around 13-14 Å (Table IV.6), however, for the complex of the SH3A domain of CIN85 with Cbl-b and the mutant Cbl-bK907L, that are able to form hetero-trimers, the values of R_g increase till 17.5 Å, indicative of the presence of larger particles.

4.4 Concluding remarks

RTKs control many important biological processes and therefore need to be tightly regulated. Downregulation of the RTKs is essential to normal cell function and failures are frequently associated to severe diseases like cancer (Bache K.G. et al., 2004). Key to this downregulation is the formation of large protein clusters where CIN85 and CD2AP play an important role by recruiting several ligand molecules. Both adaptors contain three SH3 domains that recognize atypical polyproline-arginine binding motifs present in Cbl-b/c-Cbl and CD2. Cbl-b/c-Cbl was shown to simultaneously contact two N-terminal CIN85/CMS SH3 domains providing additional molecular functions to the ubiquitin ligase (Jozic D. et al., 2005). The stoichiometry of the complexes formed by Cbl and CIN85-SH3A has been controversial (Ababou A. et al., 2008; Jozic D. et al., 2008). Using several complementary biophysical techniques, we aimed to fully characterize the interactions established by the family of adaptors including both CIN85 and CD2AP N-terminal SH3 domains, with two different cellular ligands, Cbl-b and CD2.

4.4.1 Atypical polyproline recognition: the role of arginine residues for type I and type II binding

The CIN85/CD2AP family of adaptor proteins does not bind to canonical polyproline sequences but were shown to recognize an atypical motif $\mathbf{P}_{xxx}\mathbf{PR}$ (Kowanetz K. et al., 2003; Kurakin A.V. et al., 2003; Zarrinpar A. et al., 2003; Jozic D. et al., 2005). The C-terminal arginine is driving the type II orientation in the SH3-peptide complex. Based on our results, we need to extend the polyproline sequence to include the arginine responsible for the type I interaction as present in Cbl-b: $\mathbf{R}_2\mathbf{x}_1\mathbf{P}_0\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3\mathbf{P}_4\mathbf{R}_5$ (Fig. IV.15). We use this numbering to discuss the contributions of the different residues in the binding with the SH3A domains of CIN85 and CD2AP.

ITC, NMR and SAXS experiments (black in Fig. IV.16A,E and Fig. IV.19C) show for the first time, that the SH3A domain of CIN85 is able to interact with the atypical proline-sequence of the peptide CD2, predominantly forming a dimer type II as it lacks an arginine at position \mathbf{R}_2 (section 4.3.2 and 4.3.4). In the case of peptides that present the extended atypical proline-rich sequence, including the \mathbf{R}_2 , like Cbl-b, the interaction with the SH3A domain involves both type II and type I modes of binding (section 4.3.3 and 4.3.5).

Using ITC, we have determined the stoichiometry 2:1 of CIN85-SH3A with Cbl-b peptide, pointing to the formation of a trimeric structure in solution, consisting of a Cbl-b peptide sandwiched by two CIN85-SH3A domains, as shown in the crystal structure (Fig. IV.15) (Jozic D. et al., 2005). Previous studies reported that the formation of the hetero-trimer was due to a crystallization artifact and disputed the formation of such complex in solution (Ababou A. et al., 2008; Jozic D. et al., 2008; Ababou A. et al., 2009). In our work we solved the supposed controversy as the existence of a trimeric complex in solution is in agreement with our results from ITC, NMR and the *ab initio* reconstructed shapes obtained by SAXS. In support of this, hetero-trimerization was also shown *in vivo* (Jozic D. et al., 2005). This formation is not an exception since SH3 multimerization has been observed at least in CIN85, β -PIX (Jozic D. et al., 2005), CTTN (Hashimoto S. et al., 2006) and Fyn-SH3 (Martin-Garcia J.M. et al., 2012).

Mutants Cbl-bR911A and Cbl-bR904G, designed to eliminate type II and type I interactions respectively, highlight the importance of the N-terminal (**R**₂) and C-terminal (**R**₅) arginines for the formation of the trimeric complex (Fig. IV.15). As expected, the ITC titrations indicated that both Cbl-bR904G and Cbl-bR911A preclude the formation of trimers, forming respectively type II and type I dimers with CIN85-SH3A.

In accordance, when analyzing the NMR signals of CIN85-SH3A Trp36 indole ¹⁵N-¹H cross peaks during the Cbl-b titrations, linear shift changes were observed during the titration with both Cbl-b mutants instead of the curvature observed with Cbl-b (Fig. IV.16). The directionality of the chemical shift change of the indole proton of the tryptophan can be correlated with the orientation of the peptide, that indicate the type of interaction taking place: the downfield change is related to the formation of H-bonds and is indicative of type II interaction in the SH3 domains, while the upfield shift change is related to unfavorable distances for H-bond formation, and is representative of the type I interaction of SH3 domains. These chemical shift perturbation profiles can thus be used to assess type I or type II orientation of newly studied targets for the N-terminal, but most likely also for other SH3 domains of both CIN85 and CD2AP. Figure IV.16E shows an unusual curvature and strong decrease in intensity during the titration of Cbl-b. The curvature points towards the occurrence of both type I and II interactions, while the strong broadening of NMR signals indicates extensive exchange between multiple conformers, supporting the formation of possibly type I and type II dimers but also a hetero-trimeric complex between CIN85-SH3A domain and Cbl-b.

We also studied the interaction of the SH3A domain of CD2AP with the peptides CD2 and Cbl-b, in order to completely characterize the versatile interactions with atypical proline targets within the CIN85/CD2AP family of adaptor proteins,

ITC (Fig. IV.9A) and SAXS (Fig. IV.19) experiments show that the SH3A domain of CD2AP interacts with the atypical proline-sequence peptide CD2 forming dimers. Using NMR, the observation of specific intermolecular NOEs between the tryptophan indole proton to several side-chains protons in CD2 can only be rationalized by type II binding (Fig. IV.6). Eliminating the electrostatic interactions of Arg332 (**R**₅) with the negatively charged glutamate residues in the RT loops of the SH3 domains by mutating it to an alanine, almost completely abolishes the binding (K_d is three orders of magnitude bigger for the R332A mutant). This indicates that these electrostatic interactions are the main driving force for the SH3A binding to CD2. It also confirms that the lack of a positively charged residue in the N-terminal part of the CD2 peptide at position **R**₂ results in prevalent type II binding orientation. This interaction is enthalpically driven as shown by ITC, with a favorable negative enthalpy change that compensates the positive entropy change (Tables IV.4 and IV.5).

Interestingly, the interaction of the SH3A domain of CD2AP with Cbl-b, that presents the extended atypical proline-rich sequence, consists in the formation of dimers type I and type II, confirmed by the 1:1 stoichiometry of ITC experiments and the *ab initio* reconstructed shape by SAXS.

During the NMR titration of the SH3A domain of CD2AP with Cbl-b, the signals of the Trp37 indole ¹⁵N-¹H cross peaks (Fig. IV.9E) show a curvature, as during the titration of CIN85-SH3A. However, despite that the affinity is one order of magnitude higher, in

this case there is no broadening of the signals, indicating the absence of extensive conformational exchange phenomena. In order to distinguish between the curvatures observed in the NMR titrations of CD2AP and CIN85 with Cbl-b, we tried to simulate them using the K_{ds} resulting from the fitting of chemical shift changes of the Trp indole N ϵ -H ϵ from the NMR titration of the isolated type I (with Cbl-bR911A) and type II (with Cbl-bR904G) interactions (section 4.3.3 and 4.3.5) and considering the existence of independent dimers type I and type II, together with free SH3A domains. The suitability of the model was shown by the fact that the experimental data of the isolated type II and type I interactions can be back-predicted (Fig. IV.10 and IV.18, green type II and blue type I) for both CIN85 and CD2AP-SH3A domains. Using this mathematical approach, the titration of CD2AP-SH3A with Cbl-b can be simulated perfectly (Fig. IV.10B, red). However, the titration of CIN85-SH3A with Cbl-b cannot be simulated, especially at low to intermediate concentrations of Cbl-b (Fig. IV.18, red). This indicates that besides the formation of type I and type II dimers, other multimeric states (e.g. trimer) exist in solution. These simulations are thus in agreement with the stoichiometries determined by ITC and the SAXS analysis.

We can conclude that in the case of CIN85-SH3A, the percentage of population of hetero-trimers is directly related with Cbl-b concentration. The fact that at different concentration of Cbl-b the mechanisms of clustering are different might have important biological implications.

4.4.2 SH3A domains of CD2AP and CIN85 have different decoding mechanisms: biological implications

Even though the SH3A domain of CD2AP is able to interact with Cbl-b forming dimers type I and type II, it is not able to form a trimeric complex involving Cbl-b Arg904 and Arg911 simultaneously, as observed for the SH3A domain of CIN85. This can be rationalized by superimposing its structure on the CIN85-SH3A domain in the X-ray structure of the trimeric complex (Fig. IV.22, PDB code 2BZ8). The length and structural differences of the n-*Src* loop of the SH3A domains of CIN85 and CD2AP are responsible for the ability in forming a trimer; residue Glu35 of one CD2AP-SH3A molecule would clash with Y8 of the other SH3 molecule thus preventing the formation of a trimer with Cbl-b. Moreover, Arg909 (x_3) of the Cbl-b peptide will be hindered to take up the same orientation as in the CIN85-SH3A:Cbl-b trimeric complex. Therefore, only a trimeric arrangement like the one reported in the CMS/Cbl-b crystal structure (Moncalian G. et al., 2006), involving Lys907 (x_1), would be possible. So perhaps CD2AP interaction with Cbl-b is capable of forming type II dimers with Cbl-b Arg911 (R_5), type I dimers with Arg904 (R_2) like CIN85-SH3A, and hetero-trimers involving Arg911 (R_5) in type II orientation and Lys907 (x_1) in type I. This hypothesis is supported by the observation that the affinity of isolated type I binding (Cbl-bR911A) is exactly the same for the CD2AP and CIN85-SH3A domains (Table IV.4 and IV.5) indicating a very similar interaction. According to our results, it appears that, for the SH3A domain of CD2AP, the individual dimers are clearly more stable than a trimeric arrangement, although there are clear differences in experimental conditions with respect to the crystal structures and the *in vivo* assays (Moncalian G. et al., 2006).

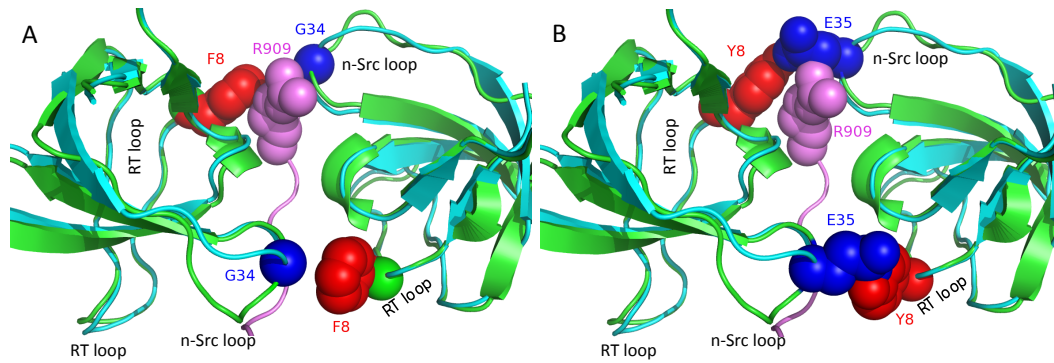


Figure IV.22. Structural comparison of CIN85-SH3A and CD2AP-SH3A domains in complex with Cbl-b. The CD2AP-SH3A structure (green) was superimposed on CIN85-SH3A domains (cyan) in the crystal structure (PDB entry **2BZ8**) of the trimeric complex with Cbl-b (purple). **A.** Residues F8 and G34 in CIN85-SH3A are represented as red and blue spheres, respectively. **B.** Corresponding residues Y8 and E35 in CD2AP-SH3A are shown as spheres. In both panels residue R909 in Cbl-b is also presented as spheres

We can conclude that the fact that the SH3A domains of this family of adaptor proteins interact in a type I or a type II mode is determined by the sequence of the ligand, in particular the pseudosymmetry of positive amino acids capable of interacting with the hydrophilic pocket in the SH3. However, the formation of a trimeric complex is related to the structural differences between the two adaptor SH3 domains, especially in the n-Src loop that will determine if the formation of a trimeric complex is sterically feasible.

We have shown that the study of systems involving versatile interactions requires the use of complementary techniques. In the case of CIN85-SH3:Cbl-b interactions, the combined use of ITC, NMR and SAXS has been crucial to clarify the previous discrepancy about the recognition mode of the SH3A domains and to contribute with novel and crucial aspects of the interaction mechanisms of these domains with atypical proline-rich sequences, providing the structural basis of the signal transduction pathways involving these two adaptor proteins.

Despite the high degree of sequence conservation and domain organization between the two adaptor molecules and many overlapping molecular functions in cultured cells, investigations with mice suggest that CIN85 and CD2AP may have distinct functions *in vivo* (Shih N.Y. et al., 1999; Grunkemeyer J.A. et al., 2005). The differential mode of recognition observed between CIN85 and CD2AP when interacting with their cellular partners would perhaps provide a molecular explanation for the distinct biological behavior observed between the two adaptors, contributing to complement the existence of another level of organization within the molecular clustering involved in receptor downregulation and signaling.

CHAPTER V

HIGH-RESOLUTION STRUCTURE OF UBZ1+2 DOMAINS OF THE UBIQUITIN-BINDING ADAPTOR PROTEIN TAX1BP1 IN SOLUTION

&

STRUCTURAL AND THERMODYNAMIC CHARACTERIZATION OF THE INTERACTION OF UBZ1+2 OF TAX1BP1 WITH UBIQUITIN*

* UBZ DOMAINS OF TAX1BP1

Part of the work presented in this Chapter is ready to be submitted with the title:

High-resolution structures of the UBZ1+2 tandem of the ubiquitin-binding adaptor protein TAX1BP1

Ceregido M.A., Spinola M., Buts L., Garcia-Pino A., Rivera-Torres J., Bravo J. and van Nuland N. A.J.

One never notices what has been done; one can only see what remains to be done.
Marie Curie.

Chapter V. UBZ DOMAINS OF TAX1BP1

5.1 Introduction

The immune system protects against injury or microbial infection by functioning as an internal defense mechanism activating an inflammatory response. It interacts continuously with the environment and uses proteins as sensors and effectors to modulate a protective response. Inflammatory responses are regulated to prevent chronic inflammation, related to many cancers and autoimmune diseases (Grivennikov S.I. et al., 2010).

Nuclear Factor kappa B (NF- κ B) acts as a regulator of genes encoding molecules that control innate and adaptive immune responses (Vallabhapurapu S. and Karin M., 2009), inflammation, cell proliferation, apoptosis and oncogenesis (Hayden M.S. and Ghosh S., 2004). Cytokines, such as tumor necrosis factor (TNF) and interleukin (IL)-1, are potent activators of NF- κ B that can initiate pro-inflammatory signaling in other cells throughout our body (Chapter I, section 1.6.1). However, the uncontrolled activation of this system leads to pathologic states such as inflammatory cardiac valvulitis, chronic inflammation, restriction of the antiviral immune response and septic shock.

For this reason, the cell develops a negative regulation feedback of the NF- κ B pathway, in which the ubiquitin adaptor protein TAX1BP1 plays an essential role as a regulator of innate immunity (Verstrepen L. et al., 2011).

TAX1BP1 (Chapter I, section 1.6) has an essential role in the negative regulation of NF- κ B signaling by contributing to the recruitment and function of the deubiquitinating enzyme A20 (Shembade N. et al., 2007). It is also involved in the negative regulation of cell growth, apoptosis, and NF- κ B dependent gene expression, but in some cases TAX1BP1 can also positively influence gene expression (Verstrepen L. et al., 2011) (see Chapter I, section 1.6.1).

TAX1BP1 is an 86 kDa protein highly conserved in eukaryotes (Chapter I, section 1.6.1). The C-terminal part of TAX1BP1 contains a tandem of two ubiquitin zinc finger domains (UBZ1 and UBZ2) with highly conserved “PPXY” motifs (where “X” is any amino acid), which are known binding motifs for modular WW domain-containing proteins (Sudol M. et al., 1995), like the E3 ligase Itch (Shembade N. et al., 2008).

Experimental evidence suggests that in the NF- κ B, the c-Jun N-terminal kinase (JNK) and the interferon regulatory transcription factor (IRF3) signaling pathways (Shembade N. et al., 2007; Shembade N. et al., 2008; Iha H. et al., 2008; Verstrepen L. et al., 2011) TAX1BP1 acts as a crucial ubiquitin-binding adaptor protein that recruits a ubiquitin-editing complex consisting of the A20 deubiquitinating enzyme, RNF11 and the E3 ubiquitin ligase Itch for the modification of specific signaling molecules (Verstrepen L. et al., 2010; Verstrepen L. et al., 2011). A recent publication suggests that phosphorylation of TAX1BP1 (at Ser593 and Ser624) by the kinase IKK α is critical for

the inducible interactions among proteins of the A20 ubiquitin-editing complex (Shembade N. et al., 2011). Furthermore, phosphorylation of TAX1BP1 triggers a conformational change, which exposes the zinc fingers and the PPXY motif (Shembade N. et al., 2011). It is known that cells lacking TAX1BP1 or Itch display persistent activation of IKK and JNK after stimulation with TNF α or IL-1 β (Iha H. et al., 2008; Chiu Y.H. et al., 2009). It has also been reported that TAX1BP1 recruits A20 to the ubiquitinated signaling proteins TRAF6 and RIP1 (Lys63-polyubiquitinated proteins), leading to their A20-mediated deubiquitination and the disruption of IL-1 and TNF-induced NF- κ B signaling, respectively (Shembade N. et al., 2007; Iha H. et al., 2008).

While it is clear that TAX1BP1 acts as a crucial ubiquitin-binding adaptor protein, it is still unknown how the UBZ domains of TAX1BP1 interact with ubiquitin. Despite some immunologic studies suggest that both UBZ domains and PPXY motifs are essential for TAX1BP1 NF- κ B inhibitory function (Shembade N. et al., 2008), the role of UBZ1 is especially unclear. Until now its interaction with ubiquitin has never been reported, and UBZ2 alone was considered essential for the ubiquitin-binding potential of TAX1BP1 (Iha H. et al., 2008; Verstrepen L. et al., 2011).

Up-to-date, the structure of only two single members of the family of UBZ domains have been reported: the NMR structure of the UBZ domain of the human DNA Y-polymerase eta (DNA pol eta) (Bomar M.G. et al., 2007) (PDB code 2I50) and the X-ray structure of UBZ domain of the human Werner helicase interacting protein 1 (WRNIP1) (Suzuki N. et al., 2010) (PDB code 3VHS).

In this chapter we report the high-resolution NMR structure of the two zinc finger domains of TAX1BP1 in tandem (UBZ1+2) (PDB code 2M7Q). The X-ray structure of UBZ1+2 (PDB code 4BMJ) was previously determined by Dr. Spinola M. (Ceregido M.A. et al., 2013). We carried out an in-depth study of the interaction of the separate UBZ1 and UBZ2 domains, and the UBZ1+2 tandem with ubiquitin by ITC and NMR. Using paramagnetic relaxation enhancement (PRE) NMR spectroscopy, we have obtained an accurate structural model of the UBZ1+2-ubiquitin complex. Finally, this research was complemented, by a set of bio-layer interferometry experiments to determine the avidity of the UBZ1+2 for the three principal types of polyubiquitin chains: Lys48-, Lys63- and linear polyubiquitin.

5.2 Experimental Procedures

5.2.1 Purification of ^{15}N - ^{13}C -labeled UBZ1+2 of TAX1BP1

The pET-28a plasmid containing the human UBZ1+2 gene cloned downstream of a hexahistidine-tag was expressed in *E. coli* BL21 (DE3) strain. For unlabeled UBZ1+2, overexpression and purification were carried out as described in Chapter III, section 3.1.

For the ^{15}N and ^{15}N - ^{13}C -labeled UBZ1+2, cells were grown in 2xTY medium at 37 °C till an OD₆₀₀ of 0.7 was reached. After two washing steps, cells were resuspended in minimal medium M9 (see Chapter III, section 3.1.3). Protein expression was induced with 0.2 mM IPTG, and cultures incubated at 20 °C for 24 h to avoid the formation of inclusion bodies. Cells were harvested by centrifugation at 4,000 rpm (BECKMAN

COULTER, J-LITER JLA-8.1000 rotor) for 20 min, resuspended in lysis buffer with 1 mM TCEP (see Chapter III, section 3.1.3), and lysed in a cell cracker (HAIVA of CONSTANTSYSTEMS Ltd.). TCEP [(Tris(2-carboxyethyl)phosphine)] was added to reduce cysteine residues of the UBZ1+2; in its absence, the protein remains unfolded due to the inefficient coordination of the zinc atoms. After centrifugation at 20,000 rpm (BECKMAN COULTER, JA-20 rotor) for 30 min at 4 °C, the supernatant was loaded on a 5 mL HisTrap Ni-chelating column (GE Healthcare) and eluted with a linear gradient of 20 to 500 mM imidazole (Fig. V.1A). The His-tagged UBZ1+2 eluted at 31% of buffer B (50 mM Tris pH 7.5, 200 mM NaCl, 500 mM imidazole and 1 mM TCEP).

To remove the His-tag, fractions containing UBZ1+2 were incubated overnight with Protease preScission (GE Healthcare) at room temperature (10 units/mg of protein were used to ensure total cleavage) (Fig. V.1C). The cleaved UBZ1+2 was further purified on a Superdex-75 16/60 gel-filtration column (GE Healthcare) (Fig. V.1B). Pure protein sample was collected in a single peak at an elution volume of 79 mL (Fig. V.1B), indicating that UBZ1+2 behaves like a monomer (8.3 kDa) according to the column calibration curve. Sample concentration was determined spectrophotometrically (UV-VIS spectrophotometer Cary 100Bio, Agilent) using an extinction coefficient $\epsilon_{280\text{nm}} = 8,480 \text{ M}^{-1}\text{cm}^{-1}$, estimated with the ProtParam tool (Expasy). The purified UBZ1+2 was reconstituted with Zn^{2+} by a stepwise addition of 2.2 molar equivalents of ZnCl_2 (always in presence of TCEP). The sample was flash-frozen in liquid nitrogen and stored at -20 °C.

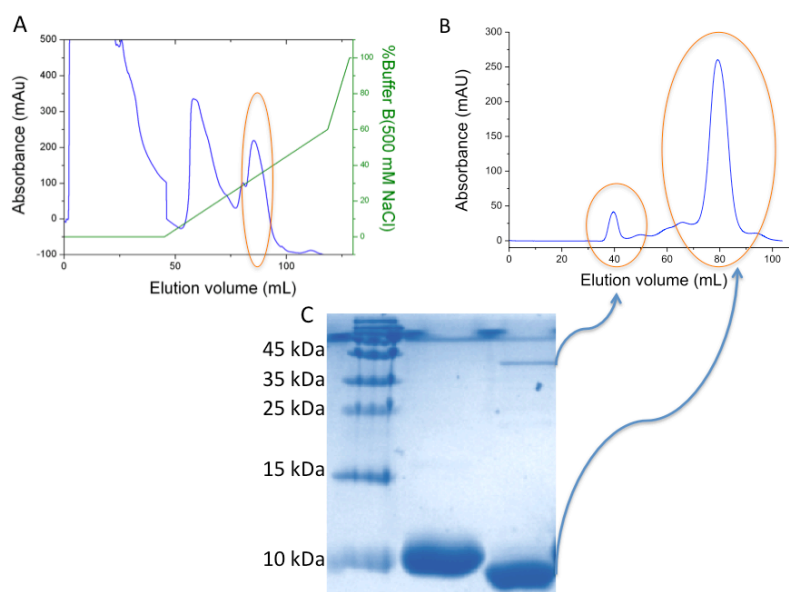


Figure V.1. Summary of the ^{15}N - ^{13}C -labeled UBZ1+2 purification. **A.** Purification of UBZ1+2 on a HisTrap Ni-chelating column. The blue line indicates $UV_{280\text{nm}}$ absorption of eluted sample, the green line shows the percentage of buffer B (500 mM Imidazole) and the orange oval indicates the collected UBZ1+2 fractions. **B.** Elution of the His-tag free UBZ1+2 from a Superdex-75 gel-filtration column. The peaks at elution volumes of 41 mL and 79 mL correspond to the PreScission protease and the UBZ1+2, respectively. **C.** SDS-PAGE showing the cleavage of the His-tag with PreScission protease. First lane: marker (Bio-Rad Life Science), molecular weights of the standards are indicated on the left side of the gel; second lane: UBZ1+2 sample before cleavage of the His-tag (10.4 kDa); third lane: UBZ1+2 sample after the cleavage (8.3 kDa). A faint high-molecular weight band seen in the third lane corresponds to the PreScission protease (46 kDa).

The UBZ1+2 studied here corresponds to the C-terminal sequence of the human TAX1BP1 protein (residue 725 - 789); however, in this work we used UBZ1+2 residue numbering 1 - 69, starting from the four N-terminal residues that remain after the cleavage of the His-tag (Fig. V.2.).

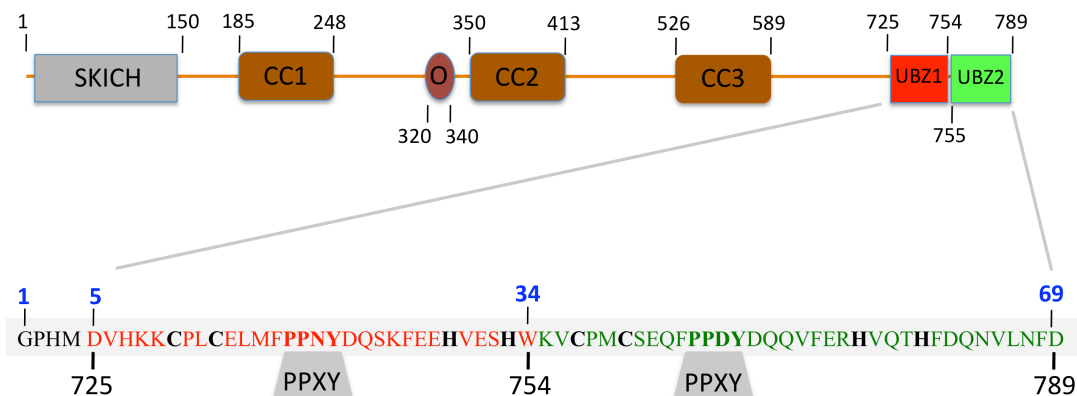


Figure V.2. Schematic domain structure of TAX1BP1. The numbers indicate the amino acids that delimit the structural domains. The N-terminus contains a SKICH (skeletal muscle and kidney enriched inositol phosphatase carboxyl homology) domain, the central part contains three Coiled-coil (CC1, CC2 and CC3) structures and a region responsible for homodimerization (O), and the C-terminal extreme contains the two zinc finger domains UBZ1 (red) and UBZ2 (green). The sequence of the UBZ1+2 tandem is indicated below, including the four residues (GPHM) that remain after the removal of the His-tag. The residue numbering used in this work is given in blue, while the numbering of the equivalent residues within the complete TAX1BP1 protein is in black. The conserved PPXY motif is highlighted in both domains and the cysteine and histidine residues that coordinate the zinc atoms are in **bold**.

Isolated UBZ1 and UBZ2 domains were purchased from ChinaPeptides (Shanghai, China). Both of them were N-terminal acetylated and C-terminal amidated. Before the experiments, the proper amount of lyophilized UBZ1 or UBZ2 was weighted on an analytical balance (precision: ± 0.1 mg) and dissolved in the appropriate buffer. In order to ensure the proper folding of the proteins, 1 molar equivalent of TCEP and 1.3 molar equivalents of $ZnCl_2$ were added to the sample to provide a reducing environment and induce the folding of the domains, respectively. The proper folding of the isolated UBZ1 and UBZ2 domains was checked by circular dichroism (CD) spectroscopy. When necessary, UBZ1 and UBZ2 domains were dialyzed using mini dialysis membranes with a 500 Da cut-off (GE Healthcare).

Single UBZ1 and UBZ2	
UBZ1	Ac-DVHKKCPLCELMFPPNYDQSKFEEHVESH-NH ₂
UBZ2	Ac-ESHWKVCPMCSEQFPDYYDQQVFERHVQTHFDQNV-NH ₂

Table V.1. Sequence of the isolated UBZ1 and UBZ2 domains purchased from ChinaPeptides (Shanghai, China). A Trp residue was included in UBZ2 domain for accurate concentration measurement.

5.2.2 Purification of ubiquitin and its single-cysteine mutants

E. coli BL21 (DE3) cells harboring the human ubiquitin in a pET-21b plasmid were grown at 37 °C in LB or M9 minimal media (in order to overexpress unlabeled or isotopically labeled ubiquitin respectively) until an OD₆₀₀ of 0.7 was reached. Ubiquitin expression was induced with 1 mM IPTG, and cultures incubated at 37 °C for 18 h. The cells were harvested by centrifugation and lysed in a cell cracker (HAIWA of CONSTANTSYSTEMS Ltd.). The lysate was centrifuged at 20,000 rpm (BECKMAN COULTER, JA-20 rotor) for 30 min at 4 °C and the ubiquitin-containing supernatant dialyzed (G2 dialysis cassettes, molecular weight cut-off 3.5 kDa, Thermo Scientific) against 50 mM Tris-HCl pH 8.0, until the conductivity of the sample was less than 5 mS/cm. The protein was purified sequentially on a source Q30 column (GE Healthcare) using a salt gradient (from 0 to 1M NaCl) and a Superdex-75 16/60 column (GE Healthcare) equilibrated in 50 mM Tris-HCl pH 8.0, 200 mM NaCl. Sample concentration was determined spectrophotometrically (UV-VIS spectrophotometer Cary 100Bio, Agilent) using an $\epsilon_{280\text{nm}} = 1,490 \text{ M}^{-1}\text{cm}^{-1}$, determined by the ProtParam tool (ExPASy). Finally, the sample was flash-frozen in liquid nitrogen and stored at -20 °C.

The plasmids encoding single-cysteine ubiquitin mutants (E51C, D39C and D32C) were prepared by GenScript (site-directed mutagenesis service, Piscataway, NJ, USA) starting from the wild type (WT) ubiquitin pET-21b vector. Ubiquitin Cys-mutants were expressed and purified analogously to the WT protein (section 5.2.2) (Fig. V.3A and B), except for the addition of 1 mM DTT to the purification buffers, to avoid the formation of intermolecular disulfide bonds.

5.2.2.1 Introduction of a paramagnetic probe in ubiquitin

Before the probe attachment, the working buffers were treated with chelating beads (Chelex Sigma-Aldrich) in order to remove traces of transition metal ions. Purified single-cysteine ubiquitin mutants (E51C, D39C and D32C) were incubated with 10 mM DTT (dithiothreitol) in 20 mM Tris-HCl at pH 8.0 for 10 min at room temperature to reduce possible intermolecular disulfide bonds. DTT was removed by passing the ubiquitin solution through a 5 mL desalting HiTrap column (GE Healthcare) (Fig. V.3C). The resulting monomeric ubiquitin was reacted with a 5-fold excess (relative to the protein) of the N-[S-(2-pyridylthio)cysteaminy] ethylenediamine-N,N,N',N'-tetraacetic acid monoamine (pEDTA, Toronto Research Chemicals) and 2 molar excess (relative to the label) of Mn²⁺ (MnCl₂ from Sigma-Aldrich) or Zn²⁺ (ZnCl₂ from Sigma-Aldrich) and incubated overnight at room temperature. Once the reaction was complete, the conjugated protein was separated from an unreacted label on a 5 mL HiTrap SP-column (GE Healthcare) (Fig. V.3D), equilibrated in 20 mM MES pH 5.0, eluted with a linear gradient of 0 - 0.5 M NaCl (30 column volumes) in the same buffer. For all mutants, the labeling reaction was highly efficient, and a single peak of conjugated protein was eluted at 21% of buffer B (see Fig. V.3D). Finally, the modified ubiquitin was dialyzed (G2 dialysis cassettes, molecular weight cut-off 3.5 kDa, Thermo Scientific) against the NMR buffer (50 mM NaPi pH 6.0) and concentrated using Amicon Ultra concentrators (Millipore). In each case, the yield of labeling was $\geq 98\%$ as estimated from the percentage of free Cys using the DTDP (4,4'-dithiodipyridine) assay (see Chapter III, section 3.3.4). The label attachment was also confirmed by mass spectrometry (De Duve Institute, platform for proteomic and protein analysis MASSPROT, Brussels, Belgium).

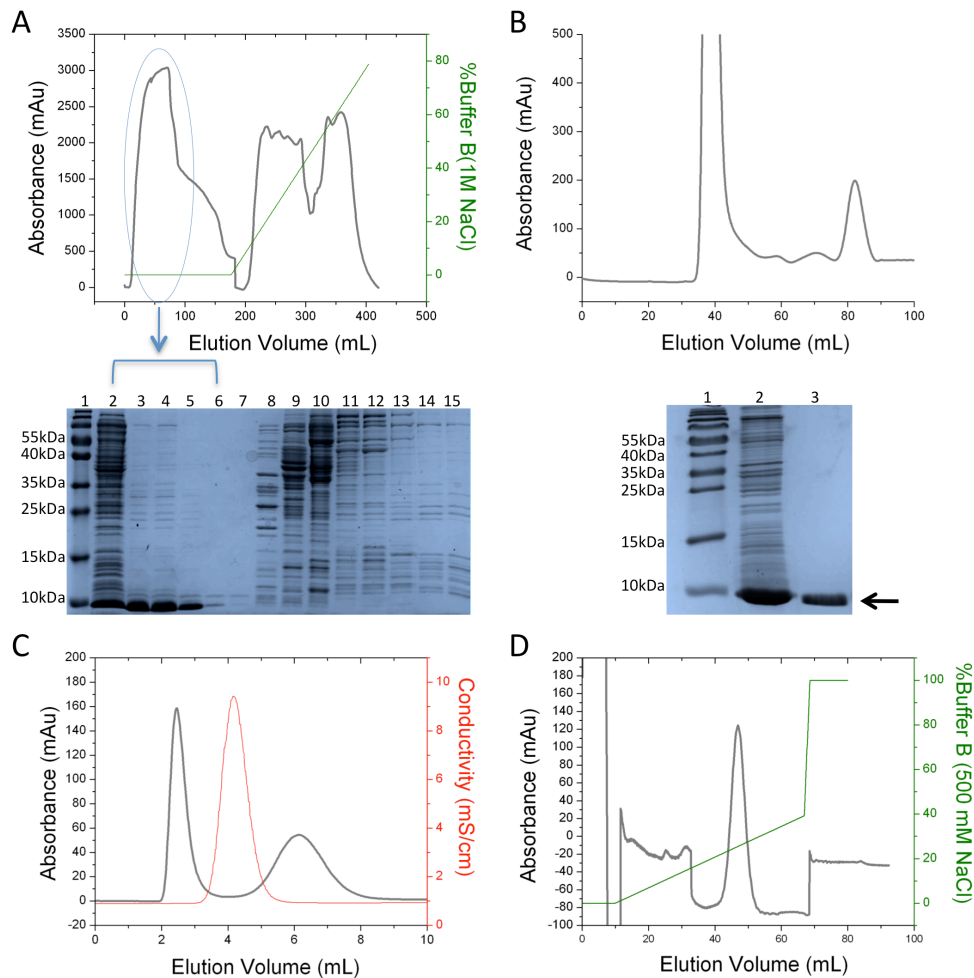


Figure V.3 Summary of the single-cysteine ubiquitin purification and labeling with the paramagnetic probe. **A.** Top: purification of the ubiquitin E51C mutant on a source Q30 anion exchange column (GE Healthcare). The protein is found in the flowthrough (indicated by a blue oval); the contaminants are eluted with the salt gradient. Bottom: SDS-PAGE of the fractions eluted from the source Q30 column. Lane 1: molecular weight marker (Bio-Rad Life Science); lane 2: protein sample injected onto the column. Lanes 3 - 6: ubiquitin-containing flowthrough fractions (shown by a blue oval on the chromatogram above); lanes 7 - 15: fractions eluted during the salt gradient. **B.** Top: elution of the ubiquitin E51C mutant from a Superdex-75 16/60 gel filtration column, appearing as a single peak at an elution volume of 81 mL. Bottom: SDS-PAGE of the ubiquitin sample eluted from the Superdex-75 column. Lane 1: molecular weight marker (Bio-Rad Life Science); lane 2: the sample injected onto the column; lane 3: the eluted sample. **C.** Chromatogram of the 5 mL desalting HiTrap column. The black and red lines indicate UV_{280nm} absorption and the conductivity in mS/cm of the eluted sample, respectively. The elution profile shows a good separation of the ubiquitin (first peak) from the DTT used to reduce the intermolecular disulfides (second peak). **D.** Purification of the conjugated ubiquitin (E51C-pEDTA-Mn²⁺) on a 5 mL HiTrap cation exchange SP-column. The black and green lines indicate respectively, UV_{280nm} absorption (for the eluted sample) and the percentage of buffer B (containing 500 mM NaCl). The pure ubiquitin-Cys sample labeled with the paramagnetic pEDTA-Mn²⁺ probe or the diamagnetic pEDTA-Zn²⁺ probe, appears in a single peak at 21% of buffer B.

5.2.3 Isothermal Titration Calorimetry

ITC experiments were performed on a MicroCal VP-ITC200 (GE Healthcare) at 25 °C in 50 mM cacodylate buffer at pH 6.0 and 1 mM TCEP. In a typical experiment a ubiquitin solution (2 mM) was placed in the syringe and injected into a sample of UBZ1+2 or UBZ2 (146 μ M) placed in the cell using a volume profile simulated in MATHCAD (PTC Mathcad) (see Chapter III, section 3.2.1). The ITC titrations of isolated UBZ1 domain were also performed on an ITC200, where 3.2 mM ubiquitin (in the syringe) was titrated into 130 μ M UBZ1 (in the cell). The binding constants and thermodynamic parameters for the interaction of the isolated UBZ domains with ubiquitin were calculated by fitting a general model of n identical and independent sites (Chapter III, section 3.2.2) to the experimental data. The experimental isotherm obtained during the titration of UBZ1+2 with ubiquitin was also analyzed using a two non-symmetrical binding site model, taking into account possible cooperativity effects, using the global analysis platform of SEDPHAT (see Chapter III, section 3.2.1). All the ITC experiments were repeated two times, and a detailed report of the thermodynamic parameters and their errors, calculated as the standard deviations from the mean, is presented in section 5.3.2.

5.2.4 Nuclear Magnetic Resonance Spectroscopy

^{15}N and ^{15}N - ^{13}C -labeled samples (0.7 mM) were exchanged into the NMR buffer (50 mM NaPi pH 6.0, 1 mM TCEP) containing 10% (v/v) D_2O and 30 μM ZnCl_2 . All NMR experiments were performed at 25 °C on Varian 600 and 800 MHz spectrometers, the latter equipped with triple-resonance pulsed-field gradient probes. The backbone and side chain ^1H , ^{15}N and ^{13}C resonances were assigned based on the spectra described in chapter III, section 3.3.3.1.

Structure calculation was done in CYANA version 2.1 (Guntert P. et al., 1997; Guntert P., 1997; Hermann T. et al., 2002) and refinement in RECOORD (Nederveen A.J. et al., 2005) as described in Chapter III, section 3.3.3.2. The twenty lowest energy structures (Fig. V.5) were used for the final analysis. Input data and structure calculation statistics are summarized in Table V.2. PROCHECK-NMR (Laskowski R.A. et al., 1996) and WHATIF (Vriend G., 1990) were used to analyze the quality of the structures. MOLMOL (Koradi R. et al., 1996) and PYMOL (<http://www.pymol.org/>) were used for further analysis and visualization

NMR titrations of the ^{15}N -labeled ubiquitin were performed with a 0.3 mM sample titrated with increasing amount of UBZ1+2, UBZ1 and UBZ2, one at a time (stock solutions of 3 - 4 mM). Series of 2D ^1H - ^{15}N -HSQC spectra were obtained, and the chemical shift changes were measured at each molar ratio. Similarly, ^{15}N - ^{13}C -labeled UBZ1+2 (0.3 mM) was titrated with unlabelled ubiquitin (stock solution of 4 mM). ^1H - ^{15}N -HSQC and ^1H - ^{13}C -HSQC spectra were recorded at each titration point. Average (^1H , ^{15}N) chemical shift perturbations were calculated as explained in Chapter III (section 3.3.3.3). The apparent dissociation constants (K_{ds}) for the ubiquitin binding to UBZ1+2 and the isolated UBZ1 and UBZ2 domain were obtained by fitting the titration curves using a cooperative binding model, allowing to extract the Hill coefficient (Chapter III, section 3.3.3.3).

The relaxation parameters ^{15}N R1, R2 and ^1H - ^{15}N steady-state NOE for ^{15}N -labeled UBZ1+2 in the free form and bound to ubiquitin at different ratios (1:0.5, 1:1.1, and 1:5.0) were measured at 600 MHz and 25 °C. Relaxation values were obtained from series of 2D experiments with coherence selection achieved by pulse field gradients using the experiments previously described (Chapter III, section 3.3.5). For data analysis and interpretation, the resonances with spectral overlap or severe line broadening were not considered. The relaxation parameters of the following backbone amides could not be obtained:

- **Free form:** Asp5, Val6, His7, Lys8, Lys9, Cys13, Leu15, Arg55, His56, Thr59.
- **Ratio 0.5:** His3, Met4, Asp5, Val6, Lys9, Leu15, Asn20, Glu28, Ser32, Trp34, Lys35, Met39, Glu42, Phe44, Asp47, Arg55, Thr59, Asp62.
- **Ratio 1.1:** His3, Asp5, Val6, Lys8, Lys9, Cys10, Leu12, Glu14, Leu15, Met16, Asn20, Asp22, Phe26, Glu28, Ser32, Trp34, Lys35, Glu42, Gln43, Phe44, Asp47, Arg55, Thr59.
- **Ratio 5.0:** His3, Asp5, Val6, His7, Lys8, Lys9, Cys10, Leu12, Cys13, Glu14, Leu15, Asn20, Asp22, Phe26, Glu28, Trp34, Lys35, Cys37, Glu42, Phe44, Asp47, Arg55, Thr59.

All the data were processed with NMRpipe (Delaglio F. et al., 1995) and analyzed in CCpNMR (Vranken W.F. et al., 2005) or NMRview (Johnson B.A., 2004). The R1 and R2 values were determined as described in Chapter III, section 3.3.5.

5.2.5 Paramagnetic Relaxation Enhancement

NMR samples contained 0.3 mM ^{15}N -labeled UBZ1+2 and 3 molar equivalents of D32C, D39C, or E51C ubiquitin-pEDTA conjugates harboring a paramagnetic (Mn^{2+}) or diamagnetic (Zn^{2+}) metal. Measurements were performed in 50 mM NaPi pH 6.0, and 6% D_2O for lock at 25 °C on a Varian 800 MHz spectrometer equipped with a triple-resonance salt-tolerant cold probe. For each ubiquitin mutant two ^1H - ^{15}N -HSQC spectra were recorded: one for the paramagnetic sample and other for the diamagnetic control. The same acquisition parameters (1024 complex points in the direct dimension, 64 increments, spectral widths of 23 ppm (^{15}N) and 14.2 ppm (^1H), 16 scans and a relaxation time delays of 3 s) were used throughout.

For the data analysis and interpretation, only the UBZ1+2 residues that show no considerable broadening or spectral overlap were considered. Based on these criteria, the following backbone amides were not analyzed: Asp5, Val6, Lys8, Lys9, Leu12, Leu15, Asn20, Glu28, Trp34, Lys35, Glu42, Phe44, Asp47, Arg55 and Thr59. Data processing, analysis and subsequent PRE-based ensemble refinement are described in Chapter III, section 3.3.4.

5.2.6 Small Angle X-ray Scattering

Samples of UBZ1+2 (7 mg/mL) and the isolated UBZ1 and UBZ2 domains (4 mg/mL) in absence and presence of ubiquitin (1.5 and 3.0 molar equivalents for the isolated domains, and 1.5, 3.8 and 6.6 molar equivalents for UBZ1+2) were prepared in 50 mM Hepes at pH 7.4, 100 mM NaCl and 1 mM TCEP.

SAXS data for characterizing the ubiquitin complexes with UBZ1+2, UBZ1 and UBZ2 domains were collected at the SWING beamline on the SOLEIL synchrotron (Gif-sur-Yvette, France) using a 17x17 cm² low-noise Avix CCD detector positioned at a distance of 2.1 m from the sample. The sample solutions were passed through a KW402.5-4F column (Shodex) (for proteins with molecular weight between 1 and 75 kDa), positioned within a vacuum chamber. 150 frames of 0.75 s each were collected, normalized to the transmitted intensity, and subsequently averaged using the image analysis software Foxtrot (SWING beamline at SOLEIL Synchrotron). The scattering curves were processed as described (Chapter III, section 3.4).

5.2.7 Bio-Layer Interferometry

Bio-layer interferometry (BLI) experiments were performed on an OctetRed 96 system (Fortebio) at 25 °C. Amine reactive sensors (AR2G) were coupled with linear, di- and tetraubiquitin, Lys63- and Lys48-linked diubiquitin chains at 25 µg/mL in 50 mM acetic acid pH 4.0, via EDC/sulfo-NHS mediated amide bond formation (Chapter III, section 3.6). Streptavidin high binding (SA) biosensors were loaded with biotinylated Lys63- and Lys48-linked tetraubiquitin, and N-terminal monoubiquitin at 25 µg/mL in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM DTT and 0.05% (v/v) tween-20. After loading, the biosensors were washed with 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM DTT, and 0.05% (v/v) tween-20. Association and disassociation steps were carried out for 600 - 900 s, followed by a regeneration step of 300 s. Different concentrations (from 0 to 300 µM) of UBZ1+2 were tested for their binding to the immobilized polyubiquitin chains. Steady-state analysis in “Data Analysis version 7.0” (Fortebio) was used to calculate the apparent affinities (K_d) (Chapter III, section 3.6.2).

5.3 Results and discussion

5.3.1 Solution NMR structure of the free UBZ1+2

5.3.1.1 Sequential assignment and structure calculations

The NMR assignments of N, H^N, H^α, H^β, CO, C^α, and C^β resonances of UBZ1+2 were obtained from the identification of intra- and inter-residue connectivities in HNCACB, CBCA(CO)NH, HBHA(CO)HN and HNCO spectra at the ¹H, ¹⁵N frequencies of every peak in the ¹H-¹⁵N-HSQC spectrum. These backbone assignments were extended to the side chain signals using correlations within CCH-TOCSY, (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE spectra. Remaining aromatic ¹H and ¹³C frequencies were assigned from the ¹H-¹³C-HSQC and ¹³C-NOESY-HSQC spectra. Side-chain NH₂ resonances of glutamines and asparagines were assigned from HNCACB, CBCA(CO)NH and 3D ¹⁵N-NOESY-HSQC. All ¹H, ¹³C and ¹⁵N resonances were verified by 3D ¹⁵N- and ¹³C-NOESY-HSQC spectra (Chapter III, section 3.3.3).

The semi-automatic assignment protocol resulted in the full backbone assignment and a ¹H side-chain assignment of 98% completeness. The ¹H, ¹⁵N and ¹³C chemical shifts have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>) under accession number 19201. Figure V.4 shows an assigned ¹H-¹⁵N-HSQC spectrum of UBZ1+2.

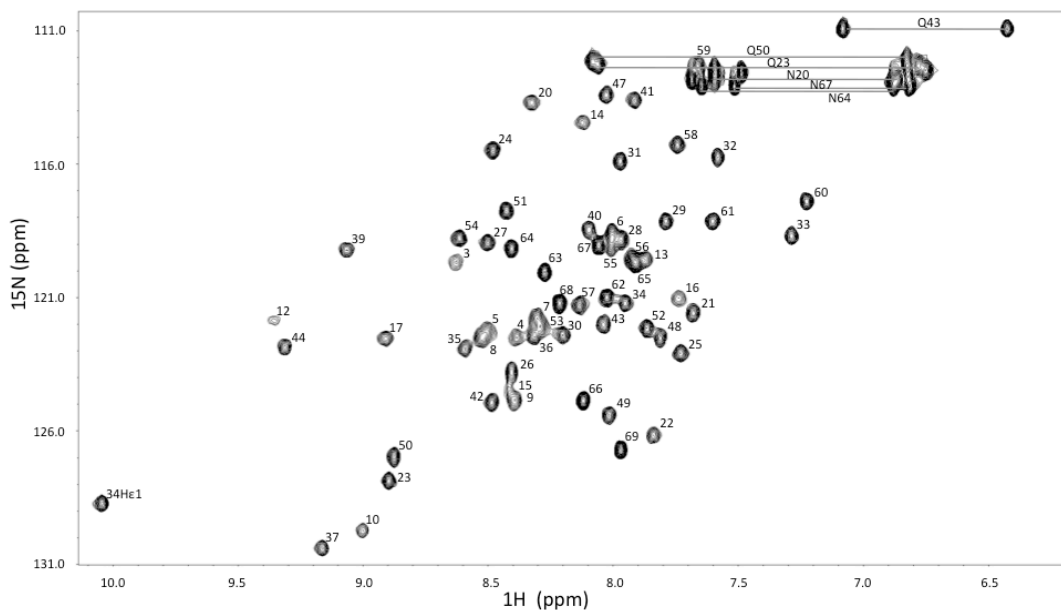


Figure V.4. Assigned 800 MHz ^1H - ^{15}N -HSQC spectrum of UBZ1+2 at pH 6.0 and 25 °C. All backbone ^1H - ^{15}N pairs are indicated by their corresponding number in the amino acid sequence. The H ϵ 1 proton of the Trp34 is indicated. Side-chain NH_2 resonances of Asn and Gln residues are connected by solid lines.

Figure V.5 shows the superposition of the ensemble of 20 lowest-energy structures derived from the automated NOE assignment protocol in CYANA and the subsequent water refinement in RECOORD. The structural coordinates and experimentally derived restraints have been deposited in the Protein Data Bank (PDB) with accession number 2M7Q.

The UBZ1+2 tandem is composed of two UBZ domains that present the classical Cys₂His₂ zinc finger fold (Berg J.M., 1988; Gibson T.J. et al., 1988), tightly connected with no linker region between them. Both UBZ domains present a compact structure comprising two short antiparallel β -sheets followed by an α -helix forming a left-handed $\beta\beta\alpha$ -unit, where the zinc atom is coordinated by Cys10, Cys13, His29 and His33 in the UBZ1 domain and by Cys37, Cys40, His56 and His60 in UBZ2. Tetrahedral coordination of the zinc atoms is crucial to maintain the 3D structure of both zinc finger domains by packing together the β -sheets with the α -helix (Fig. V.7F).

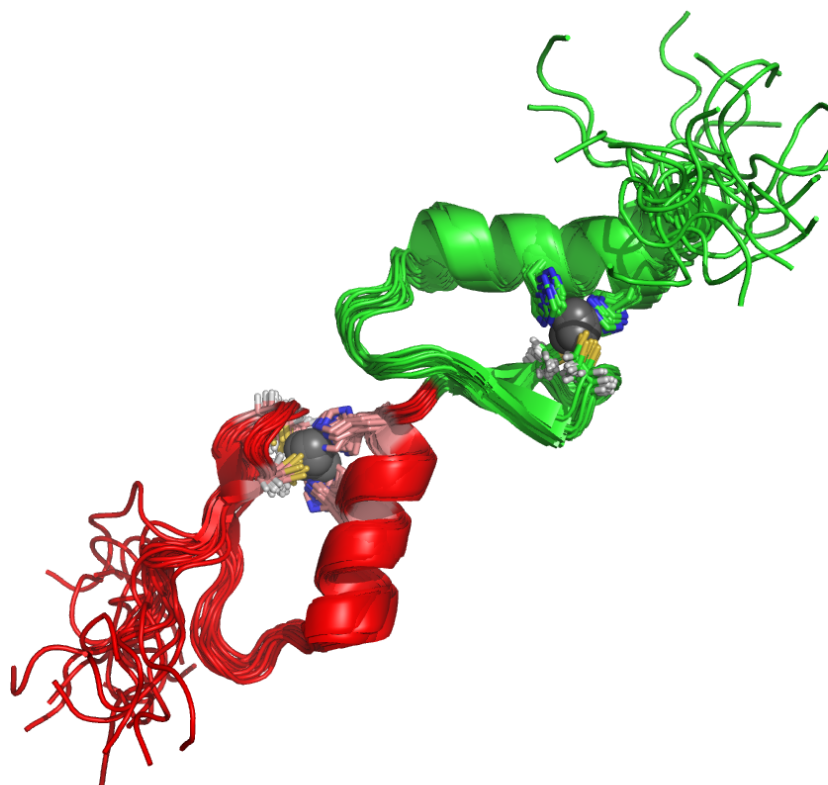


Figure V.5. Cartoon representation of the ensemble of 20 water-refined NMR structures superimposed on the backbone of residues 8 - 62. The UBZ1 and UBZ2 domains are colored red and green, respectively. The side-chains of the two Cys₂His₂ motifs coordinating the two zinc ions (grey spheres) are shown in sticks.

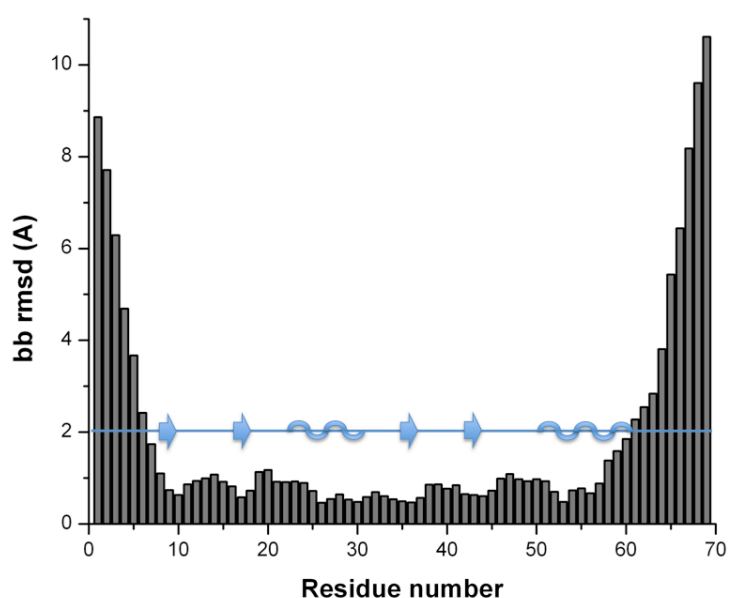


Figure V.6. Backbone variations over the ensemble of 20 NMR structures. Secondary structure elements are drawn within the graphic.

Experimental restraints^a	
Distance Restraints	1141
Intra-residue (i-j = 0)	305
Sequential (i-j = 1)	370
Medium range (2 ≤ i-j ≤ 4)	295
Long range (i-j ≥ 5)	171
Dihedral restraints	
Torsion angle (phi/psi)	40/40
Restraints statistics	
NOE violations >0.5 Å	0.2 ± 0.4
Dihedral violations >5°	0
CNS energies (kcal/mol)	
Etotal	-2393.8 ± 25.7
Evdw	-251.9 ± 12.4
Eelec	-2787.2 ± 38.9
Rmsd from average for residues 8-62 (Å)^b	
Backbone N, CA, C'	0.87 ± 0.20
Heavy atoms	1.19 ± 0.17
Ramachandran plot^c	
Most favored regions (%)	91.8
Additional allowed regions (%)	7.3
Generously allowed regions (%)	0.8
Disallowed regions (%)	0

Table V.2: Structural statistics over the twenty lowest-energy water-refined NMR structures of UBZ1+2. (^a)Restraint statistics reported for all NOEs. (^b)Coordinate precision is given as the pair-wise Cartesian coordinate root mean square deviations from the average structure over the ensemble. (^c)Values obtained from the PROCHECK-NMR analysis over all residues including the first seven highly flexible residues.

The UBZ1+2 solution structure fulfills the experimental data very well, presenting excellent Ramachandran statistics (Table V.2) and backbone r.m.s.d. of 0.87 ± 0.20 over the NMR ensemble (Table V.2 and Fig. V.6), which reflects the high quality of the 3D structure. The success of the approach can be attributed to the use of high-quality 800 MHz 3D NOESY and 2D NOESY spectra, a nearly complete set of proton assignments used for the automatic NOE assignment, and the addition of torsion angle restraints derived from the combination of selected ¹H, ¹⁵N and ¹³C chemical shifts.

However, complete structural characterization of a protein requires not only the definition of the three-dimensional position of each atom but also the study of its dynamics.

5.3.1.2 Backbone dynamics of free UBZ1+2

Backbone ¹⁵N R1, R2 and hetNOE were measured for the 50 non-overlapping cross peaks of UBZ1+2 (Fig. V.7). Almost all hetNOEs (Fig. V.7A) are lower than the theoretical upper limit value of 0.84 at 600 MHz (Kay L.E. et al., 1989). The low (negative) NOE values of the residues at the N-terminus (UBZ1 domain) and at the C-terminus (UBZ2 domain), indicate that the terminal regions are much more flexible than the rest of the protein. Elevated R2 values (Fig. V.7C) are observed especially for the

residues located in the UBZ1 domain (residues 16 - 33), whereas the R1 rates are nearly the same in the whole UBZ1+2 molecule (Fig. V.7B). Elevated R2 values are indicative of conformational exchange on the microsecond-millisecond time scale (Blake-Hall J. et al., 2004).

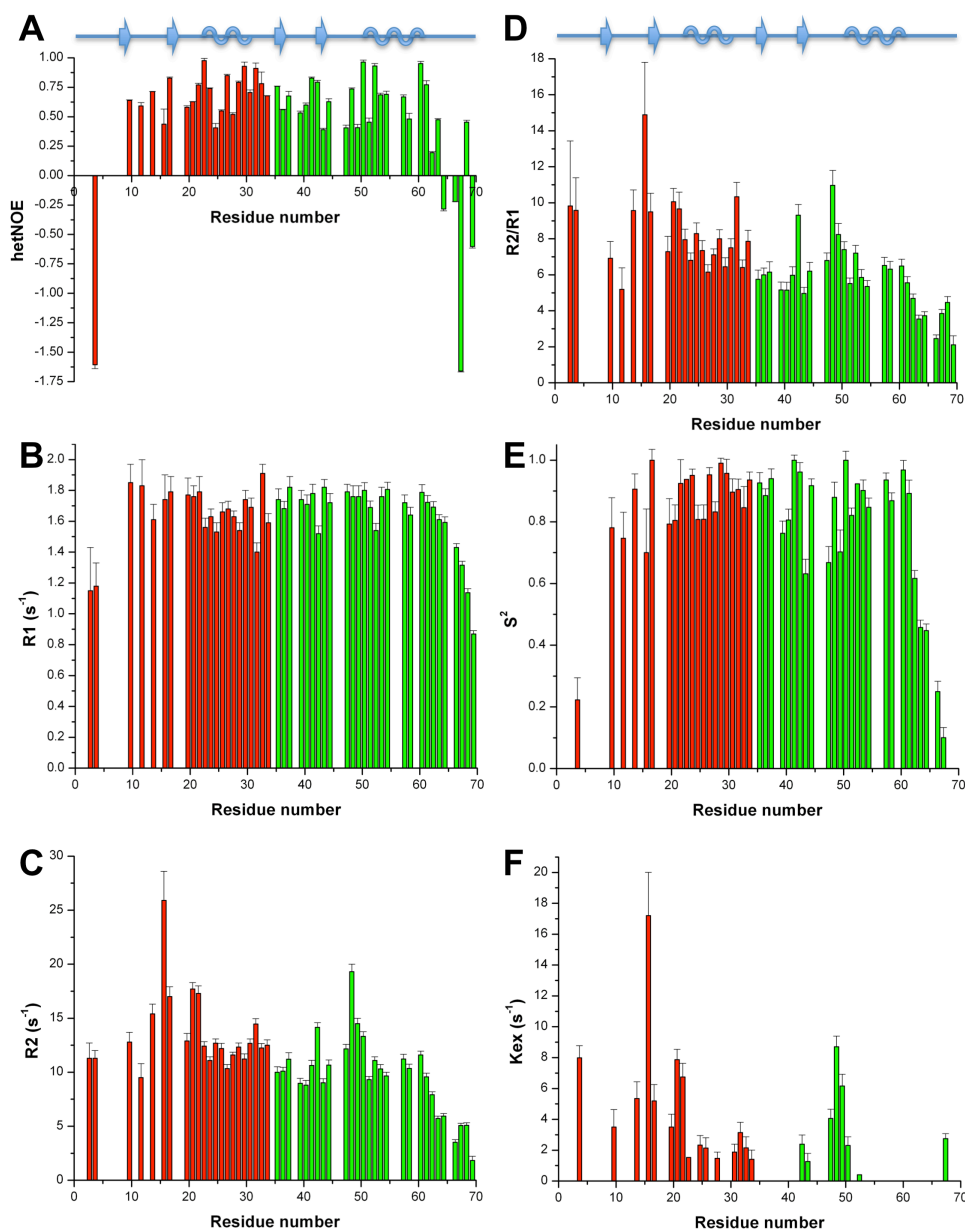


Figure V.7. Backbone dynamics of UBZ1+2 measured at 600 MHz and 25 °C. Values for residues in UBZ1 and UBZ2 are colored in red and green, respectively. **A.** ^1H - ^{15}N steady-state NOEs (hetNOE), **B.** ^{15}N R1, **C.** ^{15}N R2, **D.** R2/R1 ratios, **E.** Order parameter (S^2), and **F.** Chemical exchange term (K_{ex}). S^2 and K_{ex} were obtained from the TENSOR2 analysis using the model-free approach to describe the rotational diffusion tumbling of UBZ1+2 by an axially symmetric tensor with an anisotropy of $D_{par}/D_{perp} = 1.36 \pm 0.1$. Secondary structure elements are drawn above **A** and **D**.

Analysis of the relaxation data of UBZ1+2 using TENSOR2 (Dosset P. et al., 2000) indicates that in the most ordered regions, an average ^{15}N R2/R1 ratio (Fig. V.7D) is 6.84, corresponding to an apparent rotational correlation time τ_c of 7.73 ns. This

correlation time is higher than the value of 5.15 ns estimated for a globular protein of the same molecular weight (8.3 kDa) using the "empirical formula" based on a best fit of data from the NESG website, (www.nmr2.buffalo.edu/nescg.wiki), most likely due to the anisotropy of the system, but also because the flexible termini of the tandem can significantly enhance the τ_c (Tjandra N. et al., 1995). TENSOR2 analysis shows that UBZ1+2 has a rotational anisotropy of $D_{\text{par}}/D_{\text{perp}} = 1.36 \pm 0.1$, and the rotational diffusion tensor exhibits a prolate axial symmetry [$D_{\text{par}}=D_{\text{xx}}=D_{\text{yy}}=(1.93 \pm 0.09) \cdot 10^7 \text{ s}^{-1}$ and $D_{\text{perp}}=D_{\text{zz}}=(2.65 \pm 0.12) \cdot 10^7 \text{ s}^{-1}$], with the α -helix of the UBZ1 domain being parallel to the unique axis of the diffusion tensor (Fig. V.8). This alignment can explain the somewhat higher R2 values measured for the residues in the α -helix of UBZ1 compare to those in UBZ2 (Fig. V.7C) (Blake-Hall J. et al., 2004).

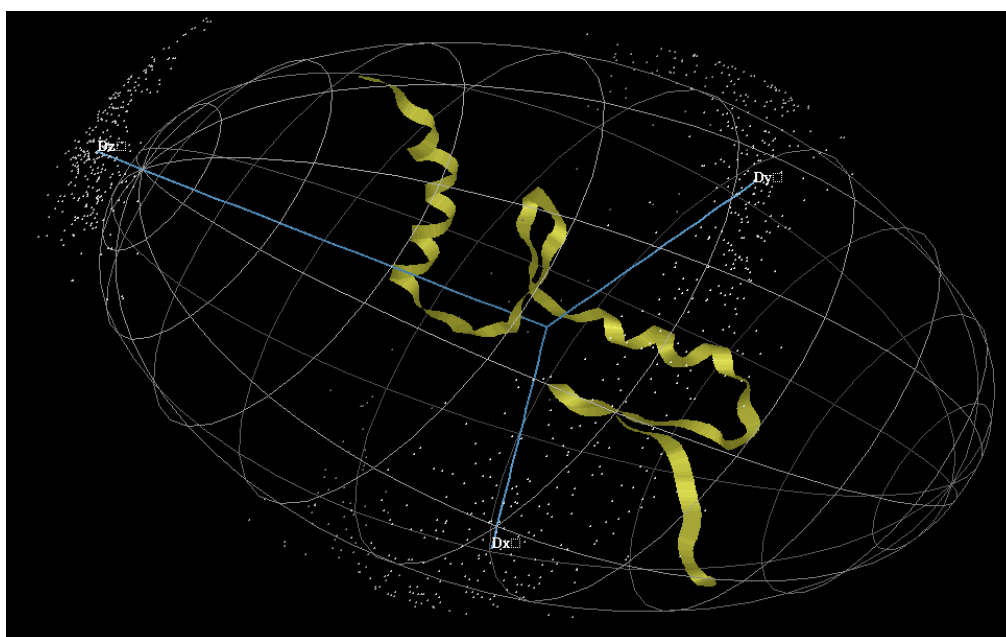


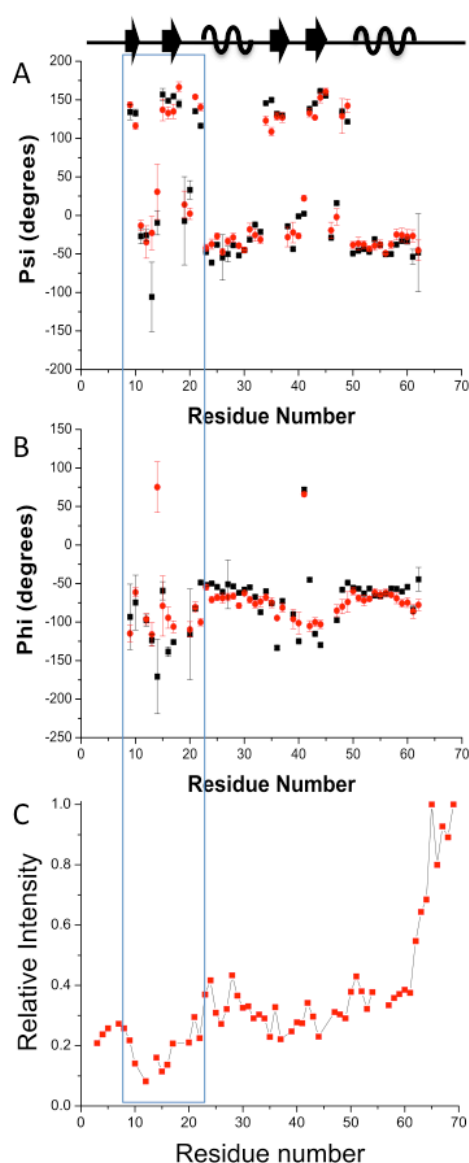
Figure V.8. Representation of the rotational diffusion tensor of UBZ1+2, exhibiting a prolate axial symmetry $D_{\text{par}}/D_{\text{perp}}=1.36 \pm 0.12$. The α -helix of UBZ1 (residues from 23 to 33) is parallel to the longitudinal axis (D_{zz}) of the diffusion tensor. White dots represent the dispersion in the directions of the diffusion tensor axes, due to experimental uncertainty as estimated from Monte Carlo simulations in TENSOR2.

Incorporation of the tensorial description into the model-free analysis (Lipari G. and Szabo A., 1982) of the local internal mobility has allowed us to describe the dynamics of UBZ1+2. For most residues, the measured relaxation data can be adequately described by a single parameter, the order parameter S^2 (Fig. V.7E), which is a measure of the motional amplitude of the N-H vector on the pico to nanosecond timescales. The S^2 profile (Fig. V.7E) shows the increase in internal mobility in the N- and C-terminal part of the protein, as can be seen from a large decrease in S^2 . The residues located in the loop connecting the β -sheet and the α -helix of UBZ2 present a small decrease in the S^2 values compared to the rest of the protein, which points to a slight increase in flexibility in this region of UBZ1+2. The local internal mobility of the residues showing conformational exchange (Fig. V.7F) can only be described with the TENSOR2 model using an additional term for exchange contributions (K_{ex}).

5.3.1.3 Comparison of the NMR and X-ray structures of the UBZ1+2 tandem

NMR spectroscopy and X-ray crystallography are complementary techniques widely used for the determination of protein structures at atomic level (Snyder D.A. et al., 2005; Yee A.A. et al., 2005). In general, independent structure determination by NMR and X-ray is useful to validate protein structures, and they can be combined to obtain a more accurate model (Shaanan B et al., 1992; Schieffer C.A. et al., 1994; Hoffman D.W. et al., 1996).

The crystallographic asymmetric unit of UBZ1+2 contains 11 symmetry-related protein molecules (Spinola M., Ph.D. thesis, 2010). Closer inspection of these reveals structural heterogeneity in the β -sheet region within the UBZ1 domain. Indeed, substantial variations of the torsion angles around the N-C $^{\alpha}$ (Phi) and C $^{\alpha}$ -CO (Psi) bonds are observed in this region by crystallographic analysis (Fig. V.9A,B), and the Phi and Psi



angles appear to be different from those in the NMR ensemble. Interestingly, a lower signal intensity of the UBZ1 residues (corresponding to those showing high backbone variation in the X-ray conformational structure) is measured in the ^1H - ^{15}N HSQC spectrum (Fig. V.9C). Such intensity decrease is indicative of extensive conformational exchange, which is in agreement with the structural heterogeneity observed in X-ray and NMR structures. Moreover, the conformational exchange within this UBZ1 is confirmed by the elevated R2 values (Fig. V.7C) and the increased K_{ex} values as seen in the NMR relaxation analysis (Fig. V.7F).

Figure V.9. Comparison of the torsion angles around the C $^{\alpha}$ -CO bond (A, Psi) and the N-C $^{\alpha}$ bond (B, Phi) for the UBZ1+2 residues 8 - 65 calculated from the X-ray (black) and NMR (red) structures. The error bars show the standard deviations calculated over 11 individual molecules in the asymmetric unit (X-ray) or 20 lowest-energy structures (NMR). C. The relative intensity of NMR signals of UBZ1+2 residues 3 - 69. The blue box highlights the differences in the Phi and Psi values of NMR and X-ray structures at the N-terminal part of the UBZ1 domain. Secondary structure elements are drawn above the plots.

Both X-ray and NMR structures of UBZ1+2 are of high quality, and the structures of the individual domains within the tandem are very similar, with backbone r.m.s.d. of 1.14 Å for UBZ1 (residues 8 - 34) and 0.91 Å for UBZ2 (residues 35 - 62). However, superposition of the full UBZ1+2 (residues 8 - 62) results in a 2.30 Å backbone r.m.s.d. between the NMR and X-ray structures. The substantial increase in r.m.s.d. is related to the changes in relative orientations of the two UBZ domains as shown in figure V.10. The hinge region involves the loop connecting the β -strand with the α -helix in UBZ2. Also the C-terminal end of the helix in UBZ2 is bent in the X-ray structure, likely as a result of the tight packing. The quantitative analysis of the relaxation parameters confirms the flexibility of the loop region in the UBZ1+2, where smaller hetNOE and S^2 values are observed in comparison with the rest of UBZ1+2 (Fig. V.7). The interdomain flexibility only became apparent when comparing the X-Ray and NMR structures, highlighting the strength of both techniques as complementary methods for the full structural description of UBZ1+2.

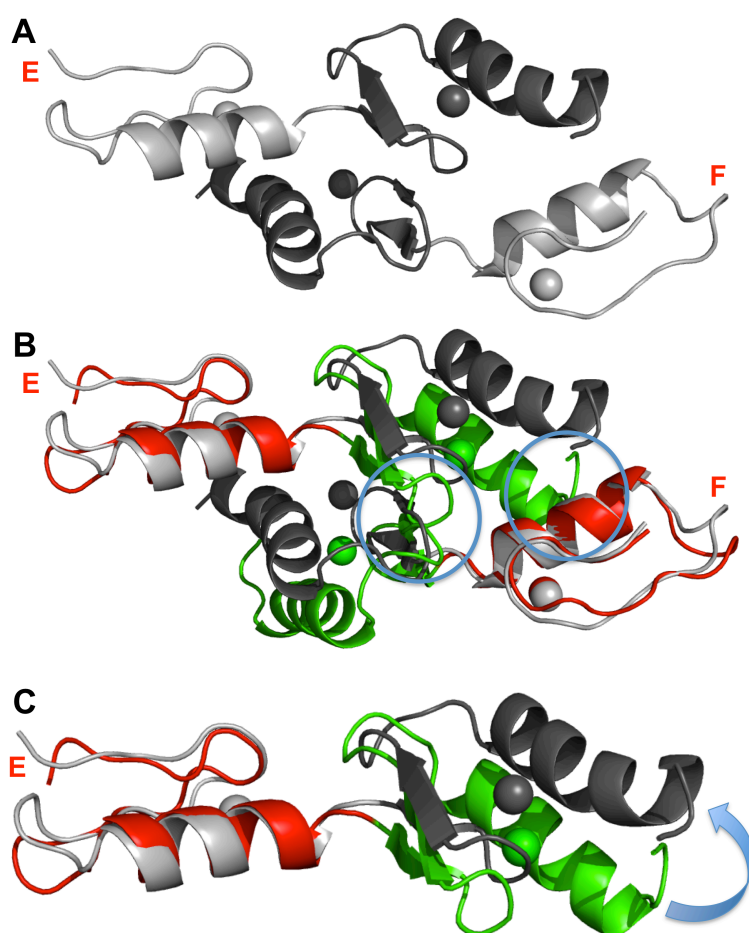


Figure V.10. Structural differences between the solution and crystal structures of UBZ1+2. *A.* Cartoon representation of X-ray molecules E and F (red letters are placed near the N-terminus of each molecule) in the asymmetric unit (Spinola M., Ph.D. thesis, 2010). UBZ1 and UBZ2 are shown in light grey and dark grey, respectively. *B.* Superposition of the lowest-energy NMR structure on the UBZ1 domain of the molecules E and F. UBZ1 and UBZ2 are shown in red and green, respectively. Regions of steric clashes are circled. *C.* Relative displacement of the two UBZ domains most likely due to the crystal packing is indicated by the arrow. For clarity only one X-ray (molecule E) and NMR structure are shown.

5.3.1.4 Coordination of the Zn^{2+} atom and the hydrophobic core of the UBZ domains

Coordinated by the two Cys (in the β -sheets) and the two His (in the α -helix), the zinc atoms are crucial for maintenance of the 3D structure of the UBZ domains. As is shown in figure V.11, in the absence of zinc, the 1D and 1H - ^{15}N -HSQC spectra of UBZ1+2 (in green) show very limited signal dispersion, with chemical shifts close to random coil values, indicating the presence of unstructured elements in the protein (Mizuguchi M. et al., 2003; Dyson H.J. and Wright P.E., 2004). In other words, the zinc fingers are not properly folded. Addition of an equimolar amount of $ZnCl_2$ (in the presence of TCEP to ensure the reduced state of the cysteines involved in the zinc coordination) induces drastic changes in the 1D and 1H - ^{15}N -HSQC spectra (Fig. V.11, pink). The spectral dispersion is greatly increased, and most of the individual peaks can be distinguished, which indicates a well-folded, Zn^{2+} -bound UBZ1+2 structure.

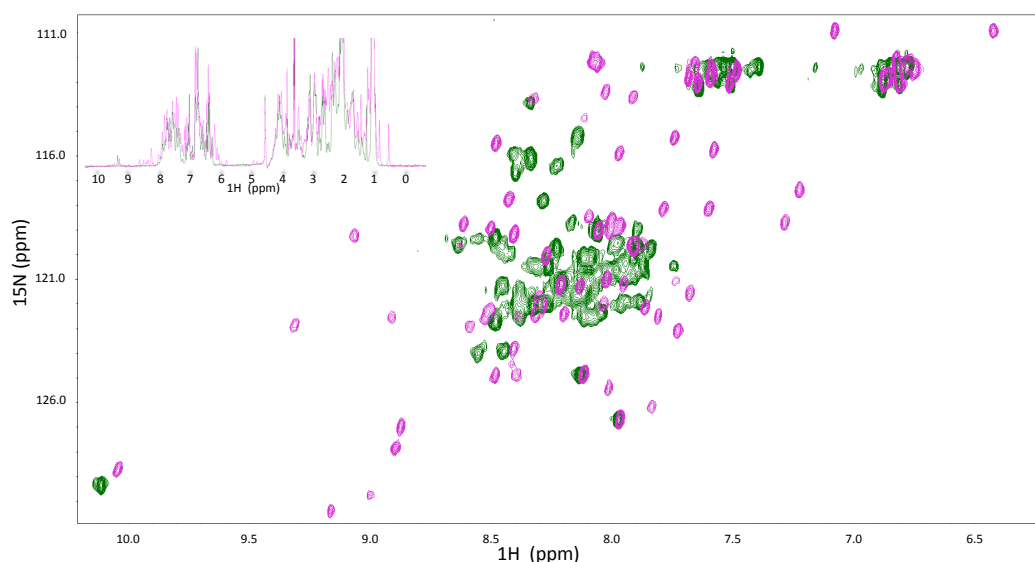


Figure V.11. Comparison of the 1H - ^{15}N -HSQC and the 1D 1H spectra of the UBZ1+2 domains in presence (pink) and absence of Zn^{2+} (green). The differences in the dispersion patterns are indicative of folded (pink) or unfolded (green) protein, and clearly highlight the importance of the Zn^{2+} atoms to maintain the 3D structure of the zinc finger domains.

These findings have been confirmed by CD spectroscopy. Figure V.12 shows how the CD spectrum of UBZ1+2 changes upon addition of increasing amounts of Zn^{2+} , up till 2 molar equivalents. After that, the spectrum does not change even if we continue adding more Zn^{2+} (Fig. V.12C, purple spectrum, ratio 1:2 and dark blue spectrum, ratio 1:2.5). The addition of Zn^{2+} results in an increase of the secondary structure content, as demonstrated by a shift of the minimum to > 200 nm, the appearance of a positive maximum at 195 nm, and the classical double-minimum in α -helix CD spectra. In contrast, the unfolded domains (Fig. V.12, in cyan) present a very low ellipticity above 210 nm and a negative band around 200 nm, characteristic of predominantly unfolded or intrinsically disordered proteins (Uversky V.N. and Longhi S., 2010; Nickerson N.N. et al., 2011). This pattern has also been observed in other zinc finger domains in presence and absence of Zn^{2+} (McColl D.J. et al., 1999; Matthews J.M. et al., 2000; Woodruff

R.V. et al., 2010). Thus, it is clear that Zn^{2+} induces secondary structure formation in UBZ1+2.

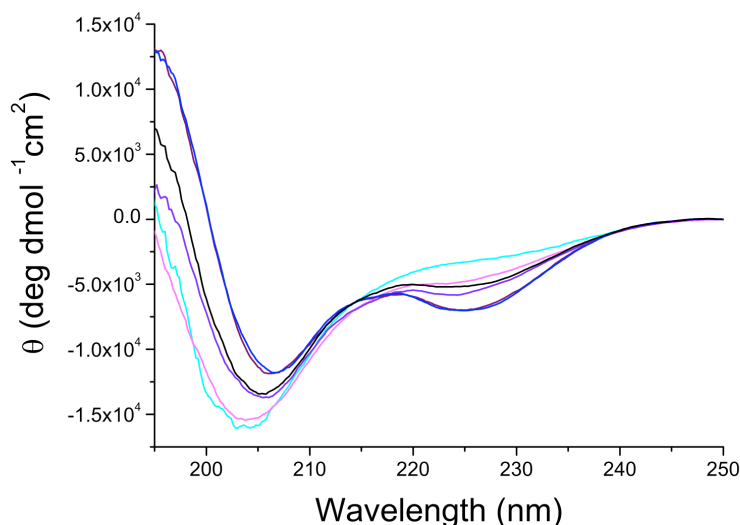


Figure V.12. Far-UV CD spectra (250 - 195 nm) showing the folding process of UBZ1+2 upon addition of increasing amounts of Zn^{2+} . The curves are colored by the UBZ1+2: Zn^{2+} ratios as cyan (ratio 1:0), magenta (ratio 1:0.5), violet (ratio 1:1), purple (ratio 1:2) and dark blue (ratio 1:2.5). The black line represents the simulation of the 1:1 ratio (violet) based on the combination of the CD spectra at ratios 1:0 and 1:2.5. Samples were prepared in 50 mM NaPi at pH 6.0 and 1 mM TCEP. Measurements were performed at 25 °C on a Jasco J-715 spectropolarimeter. For comparison, the results are normalized and expressed in molar ellipticity ($\text{deg}\cdot\text{dmol}^{-1}\cdot\text{cm}^2$) per residue (see Chapter III, section 3.5).

Interestingly, the packing of the β -sheet and the α -helix against each other forms a hydrophobic core, which places the conserved Cys and His residues toward the interior of the domain, in a position to coordinate the Zn^{2+} atom. Several residues (Leu12, Leu15, Phe17, Tyr21, Phe26, and Val30 for UBZ1; and Met39, Phe44, Tyr48, Phe53 and Val57 for the UBZ2 domain) serve to shield the Zn^{2+} ion from solvent and stabilize the metal interaction by precluding close approach of the alternative donor ligands (Fig. V.12). This hydrophobic shell surrounding the metal binding sites is very common in zinc finger domains and plays a crucial role in minimizing the change in conformational entropy upon metal binding (Narayan V.A. et al., 1997).

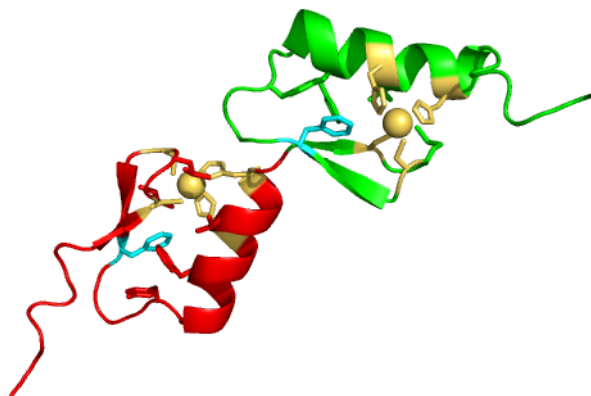


Figure V.13. Cartoon representation of the lowest energy NMR structure of UBZ1+2. Side chains of residues that form the hydrophobic core are shown as sticks (Leu12, Leu15, Phe17, Tyr21, Phe26, and Val30 for UBZ1 in red; and Met39, Phe44, Tyr48, Phe53 and Val57 for the UBZ2 domain in green). The conserved Cys and His coordinating the Zn^{2+} atoms (spheres) are

highlighted in yellow sticks. Residues Phe17 and Phe44, acting as aromatic anchors in the zinc coordination sites, are in cyan.

This hydrophobic core is further stabilized by the interaction of the aromatic side chains of the residues Phe17 with His29 (UBZ1) and Phe44 with His56 (UBZ2) (Fig.V.13). Such aromatic anchors are very common in classical zinc fingers, being involved in the edge-to-face packing to the zinc-chelating His. When the aromatic residue is located three residues after the second Cys that coordinates the Zn^{2+} (C-X₂-C-X₃-F-X₅-Φ-X₂-H), it is called “consensus”, as is the case in most of the zinc finger domains; for example, in USH-F9 (Cys₂HisCys, PDB code 1FU9) (Liew C.K. et al., 2000) and FOG-F3 (Cys₂His₂, PDB code 1SRK) (Simpson R J et al, 2004). However, in some unusual cases (less than 10%), the aromatic residue appears two residues after the second Cys that coordinate the zinc atom (C-X₂-C-X₂-Y-) and is known as “aromatic swap”. This is the case in NEMO ZF (Cys₂HisCys, PDB code 2JVX) (Cordier F. et al., 2008) and in the zinc fingers of the human Y-encoded ZFY protein, ZFY6 (Cys₂His₂, PDB code 5ZNF). The UBZ1 and UBZ2 domains of TAX1BP1 present the aromatic anchor in the “consensus” position (Fig. V.14).

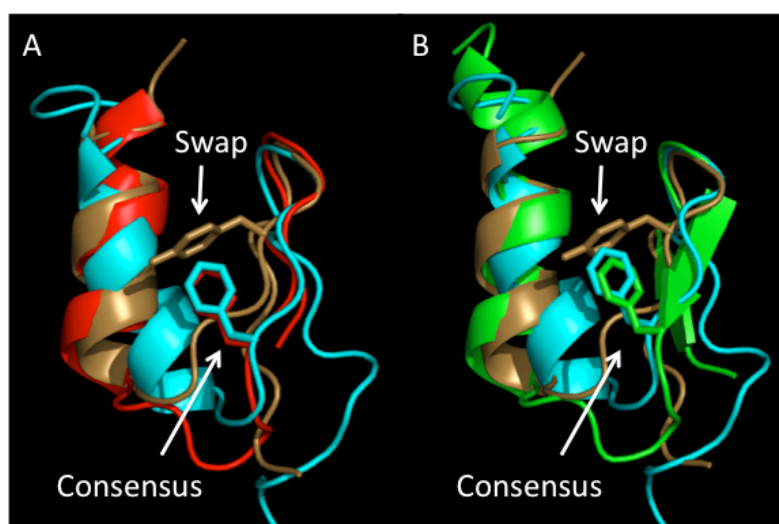


Figure V.14. Superposition of UBZ1 (A) domain (red) and UBZ2 (B) domain (green), of TAX1BP1 with the zinc finger FOG-F3 (in cyan, Cys₂His₂, PDB code 1SRK), which presents the aromatic residue in the “consensus” position, and with NEMO ZF (brown, Cys₂HisCys, PDB code 2JVX) which presents the aromatic residue in the unusual position known as “aromatic swap”. For both UBZ domains of TAX1BP1 the aromatic residues (Phe) are occupying the consensus position.

The high-resolution structure of UBZ1+2 is setting the basis for the study of the interaction with its principal target, ubiquitin. The mapping of the binding interfaces of both proteins would be crucial to obtain for the first time an accurate model of the UBZ1+2-ubiquitin complex. The study of the interaction between both proteins should provide insight into the regulation of TAX1BP1 signaling in different ubiquitin-related pathways.

5.3.2 Thermodynamic characterization of the interaction between the UBZ1+2 and ubiquitin

Isolated UBZ1 and UBZ2 domains, and the UBZ1+2 tandem were titrated with ubiquitin using ITC to quantify and compare the binding affinities and the thermodynamic parameters driving the interactions.

As shown in figures V.15B and C, the titrations of UBZ2 and UBZ1+2 with ubiquitin achieved a high degree of saturation, allowing accurate analysis. For the interaction of isolated UBZ1 with ubiquitin, the isotherm was of sufficient quality to be fitted with the model of n identical and independent binding sites (Fig. V.15A). The obtained K_d s and thermodynamic parameters are collected in Table V.3. For the interaction with UBZ1, the K_d value was close to 300 μ M, which is quite high for an accurate ITC analysis. The K_d s obtained for both UBZ2 and UBZ1+2 (Table V.3) are smaller (in the range of 30 - 40 μ M). However, the shapes of the isotherms are different, being more sigmoidal in the case of UBZ1+2. This reflects that the interaction of UBZ1+2 with ubiquitin involves not only UBZ2, but also the UBZ1 domain.

It is noteworthy that the stoichiometry of UBZ1-ubiquitin binding is not a clear 1:1 ($n = 0.7 \pm 0.2$). However, due to the poorly defined shape of the isotherm, it is not possible to establish based on the ITC titration alone if the n value indicates more than one binding event in the UBZ1 domain.

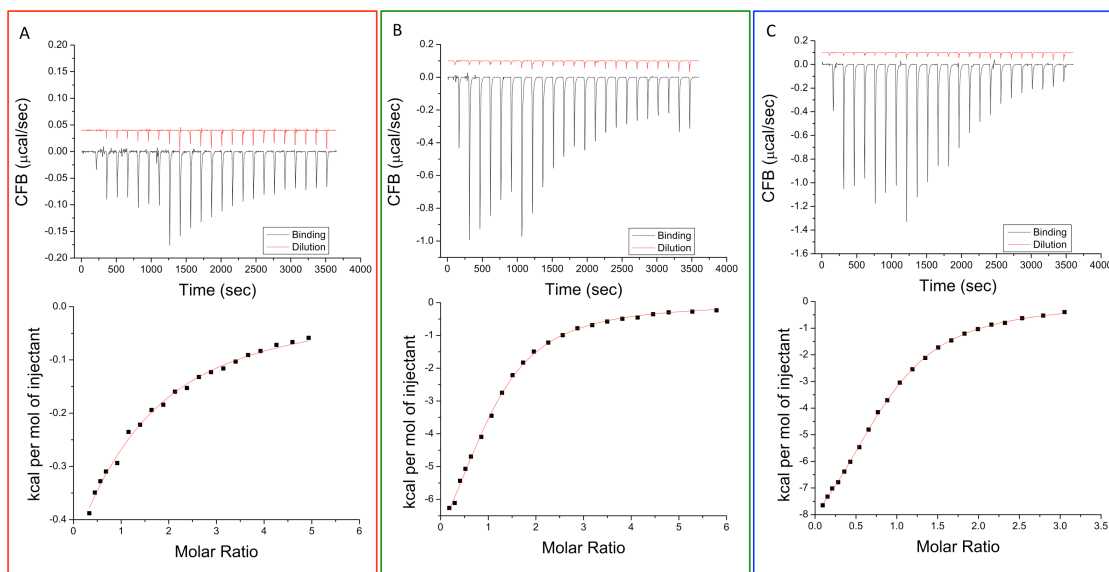


Figure V.15. ITC experiments of the isolated UBZ1 domain (130 μ M) (A), UBZ2 domain (146 μ M) (B) and UBZ1+2 (146 μ M) (C) titrated with ubiquitin (3.2 mM for the titration of UBZ1 and 2 mM for the titration of UBZ2 and UBZ1+2) in 50 mM cacodylate buffer pH 6.0 and 1mM TCEP, on an VP-ITC200 at 25 $^{\circ}$ C. In the upper panels are represented the thermograms for the binding (in black) and the dilution (in red) experiments, as heat per unit of time released after each injection. The low panels show the binding isotherms. Black squares represent experimental data and the red line corresponds to the best fit to a model considering n identical and independent sites.

PROTEINS	n	$K_d(\mu M)$	ΔH_b (kcal/ mol)	ΔG_b (kcal/ mol)	$-T \cdot \Delta S_b$ (kcal/mol)
Ubq-UBZ1	0.7 ± 0.2	294.5 ± 20.5	-1.7 ± 0.6	-4.8 ± 0.3	-3.1 ± 0.6
Ubq-UBZ2	1.0 ± 0.1	37.1 ± 1.8	-10.1 ± 0.3	-6.0 ± 0.2	4.1 ± 0.3
Ubq-UBZ1+2	0.9 ± 0.1	43.0 ± 2.3	-10.7 ± 0.4	-5.8 ± 0.2	4.9 ± 0.4

Table V.3. ITC parameters extracted from the fitting of the binding isotherms of the isolated UBZ1 and UBZ2 domains and the tandem UBZ1+2 tandem with ubiquitin (Ubq). The reported values are averaged over two experiments and the errors are calculated as the standard deviation from the mean values.

The interactions of UBZ2 and UBZ1+2 with ubiquitin are both exothermic processes, enthalpically driven (with negative enthalpic values and positive entropic terms) as was observed for the interaction of the SH3A domains of CIN85/CD2AP adaptor proteins with atypical proline-rich peptides (Chapter IV, section 4.3.1). This enthalpy-entropy compensation is common in macromolecular binding events, because the increase in the number of intermolecular interactions between the partners (enthalpic term) is offset by the decrease in degrees of freedom, resulting in an entropic penalty ($-T\Delta S > 0$).

However, the binding of the isolated UBZ1 domain to ubiquitin shows a negative enthalpy value of one order of magnitude smaller than the one obtained for the interaction of the UBZ2 domain, and a negative entropic term (larger than the enthalpic one) indicative of an entropically driven process. Entropically driven ($-T\Delta S < 0$) binding events indicate that hydrophobic interactions play an important role in the association due to an increase in solvent entropy from burial of hydrophobic groups and release of water upon binding, but also indicate a minimal loss of conformational degrees of freedom which is typical of dynamic binding events (Bronowska A.K., 2011). Briefly, weaker interactions (like is the case of UBZ1 and ubiquitin) are more entropically favourable, since the weak binding constricts less the molecular motions (Bronowska A.K., 2011).

The Gibbs energy term (ΔG_b) indicates if the binding process is thermodynamically favored or not and defines the strength of the interaction. Its partition between enthalpy and entropy is very important as these two terms reflect energetic contributions of totally different nature (Velazquez-Campoy A. and Freire E., 2005). If we compare the ΔG_b from the interaction of the isolated UBZ1 and UBZ2 with ubiquitin, we can see that both present similar values (5 - 6 kcal/mol). However, they have very different partition between enthalpy and entropy, which correspond to different binding modes driving the intermolecular interactions: enthalpically driven for the isolated UBZ2 domain and entropically driven for the isolated UBZ1 domain.

The titration of UBZ1+2 with ubiquitin was also fitted using a two non-symmetric sites model taking into account possible cooperativity effects (Chapter III, section 3.2.2). This fitting gave two distinct dissociation constants (9.7 μM and 106.1 μM) and a Hill coefficient α of 1.3, indicating a positive cooperativity.

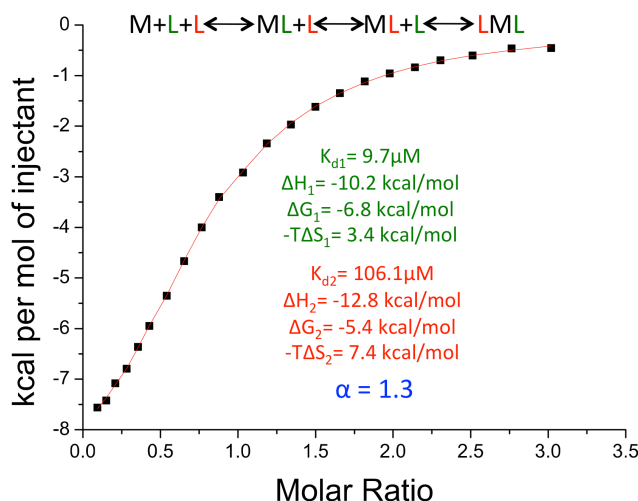


Figure V.16. Fitting of the UBZ1+2-ubiquitin ITC isotherm using the program SEDPHAT. Black squares represent experimental data, and the red line shows to the best fit to a model of two non-symmetric sites (indicated above the isotherm). The thermodynamic parameters of the binding sites with stronger and weaker affinities (corresponding to the UBZ2 and UBZ1 domains, respectively) are in green and red, respectively. In blue, the Hill coefficient obtained for the fitting, $\alpha > 1$, indicative of positive cooperativity.

In agreement with the ITC titrations of the isolated UBZ1 and UBZ2 domains (Table V.3), the present analysis indicates the existence of two ubiquitin-binding sites in the tandem UBZ1+2. The tighter ubiquitin binding to the UBZ domains in UBZ1+2 compared to that of the isolated UBZ1 and UBZ2 supports the existence of a positive cooperativity within the UBZ1+2 tandem as reflected in the Hill coefficient $\alpha > 1$. Analysis of the thermodynamic parameters obtained from the fitting in SEDPHAT shows that, when in tandem, UBZ1 and UBZ2 domains bind to ubiquitin in an enthalpically driven process. Interestingly, the enthalpic contribution of the weaker binding site (Fig. V.16, in red) is ten times bigger than that observed for the titration of the isolated UBZ1 domain. Reflecting the rise in the strength of the intermolecular interactions, this enthalpy increase is in agreement with the affinity enhancement, indicating that when UBZ1 is forming part of the tandem with UBZ2, its interaction with ubiquitin is more favorable, probably promoted by the stronger ubiquitin binding to the UBZ2 domain.

5.3.3 Mapping the interaction of UBZ1+2 with ubiquitin by NMR

NMR was used to map the binding interfaces involved in the interaction of UBZ1+2 with ubiquitin, from the point of view of both partners. The exquisite sensitivity of NMR to environmental changes of the residues was especially useful for the study of the low affinity UBZ1 binding event.

Series of ^1H - ^{15}N HSQC spectra were recorded for the ^{15}N -labeled ubiquitin sample during the titration with the isolated UBZ1 and UBZ2 domains as well as the UBZ1+2 tandem. The reverse experiment was also carried out by collecting a series of ^1H - ^{15}N -HSQC spectra for the ^{15}N - ^{13}C -labeled UBZ1+2 sample upon addition of increasing amounts of unlabeled ubiquitin.

5.3.3.1 Titrations of ^{15}N -labeled ubiquitin with isolated UBZ1 and UBZ2 domains and the UBZ1+2 tandem

In first place, ^{15}N -labeled ubiquitin was titrated with unlabeled UBZ1+2. The overlay of the ^1H - ^{15}N -HSQC spectra recorded at each titration ratio (Fig. V.17) shows the chemical shift perturbations for some of the ubiquitin resonances caused by the binding of UBZ1+2. Most of the binding shifts are in fast exchange on the NMR chemical shifts time scale (e.g. Gln49, Fig. V.17) and can be used to estimate the K_{d} s.

However, the resonances of some residues (e.g. Val70, Fig. V.17A) are in intermediate exchange. This is reflected by the progressive changes in the peak positions and the concomitant line-broadening, which initially increases and then decreases towards the end of the titration, when the saturation is reached. Three residues (Arg42, Leu43 and Leu71) broaden so much that their signals disappear and do not reappear at saturation, indicating a very large binding shift (Fig. V.20B, blue).

To determine the K_{d} s (Fig. V.17B), the titration curves of the residues in fast-to-intermediate exchange were fitted using a cooperative binding model (see section 3.3.3.3), which relates the chemical shift perturbation of each backbone amide to the ratio of protein concentrations at every titration step, and provides the Hill coefficient (α). The average K_{d} value ($12.6 \pm 3.4 \mu\text{M}$) is somewhat smaller than the obtained by ITC (Table V.3), but still in the same affinity range. The Hill coefficient $\alpha = 1$ indicates no cooperative binding effect within the ubiquitin molecule (Fig. V.17C).

The plot of the chemical shift perturbations vs the ubiquitin residue number (Fig. V.20) reveals that ubiquitin binds to UBZ1+2 through the hydrophobic patch, usually known as the canonical or Leu8-Ile44-Val70 motif. Reported in several related papers (Mueller T.D. et al., 2004; Hicke L. et al., 2005; Ohno A. et al., 2005; Wang Q. et al., 2005; Choi J.Y. et al., 2007), this canonical binding surface corresponds to the recognition interface for the ubiquitin-binding domains containing an α -helix (Hurley J.H. et al., 2006).

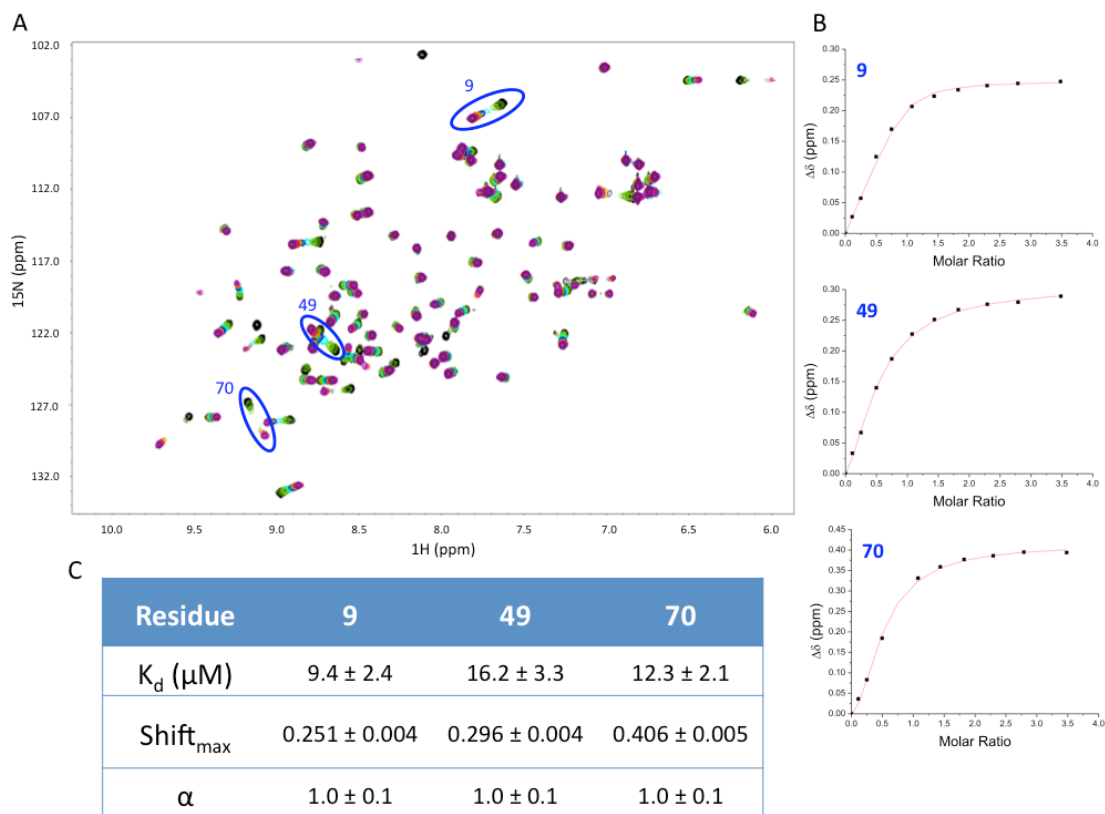


Figure V.17. *A.* Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the ^{15}N -labeled ubiquitin NMR titration with unlabeled UBZ1+2. Samples were prepared in 50 mM NaPi buffer at pH 6.0 and 1 mM TCEP and experiment were recorded on a Varian 800 MHz spectrophotometer at 25 °C. Most of the affected residues are in fast-to-intermediate exchange and show linear trajectories. Residues Thr9, Glu49 and Val70 are indicated by blue ovals (blue color code refers to UBZ1+2 throughout the text). *B.* Chemical shift changes of the highlighted residues vs the molar ratio. The dots correspond to the experimental data, and the red lines show the best fit to a model that takes into account possible cooperativity effects. *C.* The fitted parameters for titrations of the selected residues shown in (B). The average value of the calculated dissociation constant (K_d) is 12.6 ± 3.4 μM . $\text{Shift}_{\text{max}}$ is the maximum chemical shift value achieved upon saturation.

To determine if ubiquitin uses the same binding interface to recognize the individual domains, ^{15}N -labeled ubiquitin was titrated with increasing amounts of isolated UBZ1 and UBZ2, one at a time.

NMR titration of ubiquitin with the UBZ1 domain (Fig. V.18A) shows that basically the same ubiquitin residues are affected as in the UBZ1+2 experiment described above. Small differences are found in the average chemical shift perturbations of the residues Gly10 and Glu50, which are slightly less affected during the titration with UBZ1, and the residues Gln41, Arg74 and Gly75, which average $\Delta\delta$ values are a bit more affected by the interaction with the isolated UBZ1 than with UBZ1+2.

In general, the chemical shift perturbations of the ubiquitin resonances are smaller than the ones observed during the titration with UBZ1+2. All affected residues are in fast-exchange and contrary to the case of UBZ1+2, none of the peaks show considerable line-broadening. As shown by the fitting of some binding shifts, saturation is not

reached at the 1:3.0 UBZ1:ubiquitin ratio (Fig. V.18B). This is a clear indication of the weak interaction between UBZ1 and ubiquitin, with the average K_d value of $346.8 \pm 37.3 \mu\text{M}$, which is in good agreement with the value obtained by ITC. Being common in transient protein interactions (Worrall J.A. et al., 2002; Volkov A.N. et al., 2005), small chemical shift perturbations are indicative of dynamic complexes, which might also be the case for the entropically driven binding of the isolated UBZ1 and ubiquitin, as suggested by ITC. Again, $\alpha = 1.0$ reflects a non-cooperative effect within ubiquitin (Fig. V.18C).

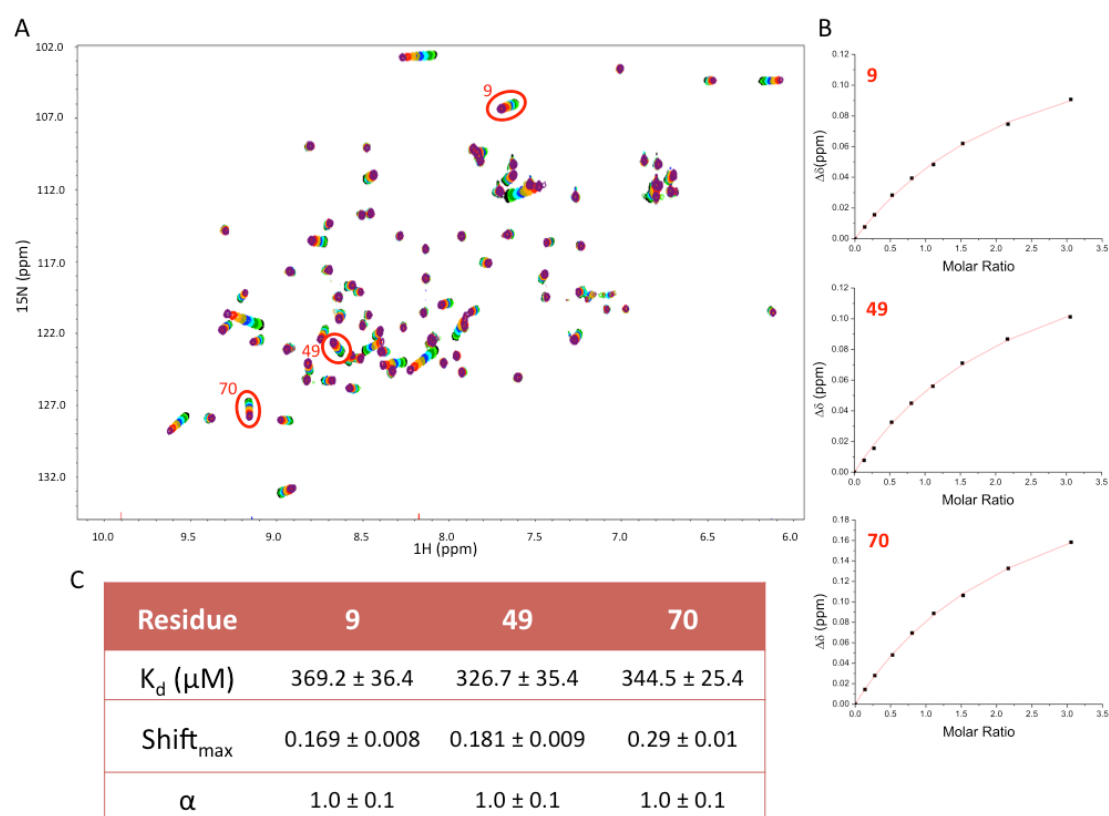


Figure V.18. *A.* Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the ^{15}N -labeled ubiquitin NMR titration with the unlabeled UBZ1. Samples were prepared in 50 mM NaPi buffer at pH 6.0 and 1 mM TCEP, and spectra were recorded on a Varian 800 MHz spectrophotometer at 25 °C. Most of the affected residues are in fast exchange and show linear trajectories. Residues Thr9, Glu49 and Val70 are indicated by red ovals (red color code refers to UBZ1 throughout the text). *B.* Chemical shift changes of the highlighted residues vs the molar ratio. The dots correspond to the experimental data and the red lines show the best fit to a model that takes into account possible cooperativity effects. *C.* The fitted parameters for titrations of the selected residues shown in (B). The average value of the calculated dissociation constant (K_d) is $346.8 \pm 37.3 \mu\text{M}$. $\text{Shift}_{\text{max}}$ is the maximum chemical shift values achieved upon saturation.

For the titration with isolated UBZ2, the same ubiquitin residues as in the previous NMR experiments are affected (Fig. V.19A), and the size of the chemical shift perturbations is similar to that observed for the interaction with UBZ1+2 (see above). As in the case of UBZ1+2 binding, the UBZ2-bound ubiquitin resonances are in fast-to-intermediate exchange regime. The average K_d value $36.5 \pm 5.5 \mu\text{M}$ obtained from this NMR titration agrees well with the values obtained by ITC (Table V.3).

The large binding shifts of UBZ2 in comparison with UBZ1 might be associated with the formation of a predominantly single-orientation complex (Prudêncio M. and Ubbink M., 2004) as expected for the 1:1 enthalpically driven interaction detected by ITC.

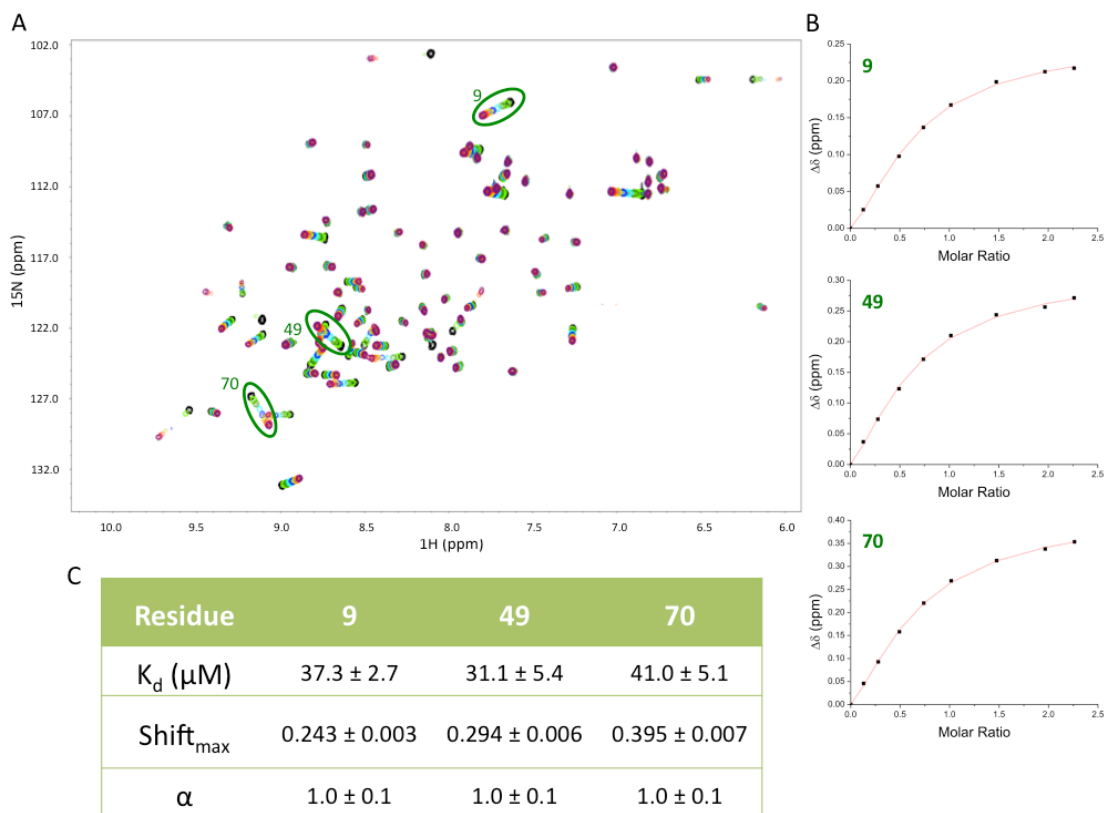


Figure V.19. *A.* Overlay of the ^1H - ^{15}N -HSQC spectra recorded during the ^{15}N -labeled ubiquitin NMR titration with unlabeled UBZ2. Samples were prepared in 50 mM NaPi buffer at pH 6.0 and 1 mM TCEP, and spectra were recorded on a Varian 800 MHz spectrophotometer at 25 °C. Most of the affected residues are in fast-to-intermediate exchange and show linear trajectories. Residues Thr9, Glu49 and Val70 are indicated by green ovals (green color code refers to UBZ2 throughout the text). *B.* Chemical shift changes of the highlighted residues vs the molar ratio. The dots correspond to the experimental data and the red lines show the fitting by a mathematical model that takes into account possible cooperativity effects. *C.* The fitted parameters for titrations of the selected residues shown in (B). The average value of the calculated dissociation constant (K_d) is $36.5 \pm 5.5 \mu\text{M}$. $\text{Shift}_{\text{max}}$ is the maximum chemical shift values achieved upon saturation.

The chemical shift perturbation analysis reveals that ubiquitin binds the isolated UBZ domains and UBZ1+2 tandem in a uniform manner (Fig. V.20), always through its canonical binding site (Leu8-Ile44-Val70) (See Fig. V.20A). Since ubiquitin interacts with the UBZ domains through a unique binding site, it is not possible to determine any kind of cooperativity effect within ubiquitin, which is in agreement with $\alpha = 1$ values.

In agreement with the ITC results, the biggest K_d value ($346.8 \pm 37.3 \mu\text{M}$) has been found for the ubiquitin interaction with the isolated UBZ1, while the isolated UBZ2 shows stronger binding with the K_d value of $36.5 \pm 5.5 \mu\text{M}$. The lower apparent K_d of $12.6 \pm 3.4 \mu\text{M}$ determined for the interaction with UBZ1+2 indicates that the tandem motif increases the binding affinities of ubiquitin with both the UBZ1 and UBZ2 domains.

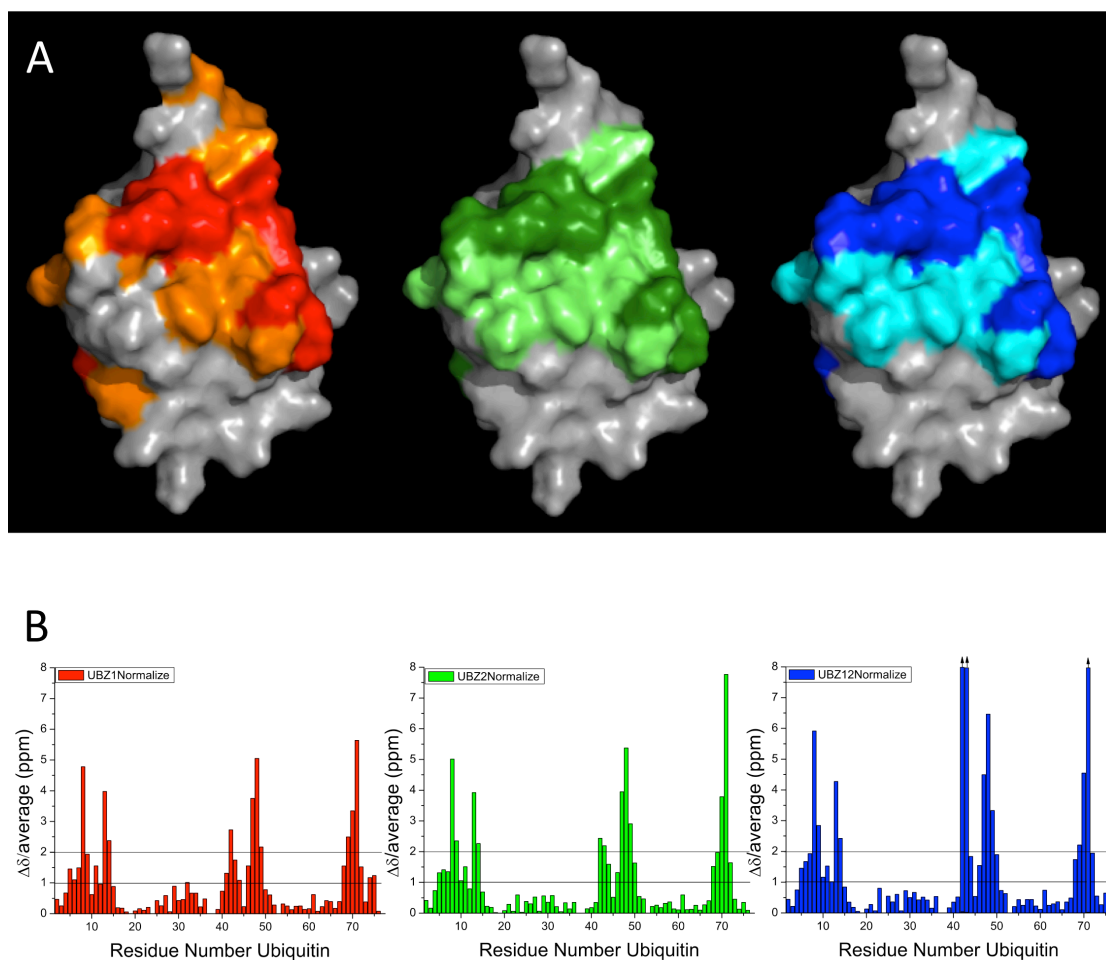


Figure V.20. *A.* Surface representation of the solution structure of ubiquitin (in grey, PDB code 1D3Z) indicating the residues affected during the titration with the isolated UBZ1 (in red) and UBZ2 (in green) domains, and the UBZ1+2 tandem (in blue). Light and dark colors correspond to the residues with $\Delta\delta/\text{average} = 1 - 2$ and $\Delta\delta/\text{average} > 2$, respectively. *B.* Normalized chemical shift perturbations ($\Delta\delta/\text{average}$) of the backbone $^1\text{H}-^{15}\text{N}$ resonances of ubiquitin vs residue number, at the end of the titration with the isolated UBZ1 (in red) and UBZ2 (in green) domains, and UBZ1+2 (in blue). The ubiquitin residues that disappeared during the titration with UBZ1+2 (Arg42, Leu43 and Leu71) are presented by blue bars with an arrow on top, indicative of a big shift.

5.3.3.2 Determining the ubiquitin-binding interface on UBZ1+2 by chemical shift mapping

The ^{15}N - ^{13}C -labeled UBZ1+2 tandem was titrated with increasing amounts of ubiquitin till a ratio 1:7.0 to ensure that both domains, also UBZ1, will be completely saturated.

The UBZ1+2 amide chemical shift changes during the titration with ubiquitin are shown in figure V.21. Interestingly, the binding shifts follow different patterns depending on the residues. Most of the residues affected are in fast exchange (e.g. Lys9, His33, Glu42, His60, Phe61), while others (e.g. Asn20, Tyr21, Glu28, Trp34, Met39, Phe44, Asp47, Gln58) are in intermediate exchange and experience a strong line-broadening at intermediate UBZ1+2:ubiquitin ratios, which decreases again when saturation is reached. HNCACB, CBCA(CO)NH and ^{15}N -NOESY- HSQC spectra have been used to assign the peaks corresponding to the bound form of some residues (Asn20, Glu28, Trp34 and Asp47), which trajectories could not be visually followed due to the severe line-broadening. For several residues (Cys10, Leu12, Glu14 and Lys35), the signals disappear during the first steps of the titration and do not reappear towards the end. Most of the chemical shift perturbation trajectories are linear, indicative of a single binding event (see Chapter III, section 3.3.3.3). However, some residues (Ser32, Trp34, Gln43, Asp47, Tyr48 and Asp49; Fig.V.21) show a clear curvature, related to two or more binding events (see Chapter III, section 3.3.3.3). It is interesting to note that the residues showing curvature follow three different patterns. The chemical shift perturbation of the residues Ser32 and Gln43 (located in the region between both UBZ domains, see Fig. V.24, yellow) are in fast-intermediate exchange and change the trajectory after ratio 1:0.51. Residues Tyr48 and Asp49 (located in the hinge loop that connect the β -sheet with the α -helix of UBZ2, see Fig. V.24, red) are also in fast-intermediate exchange, however the change of trajectory takes place after a ratio 1:2.40. Finally, residues Trp34 (in the region between both UBZ domains, Fig. V.24, black) and Asp47 (located in the hinge loop that connect the β -sheet with the α -helix of UBZ2, between two Pro and a Tyr48 that form part of the hydrophobic core, Fig. V.24, black) are in intermediate exchange disappearing after the ratio 1:0.36 and reappearing when saturation is reached; however, the directions of chemical shift changes in the initial and final titrations points indicate the existence of a strong curvature.

Backbone amide chemical shift perturbation can be interpreted qualitatively either as a consequence of direct ligand interactions or ligand-induced conformational changes within the protein (Zuiderweg E.R., 2002). In this sense, the presence of a curvature for some residues can arise from two or more binding events (Zuiderweg E.R., 2002) or be due to secondary binding effects. Considering that both UBZ1 and UBZ2 interact with ubiquitin, it is logic to assume that the curvature shown by the residues in the area between both domains (see Fig. V.24 yellow) reflects the binding of more than one ubiquitin molecule. However, for the residues located in hinge-regions and loops (Fig. V.24 red and black), the change of the chemical shift perturbation trajectories could be due to binding-induced conformational changes.

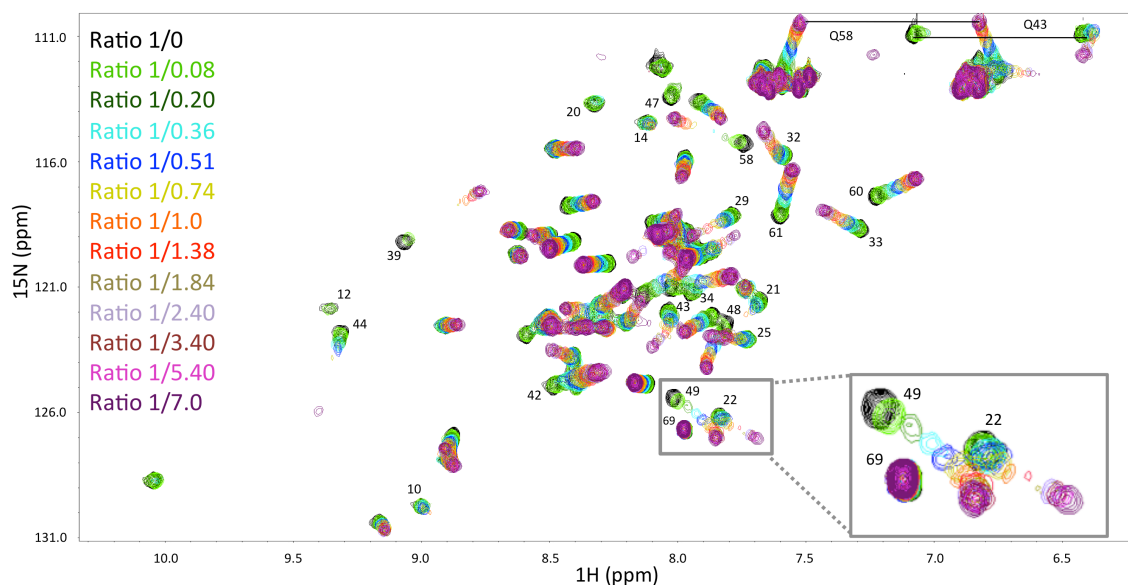


Figure V.21. ^1H - ^{15}N -HSQC spectra recorded during the NMR titration of ^{15}N - ^{13}C -labeled UBZ1+2 tandem with unlabeled ubiquitin. Samples were prepared in 50 mM NaPi buffer at pH 6.0 and 1 mM TCEP, and experiments were recorded on a Varian 800 MHz spectrophotometer at 25 °C. The color code corresponding to the different ratios of the titration is given on the left. The zoomed in region shows three representative patterns found during the titration: Asp69 (not involved in the interaction) experiences no binding shifts; Asp49 (implicated in two or more binding events or conformational changes) shows a curved trajectory; and Asp22 (implicated in the conformational exchange) follows a linear trajectory and shows signal broadening. Some of the representative residues in fast and intermediate exchange, together with those showing curvature, have been identified by the labels. Also the most affected NH_2 side chains (Gln43, showing strong curvature and Gln58, showing a linear pattern) have been indicated.

Using NMR chemical shift perturbation analysis to map protein-protein interaction at a residue level, we observed that residues located in both UBZ1 and UBZ2 domains are involved in binding to ubiquitin. UBZ1 residues whose backbone amides are more affected upon binding with ubiquitin are Cys10, Leu12, Cys13, Glu14, Asn20, Tyr21, Asp22, Lys25, Glu28, His29, Val30, Ser32, His33 and Trp34; and in the UBZ2 domain, residues Lys35, Met39, Ser41, Glu42, Gln43, Phe44, Asp47, Tyr48, Asp49, Val52, Phe53, Val 57, Gln58, His60, Phe61 and Asp62. Figure V.22 shows that NH_2 side chains of Gln43 and Gln58 are also affected during the titration. Interestingly the side chain NH_2 groups of both residues show the same binding patterns as their backbone amides: Gln43 exhibits curvature, while Gln58 follows linear chemical shift changes. The other Asn and Gln side chains are not affected, except for the NH_2 side chains of Gln23 and Gln50, which broaden out during the titration (they are indistinguishable due to strong overlap, Fig. V.22).

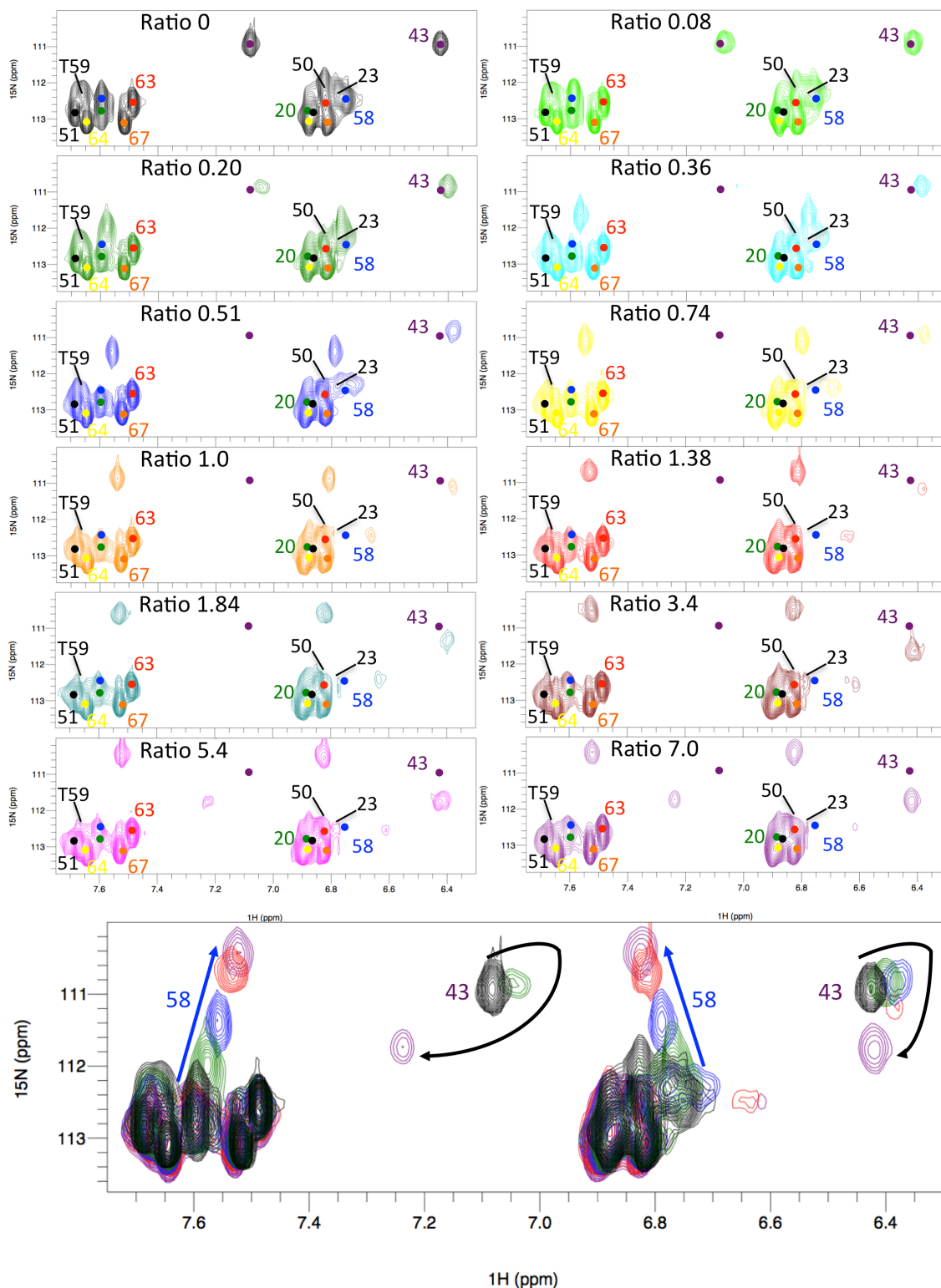


Figure V.22. The NH₂ side chain region of the ¹H-¹⁵N-HSQC spectra shown in figure V.21. The NH₂ side chains of especially Gln43 and Gln58 (and also Gln50 and/or Gln23) are strongly affected upon binding to ubiquitin. HδNδ and HεNε assignment of Asn and Gln residues, respectively, are indicated. At all titration points, the ¹H-¹⁵N peak position of each residue in free UBZ1+2 is indicated with a dot colored according to the residue number.

Since UBZ1+2 contains two ubiquitin-binding domains, the chemical shift perturbation of UBZ1+2 at different ubiquitin titration points were used to follow how the two UBZ domains are affected during the titration (Fig. V.23).

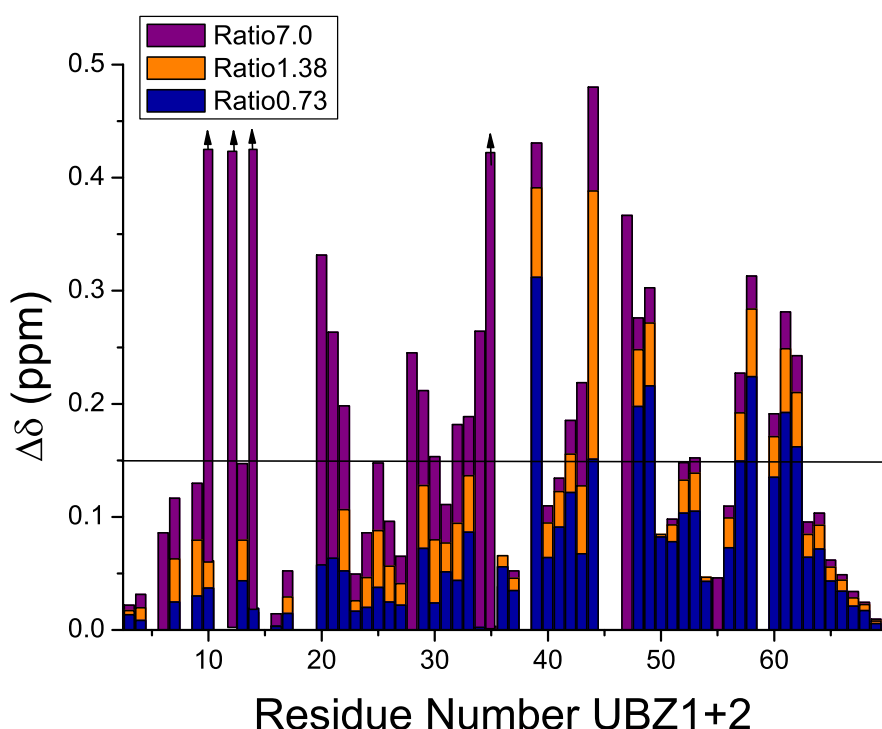


Figure V.23. Backbone amide chemical shift perturbations vs the UBZ1+2 residue number at three different ratios during the titration with ubiquitin (1:0.73 in dark blue; 1:1.38 in orange; 1:7.0 in purple). Residues of UBZ1+2 that disappear during the titration with ubiquitin are represented by purple bars with an arrow on the top. The missing data points correspond to residues Asp5, Lys8, Leu15 and Thr59 (whose chemical shift perturbations could not be analyzed due to spectral overlap or signal disappearing) and prolines 11, 18, 19, 38, 45 and 46. The black line represents the chemical shift cut-off based on the mean chemical shift perturbation ($\Delta\delta$) plus one standard deviation at ratio 1:7.0.

Mapping the chemical shift perturbations on the UBZ1+2 structure at different protein ratios reveals how the binding of ubiquitin is gradually affecting the tandem of both UBZ domains (Fig. V.24). It is clear that residues affected at the beginning of the titration are located in the UBZ2 domain. When UBZ2 domain is almost saturated (ratio 1:1.38), residues in the UBZ1 domain start to show a gradual increase in chemical shift perturbations upon further addition of ubiquitin. This pattern likely indicates sequential binding of ubiquitin molecules, where the first ubiquitin molecule binds to the high-affinity UBZ2 site, followed by the second, weaker binding event at the UBZ1 domain. This interaction scenario is in agreement with the K_d s determined for the individual UBZ1 (K_d of $346.8 \pm 37.3 \mu\text{M}$) and UBZ2 (K_d of $36.5 \pm 5.5 \mu\text{M}$) domains.

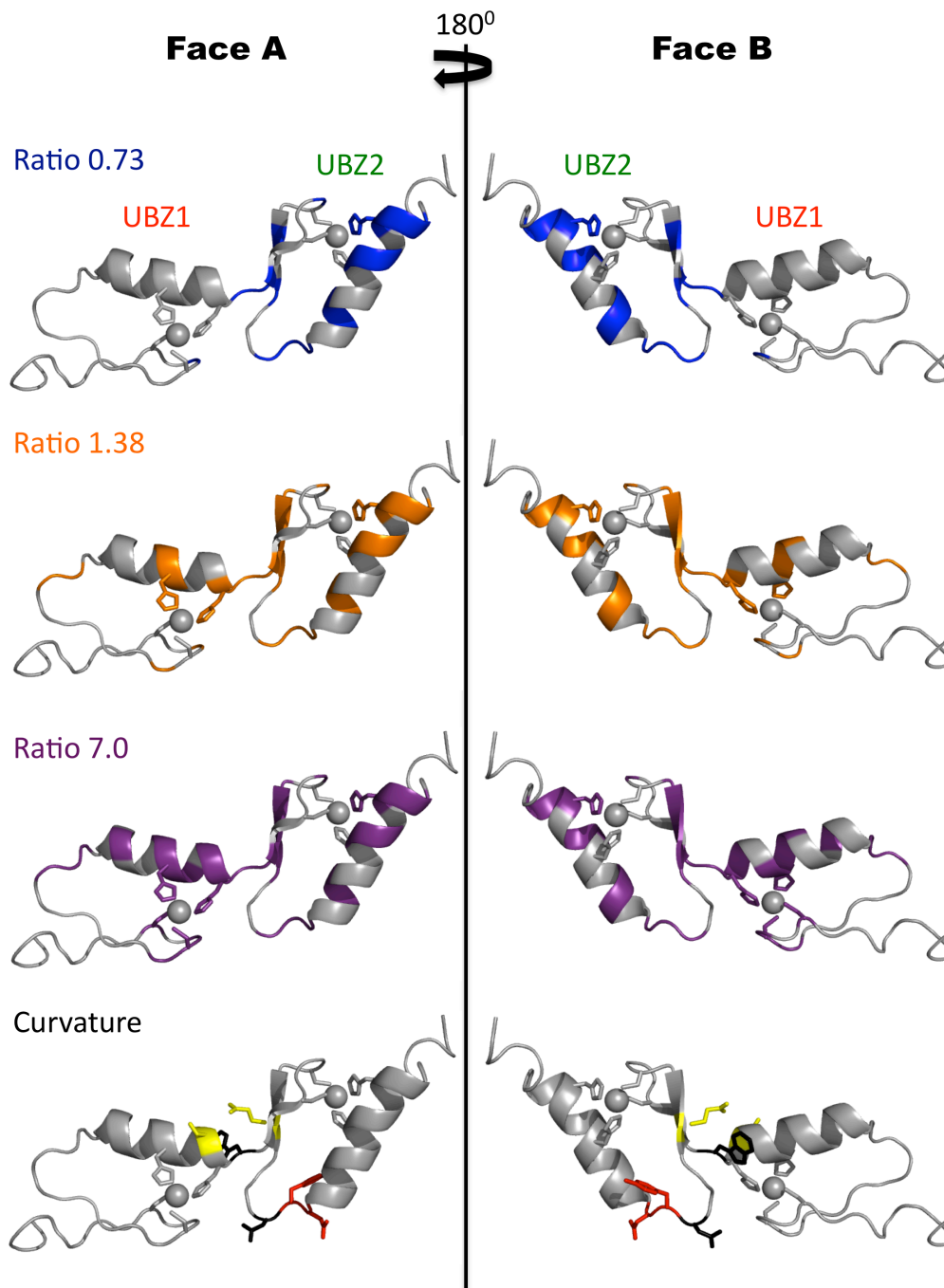


Figure V.24. Chemical shift perturbation mapping on the lowest-energy UBZ1+2 NMR structure (grey cartoon) at different UBZ1+2:ubiquitin ratios. Side chains of the Cys and His residues that coordinate the zinc atoms (grey spheres) are shown in sticks. The residues affected upon binding to ubiquitin are colored according to the scheme in figure V.23. Every molecule is represented in two orientations (related by 180° rotation) denoted as **Face A** and **Face B**. At each protein ratio, the chemical shift cut-off is the mean $\Delta\delta$ value plus one standard deviation. The bottom panel shows the residues with curved chemical shifts perturbation trajectories. The different coloring corresponds to the three different curving patterns observed during the titration: yellow (Ser32 and Gln43), red (Tyr48 and Asp49) and black (Trp34 and Asp47). Corresponding side chains are shown as sticks.

Figure V.24 shows how the binding shifts of UBZ1+2 resonances evolve upon addition of the increasing amounts of ubiquitin. At the beginning of the titration, the biggest chemical shift perturbations are observed in the UBZ2 domain, affecting the residues located on face A at the C-terminal part of the α -helix (residues Val57, Gln58, His60, Phe61 and Asp62) but also the Met39 in the loop connecting the two β -sheets of UBZ2 that is oriented to the face A of UBZ2 and likely interacts with ubiquitin in coordination with the C-terminal part of the α -helix. The binding of ubiquitin to this region is affecting the residues involved in the tetrahedral coordination of the Zn^{2+} atom, especially His60. Phe44 is also strongly affected despite being not exposed to solvent; this effect can be explained considering the important role of Phe44 in the hydrophobic core of UBZ2. This Phe acts as an aromatic anchor of the $Cys_2His_2 Zn^{2+}$ -coordination by a strong interaction with the His residues (section 5.3.1.4). It is plausible that ubiquitin binding to the C-terminal part of UBZ2, which affects the His60, is also felt by Phe44 and other aromatic residues in the UBZ2 hydrophobic core (e.g. Phe53 and Tyr48). This, in turn, could lead to the rearrangement of the loop connecting the β -sheets with the α -helix of UBZ2 as evidenced by considerable chemical shift perturbations of residues Asp47, Tyr48 and Asp49 located in this region. This loop corresponds to the hinge region of the tandem (section 5.3.1.3), which might be especially sensitive to any type of rearrangement within the hydrophobic core, because it coordinates the relative orientation of the α -helix of UBZ2 with respect to UBZ1. While not observed in the 1H - ^{15}N -HSQC spectra, Pro45 and Pro46 are also part of this loop and show a big effect in the 1H - ^{13}C -HSQC spectra recorded during the titration (Fig. V.26 and V.27), where the $^1H\delta$ - $^{13}C\delta$ of Pro46 disappears and the $^1H\alpha$ - $^{13}C\alpha$ of Pro45 displays curvature.

Located between the two UBZ domains, residues Trp34 (in UBZ1) and Lys35 (in UBZ2) occupy a key position and are strongly affected upon binding. Their signals disappear in the spectra before reaching ratio 1:0.73 probably due to the extensive conformational exchange. As shown by the high number of NOEs detected in the 2D and 3D NOESY spectra, Trp34 is involved in a wide range of interactions with residues from both domains, including the hydrophobic core of UBZ2. Interestingly, the early chemical shift perturbation effect observed in Leu12 $^1H\delta$ - $^{13}C\delta$ (Fig. V.28) can be explained by the spatial connection between Leu12 and Pro46 (confirmed by NOE analysis, Fig. V.25).

It is important to take into account that both UBZ domains are strongly associated by the absence of a linker between them and by the existence of well-defined hydrophobic cores in each domain, which contribute to the rearrangement of the molecule upon binding ubiquitin in order to maintain the stability of the Zn^{2+} -coordination, crucial for the proper folding of the protein. Due to these peculiar circumstances it is logic to think that residues located in loops, hinge regions and not exposed areas like the hydrophobic core, are affected not by direct binding of ubiquitin, but by secondary binding effects such as conformational changes or side chain reorganization, while the residues exposed in the C-terminal part of the UBZ2 helix and Met39 are affected due to direct interactions with ubiquitin (Fig. V.25).

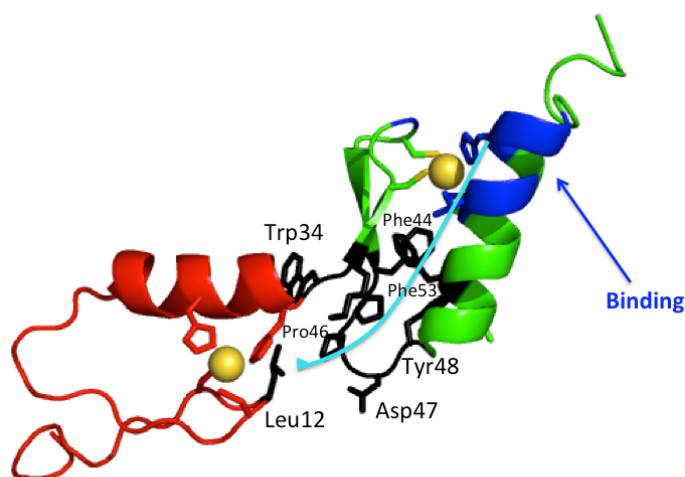


Figure V.25. Cartoon representation of the UBZ1 (in red) and UBZ2 (in green) domains of TAX1BP1 in tandem (Face A). Black sticks represent the side chains of the residues indirectly affected by the binding of ubiquitin to the α -helix of UBZ2 at the beginning of the titration (ratio 1:0.73). In blue, the region of the helix involved in direct binding to ubiquitin. The binding effect in the C-terminal part of the helix can be transmitted through the hydrophobic core to Trp34 and Leu12 located in UBZ1 domain, as indicated by a cyan arrow. Zn^{2+} atoms are shown as yellow spheres and side chains of the Cys and His responsible of Zn^{2+} -coordination are represented in sticks.

In general, upon increasing the ubiquitin concentration, the same UBZ2 resonances remain affected, and new residues located in UBZ1 start showing the binding shifts. The chemical shift perturbations of UBZ2 residues did not change much above ratio 1:1.38 (when the domain is saturated); however, residues located in UBZ1 domain continue to be affected till ratio 1:7.0. Under the saturation conditions, the side chains of the UBZ1 residues whose backbones are affected by ubiquitin binding are involved in Zn^{2+} coordination (His10, Cys13, His29 and His33), oriented to face A of the tandem (Glu14, Asn20 Asp22, Lys25, Glu28 and Ser32), or positioned to face B of the tandem (Leu12, Tyr21, Val30 and Trp34). Close inspection of the ^{13}C -HSQC spectra shows that H δ protons of Pro18 and Pro19 also start to be strongly affected above 1:1.38 ratio.

The information obtained from the UBZ1+2 NMR titrations with ubiquitin has been crucial to establish which residues between both domains are responsible for the communication within the tandem. As can be observed in figure V.27, at least one of the two Pro H δ signals disappears at low ratios during the titration (Pro11 and Pro46). These proline residues are very close to Leu12, whose NH signal disappears in the ^{15}N -HSQC spectra at the beginning of the titration, and their methyl H δ C δ signals show substantial perturbations (including slight curvature) during the whole titration (Fig. V.28). The spatial proximity between Leu12 and Pro46 is evident by the detection of intramolecular NOEs between the methyl protons of Leu12 and H β and H γ protons of Pro46. In the same region, intramolecular NOEs (e.g. Leu12-Trp34, Trp34-Pro46, Trp34-Glu31, Trp34-Gln43) are also detected.

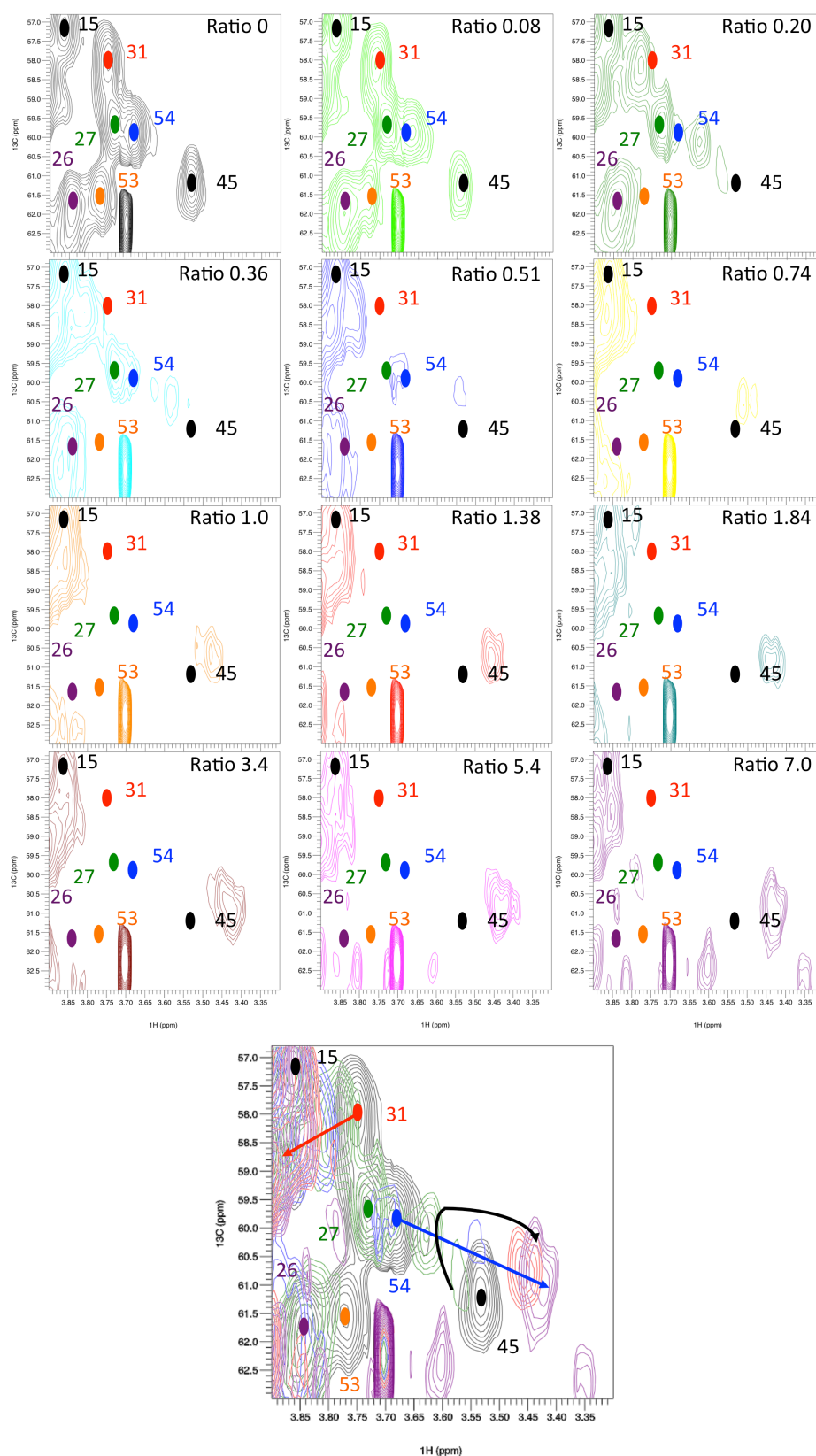


Figure V.26. Selected $H\alpha C\alpha$ region of the 1H - ^{13}C -HSQC spectra of UBZ1+2 during the titration with ubiquitin at different ratios. The color-coding is the same as in the corresponding 1H - ^{15}N -HSQC titration (Fig. V.21). $H\alpha C\alpha$ assignment is indicated. At all titrations points the 1H - ^{13}C peak position in free UBZ1+2 is indicated with a dot colored according to the residue number.

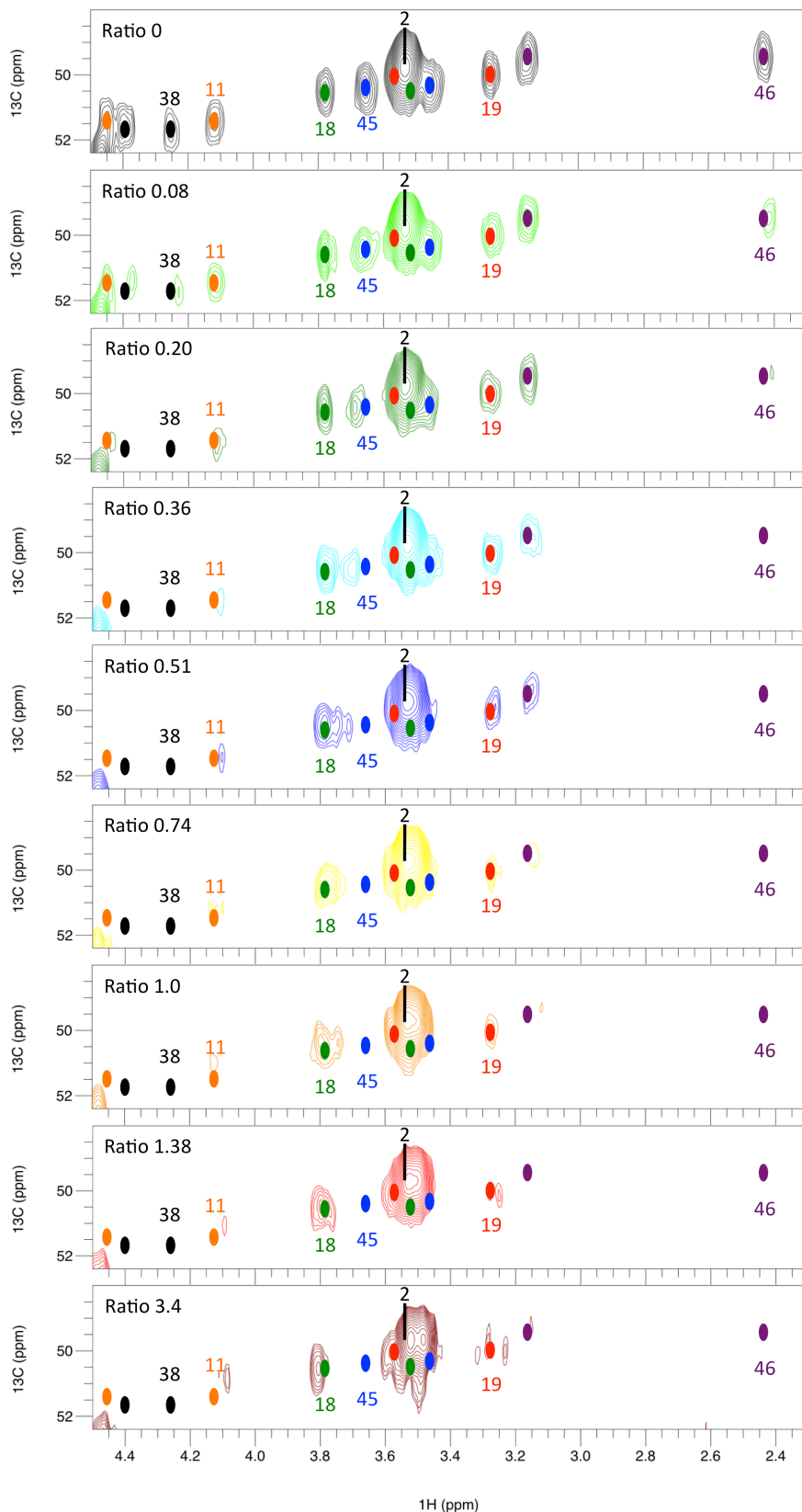


Figure V.27. Prolines $H\delta C\delta$ region of the 1H - ^{13}C -HSQC spectra of UBZI+2 during the titration with ubiquitin at different ratios. The color-coding is the same as in the corresponding 1H - ^{15}N -HSQC titration (Fig. V.21). Resonance assignments are indicated. At all titration points the 1H - ^{13}C peak position in free UBZI+2 is indicated with a dot colored according to the residue number.

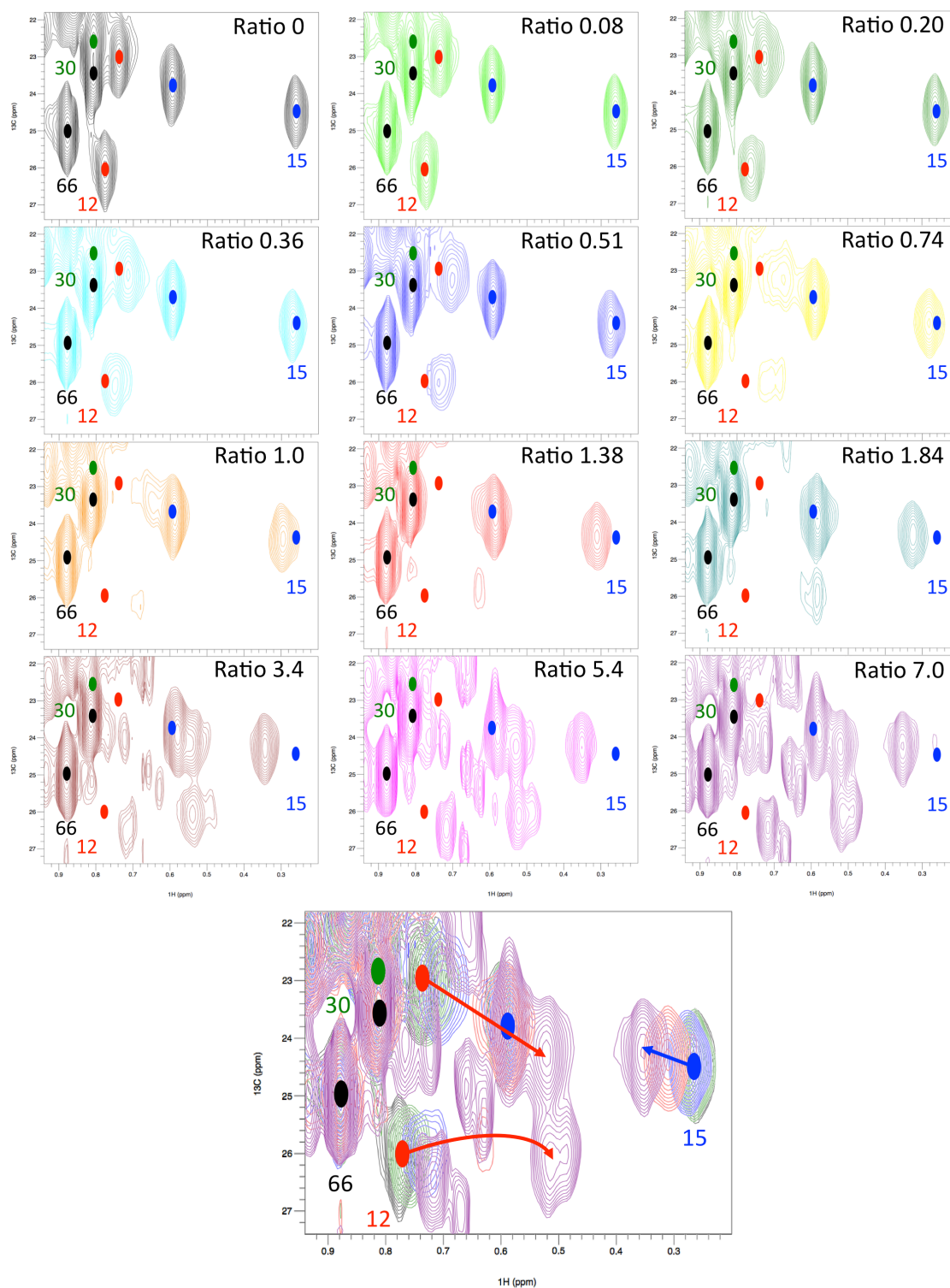


Figure V.28. Selected methyl region of the ^1H - ^{13}C -HSQC spectra of UBZ1+2 during the titration with ubiquitin at different ratios. The color-coding is the same as in the corresponding ^1H - ^{15}N -HSQC titration (Fig. V.21). Side-chain assignments are indicated. At all titration points the ^1H - ^{13}C peak position of each methyl in free UBZ1+2 is indicated with a dot colored according to the residue number.

As can be seen in figure V.26, the H α C α signal of Glu31 is affected at early stages of the titration, despite the residue is located in the UBZ1 domain. This is mainly due to its position between both UBZ domains, with its CO being involved in a hydrogen bond with the N ϵ 2 group of Gln43. The N ϵ 2 group of Gln43 also makes a hydrogen bond with the CO of Trp34. It is worth to mention that Gln43 is located in a key position close to the α -helix of UBZ1 domain, sensing environmental changes coming from both faces of the tandem. Probably this is why it is one of the residues that experience a drastic curvature in the chemical shift perturbation profile for both the backbone NH (Fig. V.21) as well as for the side chain H ϵ N ϵ (Fig. V.22).

We can conclude that Pro11, Leu12, Glu31, Trp34, Gln43 and Pro46 are structurally interconnected. Moreover, they define the interface between both UBZ domains and seem to communicate the conformational changes elicited in the UBZ2 domain upon binding to ubiquitin to the UBZ1 (Fig. V.29). In turn, this could lead to possible reorientations, which facilitate ubiquitin binding to the UBZ1 domain.

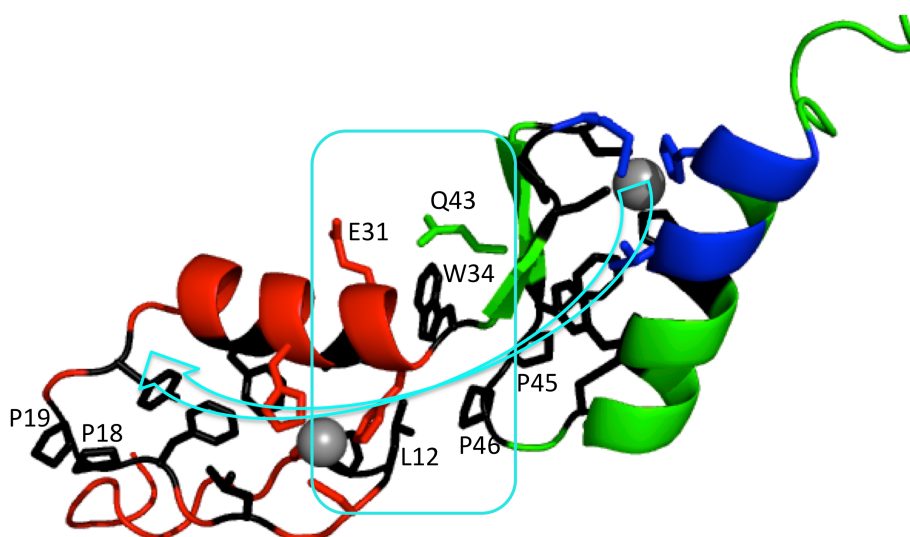


Figure V.29. Communication within the tandem of UBZ1 (red) and UBZ2 (green) domains through the hydrophobic core (black sticks). The cyan arrow indicates the direction of the communication flow upon binding of ubiquitin to the C-terminal part of the α -helix and Met39 in UBZ2 (blue). Glu31-Gln43 and Leu12-Pro46 together with Trp34 are the key interface residues.

The in-depth chemical shift analysis described above indicates possible cooperative effects in the ubiquitin binding to UBZ1+2. To confirm this hypothesis, we compared the binding curves of the individual residues located in both domains ($\Delta\delta$ vs Molar ratio). Figures V.30 and V.31 show the binding curves for some of the residues located in UBZ1 and in UBZ2, respectively. They present two clearly different patterns. The residues located in the UBZ1 domain show a sigmoidal behavior, different from the exponential one observed for the residues located in the UBZ2 domain. As can be seen in figure V.31, the binding curves of the UBZ2 residues saturate at ratio ~ 1.0 , at which point the UBZ1 residues start to experience a large increase in the chemical shift changes (Fig. V.30, inflexion point at ratio 1.0). Sigmoidal curves are indicative of cooperative binding (Ferrell J.E., 2009). Indeed, when the experimental chemical shifts were analyzed with a single-binding site model (see Chapter III, section 3.3.3.3), only the titration curves for the residues located in the UBZ2 domain could be accurately

fitted (Fig. V.31), and not those for the residues in UBZ1 (Fig. V.30). Binding curves of the residues located in UBZ1 domain can only be adequately fitted using an equation that includes the Hill coefficient, thus taking into account possible cooperativity effects (Chapter III, section 3.3.3.3).

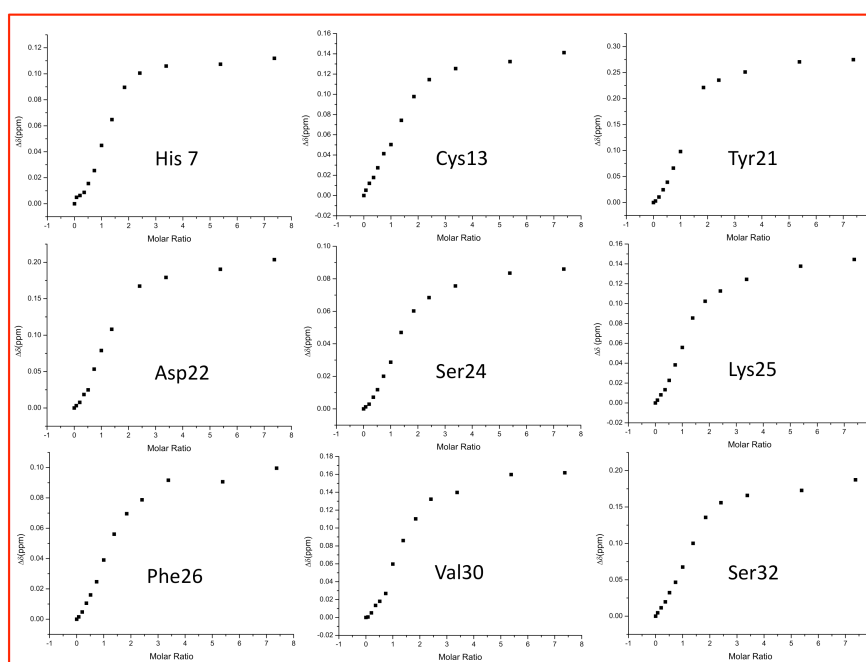


Figure V.30. Chemical shift changes vs molar ratio for some of the residues in the UBZ1 domain affected during the titration with ubiquitin. The titration curves show a sigmoidal shape characteristic of cooperative binding.

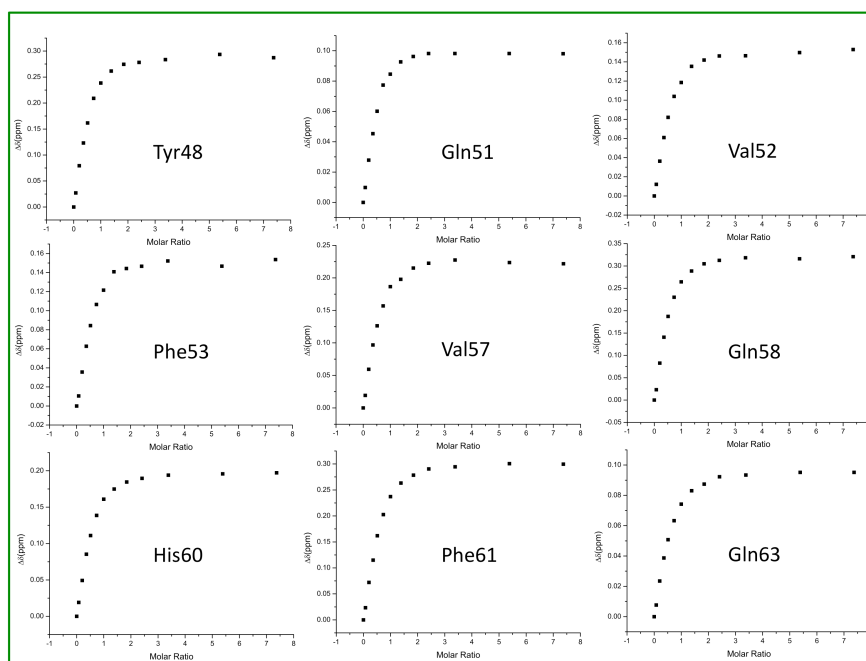


Figure V.31. Chemical shift changes vs molar ratio for some of the residues in the UBZ2 domain affected during the titration with ubiquitin. The titration curves show an exponential shape characteristic of single binding interactions.

Such fits, shown in figure V.32, allow calculating the individual K_d s for the ubiquitin interaction with the residues located in UBZ1 and UBZ2 domains and provide the Hill coefficient, which indicates if there is a cooperativity effect.

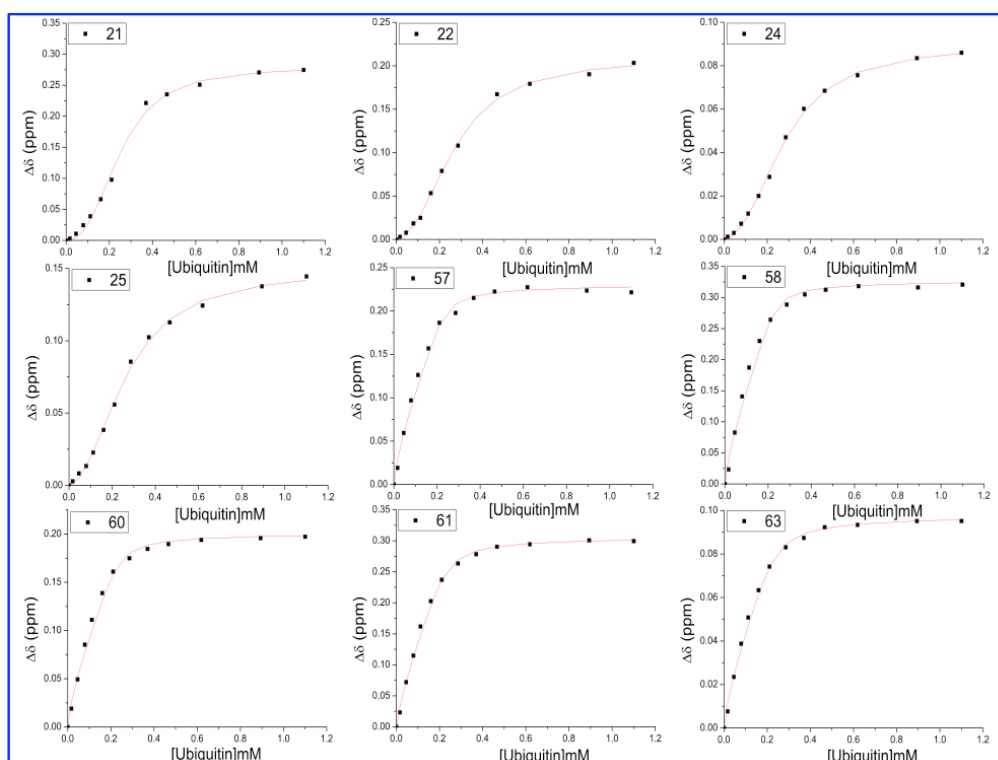


Figure V.32. Fitting (red line) of the experimental data (black dots) obtained during the titration of ^{15}N - ^{13}C -labeled UBZ1+2 with unlabeled ubiquitin (Fig. V.21). A mathematic model accounting for possible cooperativity effects with the Hill coefficient (α) fits the titration curves (Chapter III, Eq.20). Residues represented are: Tyr21, Asp22, Ser24, Lys25, Val57, Gln58, His60, Phe61 and Gln63.

UBZ1+2 residue number	K_d (μM)	Shift _{max}	α
21	144.1 ± 9.9	0.280 ± 0.007	2.0 ± 0.1
22	165.3 ± 8.9	0.209 ± 0.004	1.7 ± 0.1
24	175.9 ± 6.5	0.089 ± 0.001	1.8 ± 0.1
25	159.3 ± 8.4	0.149 ± 0.003	1.6 ± 0.1
57	7.3 ± 1.8	0.228 ± 0.009	1.0 ± 0.1
58	5.9 ± 1.0	0.323 ± 0.009	0.9 ± 0.1
60	7.1 ± 1.7	0.198 ± 0.009	1.0 ± 0.1
61	10.9 ± 2.0	0.303 ± 0.005	1.0 ± 0.1
63	12.2 ± 3.2	0.096 ± 0.001	1.0 ± 0.1

Table V.4. Fitting parameters of the titration curves shown in the Fig. V.32. The first column lists the residues of UBZ1 (red) and UBZ2 (green). The second column reports the values of the binding constant (K_d), the third column the maximum chemical shift changes upon saturation, and the fourth column the corresponding Hill coefficient (α). Uncertainties are the errors of the fits.

The general equation accounting for possible cooperative effects fits well the experimental data, irrespectively of the residue location in either UBZ1 or UBZ2 domain. The difference in shapes (sigmoidal or exponential) is reflected in the Hill coefficient (α). For the residues located in UBZ2, $\alpha = 1$ is found, which implies the absence of cooperativity as expected from the exponential shape of the binding curve. However, for the residues located in the UBZ1 domain $1 \ll \alpha \leq 2$ is obtained, which indicates positive cooperativity, in agreement with the sigmoidal shape of the binding curve.

The K_{dS} obtained from the titration of the tandem UBZ1+2 with ubiquitin, are in the order of 10 μM ($8.7 \pm 3.7 \mu\text{M}$) for the residues located in the UBZ2 domain and around 150 μM ($161.1 \pm 13.3 \mu\text{M}$) for those in the UBZ1 domain. These values are in good agreement with those obtained from the ITC analysis of the UBZ1+2-ubiquitin interaction with the model of two non-symmetric binding sites (Fig. V.16), which also showed positive cooperativity. In agreement with the ITC analysis presented above (section 5.3.2), the NMR titrations show that the tandem of UBZ1 and UBZ2 domains increases the binding affinities for ubiquitin as compared with single, isolated UBZ domains (UBZ1 from $346.8 \pm 37.3 \mu\text{M}$ to $161.1 \pm 13.3 \mu\text{M}$ and UBZ2 from $36.5 \pm 5.5 \mu\text{M}$ to $8.7 \pm 3.7 \mu\text{M}$). Either in isolation or in tandem, UBZ2 always binds ubiquitin one order of magnitude tighter than does UBZ1.

Considering UBZ1+2 as a multi-subunit protein, we can explain the affinity increase for both UBZ domains by a cooperative domain-domain interaction, often referred to as enforced proximity or the avidity effect, which is common in protein-protein interactions (Ferrel E.J., 2009). Based on the chemical shift perturbation analysis we have hypothesized that the hydrophobic core of UBZ1 and UBZ2 is implicated in the rearrangement of the tandem upon binding of ubiquitin to UBZ2. By the analysis of the ^{15}N - and ^{13}C -HSQC titrations we have been able to define the residues whose backbone and side chain atoms form the interface between the UBZ domains (Fig. V.29), crucial for the communication within the tandem.

Other ubiquitin-binding domains, like the UIM12 motif of Ataxin-3 (AT3), show increased binding affinities in tandem as compared to single, isolated ones (Song A.X. et al., 2010). In case of AT3, the binding cooperativity is enabled by a short linker region between the two UIMs (only three residues, Glu239, Ile240 and Asp241). However, other UIM domains, such as those presented in S5a (subunit 5a of the 26S proteasome) and Vps27p (Vacuolar protein sorting-associated protein 27), do not show cooperativity, because they contain a long (20 - 50 residues) non-structured linker between the two UIMs (Swanson K.A. et al., 2003; Fujiwara K. et al., 2004; Wang Q. et al., 2005). For the UBZ1 and UBZ2 domains of TAX1BP1, the situation is more peculiar since there is no real linker region. This raises a possibility that the positive cooperativity is mediated by the conformational changes in the loop and hinge region of the tandem upon the tighter UBZ2-ubiquitin binding event (i.e. conformationally driven cooperativity).

Until now, the interaction of UBZ1 with ubiquitin has passed unnoticed. For example, co-immunoprecipitation studies and GST-pull down assays on TAX1BP1 have only identified the UBZ2-ubiquitin interaction (Verstrepen L. et al., 2011; Iha H. et al., 2008). The chemical shift mapping of the UBZ1+2-ubiquitin binding on both protein surfaces, together with the determination of the K_{dS} for the binding events to both UBZ

domains (isolated and in tandem), has been crucial to establish the role of UBZ1 in the interaction mechanism of TAX1BP1 with ubiquitin and to determine how the interaction of ubiquitin is communicated through the UBZ2 domain to UBZ1.

5.3.4 Backbone dynamics of UBZ1+2 tandem in complex with ubiquitin

NMR is a very useful tool that enables detailed characterization of local and global dynamic properties of proteins in aqueous solution. Since the beginning of the 90's, heteronuclear ^{15}N relaxation studies have been widely used to characterize the backbone dynamics and motional properties of many protein molecules (Barbato G. et al., 1992; Palmer A.G. et al., 1996; Kay L.E., 1998; Prompers J.J. et al., 1999; Malmendal A. et al., 1999; Ye J. et al., 1999).

^{15}N relaxation measurements (R1, R2 and hetNOE) were used to evaluate the backbone dynamics of the free and bound state of UBZ1+2 at different ratios with ubiquitin (1:0.5; 1:1.1; 1:5.0). For some residues, the relaxation parameters could not be measured either due to spectral overlapping or because of the severe line-broadening of some cross-peaks during the titration with ubiquitin (Fig. V.16), resulting in too low intensities of the signals for a reliable measurement (section 5.2.4).

Section 5.3.1.2 includes an in-depth analysis of the UBZ1+2 dynamic properties in the free form. Due to the complexity of the interaction model between UBZ1+2 and ubiquitin, including more than one binding site and taking into account the anisotropic shape of the system, we did not perform an exhaustive analysis of the overall and internal mobility of the complex. The main objective was to observe how the relaxation parameters of UBZ1+2 changed upon binding to ubiquitin.

The comparison of the R1, R2 and hetNOE values for the free UBZ1+2 and the complex between UBZ1+2 and ubiquitin at different ratios is shown in figure V.33. We noticed that, on average, the hetNOE values did not change a lot, the R1 values decreased, and the R2 values increased upon complex formation. R2 is directly dependent of the molecular rotational correlation time (τ_c) and thus on the molecular weight of the complex. The increase in the molecular weight of the complex upon ubiquitin binding gave rise to higher τ_c , as the motions became slower, which increased the R2 values.

In general the backbone amides with large hetNOE values (0.7 - 1.0) are expected to have lower flexibilities and to be located in the rigid region (α -helices and β -sheets), while those with small NOE values (< 0.6) are considered to be located in the loops or linker regions that have high flexibility (Farrow N.A. et al., 1994).

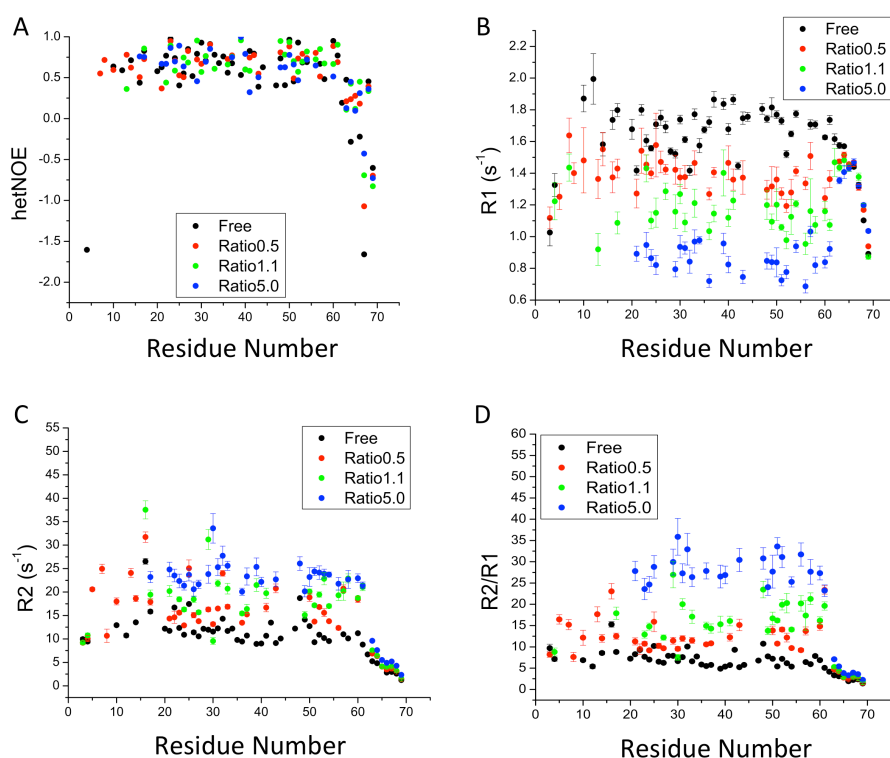


Figure V.33. Backbone dynamics of UBZ1+2 in the free form (black) and in complex with ubiquitin at different ratios: 1:0.5 (red), 1:1.1 (green) and 1:5.0 (blue). Measured hetNOE (A), ^{15}N longitudinal (R1) and transverse (R2) relaxation rates (B and C, respectively) values, are represented vs the residue number of UBZ1+2. D. Derived R2/R1 ratios plotted for each residue of UBZ1+2.

Almost all hetNOE values for UBZ1+2 residues in either free or ubiquitin-bound form are lower than the theoretical upper limit value of 0.84 at 600 MHz (Kay L.E. et al., 1989). Negative hetNOE values for the N-terminal (Met4) and the C-terminal (Asn64, Leu66, Asn67 and Asp69) residues show that these regions are much more flexible than other structural elements of the protein. However, upon increasing ubiquitin concentration, hetNOE values of the C-terminal residues of UBZ2 become less negative, indicating a loss of flexibility due to the complex formation (Fig. V.33A). The binding of ubiquitin is not affecting the hetNOE values of the residues located in the α -helix or β -sheet, showing that these regions of the UBZ1+2 do not become more rigid upon binding.

The ^{15}N R2/R1 ratio provides a good measure of the global tumbling rate of the NH vector and is independent of rapid internal motions and the magnitude of the chemical shift anisotropy (Sahu S.C. et al., 2000). Figure V.33D shows an increase in the R2/R1 ratio upon complex formation. The interaction of UBZ1+2 (8.3 kDa) with one or two ubiquitin molecules (8.6 kDa each), leads to the formation of a complex with higher molecular weight, which is reflected in the increase of the transverse relaxation rate (R2) of the system. However, the ratio R2/R1 almost does not change for residues located in the flexible C-terminal part (from 64 till 69), reflecting the negligible contribution of binding to the tumbling of this protein region, in agreement with the chemical shift mapping experiments (section 5.3.3.2).

For rigid protein regions in the limit of slow motion ($\tau_c > 0.5$ ns), τ_c can be described as a function of the R2/R1 ratio (Kay L.E. et al., 1989) (Fig. V.32D). In this way, average ^{15}N R2/R1 ratios of 6.8 (free UBZ1+2), 12.2 (ratio 1:0.5), 17.1 (ratio 1:1.1) and 28.3 (ratio 1:5.0) correspond to an apparent overall rotational correlation time τ_c of 7.7, 10.6, 12.8 and 16.7 ns, respectively. Since the rotational correlation time of a molecule is directly related to its molecular weight, NMR relaxation data clearly confirm the formation of an UBZ1+2-ubiquitin complex in solution. The interaction seems to involve more than one ubiquitin molecule, because the correlation time at a ratio 1:1.1 (when the binding site of UBZ2 is saturated) is significantly different from that at a ratio 1:5.0 (when UBZ1 domain should be almost saturated).

5.3.5 Paramagnetic tagging for structural determination of the UBZ1+2-ubiquitin complex

Here, we used NMR spectroscopy to elucidate the UBZ1+2-ubiquitin complex in solution. Upon binding to ubiquitin, the resonances of UBZ1+2 experience large chemical shifts changes. From the analysis of the chemical shift perturbations, it was possible to map the binding interfaces of UBZ1+2 (Fig. V.24) and ubiquitin (Fig. V.20). However, chemical shift perturbations do not provide information about the relative orientations of the protein molecules. The traditional NMR approach for structure determination of protein complexes in solution relies on the use of intermolecular NOEs (Gordon S.L. and Wüthrich K., 1978; Braun W. et al., 1981). However, this method has a limited use for weak protein complexes as it requires a relatively tight ($K_d \leq 10$ μM) interaction between molecules (Zuiderweg E.R., 2002). Furthermore, NOE values are restricted to distances less than 5 Å (see Chapter III, section 3.3.2.3), which makes this methodology unsuitable for characterization of long-range interactions typically found in weak protein complexes. For the last two decades, structural characterization of weak protein interactions has been greatly enhanced by the advent of the NMR techniques that use residual dipolar couplings (RDCs) and paramagnetic effects (PRE) to provide long-range intermolecular restraints (Zuiderweg E.R., 2002; Bonvin A.M. et al., 2005).

Based on the NOESY spectra of ubiquitin in complex with the isolated UBZ1 and UBZ2 domains at 1:2 ratio and the UBZ1+2-ubiquitin at ratio 1:5, several intermolecular NOEs were ambiguously identified between ubiquitin (most likely from Ile8 H δ 2, Thr9 H γ 2, Ile44 H δ 2 and Val70 H γ 1 and H γ 2) and some side chains of residues located within UBZ2 (likely Phe61 and Gln58) and UBZ1 (likely Gln23). However, this information was not sufficient to determine a reliable model of the UBZ1+2-ubiquitin complex. In order to determine the orientations of the ubiquitin molecules bound to UBZ1+2, we needed to obtain an adequate set of intermolecular distance restraints. Besides, it was also necessary to take into account the weak interaction of UBZ1 domain with ubiquitin, which could implicate the formation of lowly populated states of the complex under the equilibrium conditions. In order to satisfy these requirements, paramagnetic NMR was used to obtain an accurate model of the UBZ1+2-ubiquitin complex.

5.3.5.1 PREs in the complex of ^{15}N UBZ1+2 with ubiquitin

Distance restraints from paramagnetic relaxation enhancement (PRE), together with the exquisite sensitivity of the PRE NMR to detect transient, lowly populated intermediates in biomolecular interactions (Iwahara J. and Clore G.M., 2006; Tang C. et al., 2006; Volkov A.N. et al., 2006; Clore G.M. et al., 2007; Clore G.M. and Iwahara J., 2009), provided sufficient structural information to determine a reliable model of the complex of UBZ1+2 with ubiquitin and define the conformational space sampled by ubiquitin during the interaction with the weak binding site.

As explained in section 5.2.5, three ubiquitin mutants (D32C, D39C and E51C) were designed to strategically place a cysteine residue at the periphery of the ubiquitin binding surface as established by chemical shifts perturbation mapping. The introduced solvent-exposed Cys allowed the insertion of a metal chelator probe (pEDTA) able to bind paramagnetic (Mn^{2+}) and diamagnetic (Zn^{2+}) divalent metals with high affinity (see section 5.2.2.1 and Chapter III, section 3.3.4). These modifications neither changed the net charge of ubiquitin nor perturbed the UBZ1+2 binding equilibrium (see section 5.2.2.1, and Fig. V.37A). Merely two ^1H - ^{15}N -HSQC spectra per ubiquitin mutant (one with the paramagnetic sample and the other with the diamagnetic control) of the ^{15}N -labeled UBZ1+2 bound to ubiquitin-pEDTA conjugate were sufficient to determine intermolecular distance restraints for the amide protons. Consequently, this information was used in a rigid-body docking protocol to deduce the relative orientation of the two proteins in the complex (Tang C. et al., 2006; Volkov A.N. et al., 2010).

All NMR samples of ^{15}N -labeled UBZ1+2 in complex with ubiquitin mutants were prepared at 1:3 ratio to ensure the interaction of ubiquitin with both UBZ domains. Larger excess of the conjugated ubiquitin was avoided, because a high concentration of paramagnetic probe attached to ubiquitin might cause non-specific effects. Based on the K_{d} s determined from the UBZ1+2-ubiquitin NMR titration, 100% UBZ2 and 65.5% UBZ1 were estimated to be bound to ubiquitin at the protein concentrations used.

The backbone amide resonances of UBZ1+2 in complex with pEDTA-conjugated ubiquitin were monitored using ^1H - ^{15}N -HSQC spectra recorded for the paramagnetic sample and the diamagnetic control. Two pEDTA- Mn^{2+} probes (D39C and E51C) were located close to the binding site (mapped by NMR chemical shift changes) and gave rise to PREs, while the pEDTA- Mn^{2+} attached to the mutant D32C (located at the back of the protein, negative control) showed no significant effects. As an example, figure V.34 shows the paramagnetic effects experienced by the resonances of UBZ1+2 bound to ubiquitin E51C-pEDTA- Mn^{2+} (note the missing signals, broadened beyond detection due to large intermolecular PREs, in the red spectrum). Characterized by r^{-6} distance-dependence between the nucleus and the paramagnetic center, this effect is produced by the unpaired electrons of the paramagnetic probe, which increase the relaxation rate of the neighboring nuclei due to the magnetic dipolar interactions.

For each observed amide proton of UBZ1+2, the intensity ratio ($I_{\text{para}}/I_{\text{dia}}$) of the resonances in the paramagnetic and diamagnetic samples were determined, and used to calculate the experimental PRE rates Γ_2 , necessary for the ensemble refinement (see Chapter III, section 3.3.4.2). This procedure was repeated with the three ubiquitin mutants (D32C, D39C and E51C), one at a time. The final dataset, comprising the Γ_2

values from all ubiquitin mutants, was used as an input in the subsequent structure calculations.

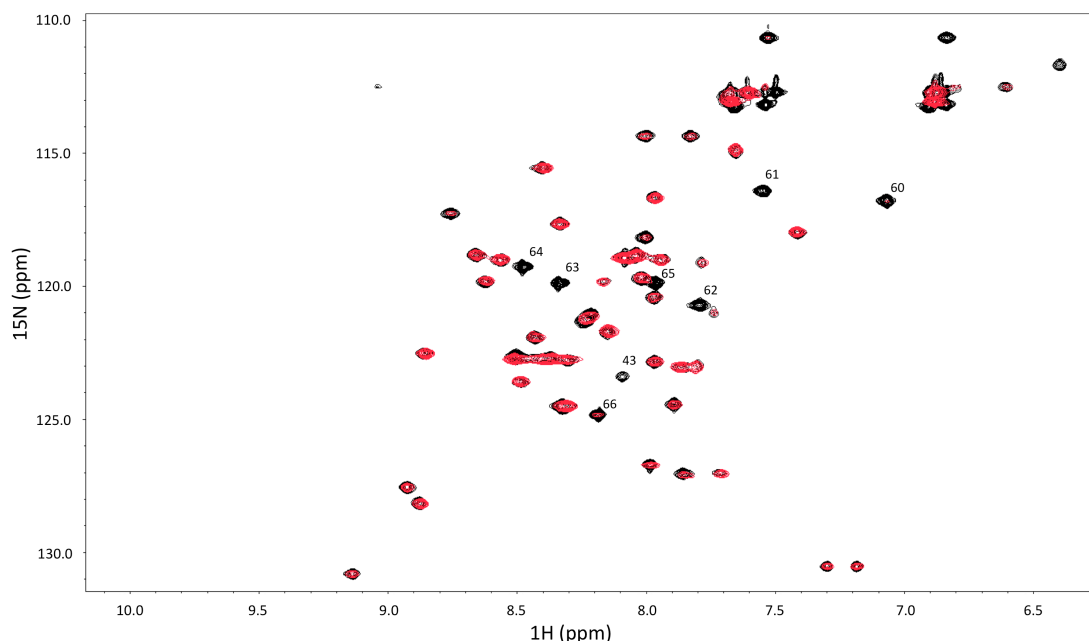


Figure V.34. ^1H - ^{15}N -HSQC spectra of ^{15}N -labeled UBZ1+2 in presence of 3 molar equivalents of ubiquitin-E51C-pEDTA- Zn^{2+} (diamagnetic, black) and ubiquitin-E51C-pEDTA- Mn^{2+} (paramagnetic, red). The experiments were conducted in 50 mM NaPi pH 6.0 at 25 °C on Varian 800 MHz spectrometer. Several backbone amide resonances strongly affected by the paramagnetic probe are indicated by the labels.

5.3.5.2 Structure calculations of the UBZ1+2-ubiquitin complex

The restraints from all three pEDTA- Mn^{2+} positions were simultaneously used in the structure calculation of the protein complex. To take into account the rotational freedom of the pEDTA- Mn^{2+} probe attached to the surface of ubiquitin, we performed ensemble averaging of the intermolecular distance restraints using multiple starting points of the Mn^{2+} atom as described in Chapter III (section 3.3.4). For the three spin-labeled ubiquitin variants, the ensembles of Mn^{2+} atom coordinates are shown in figure V.35.

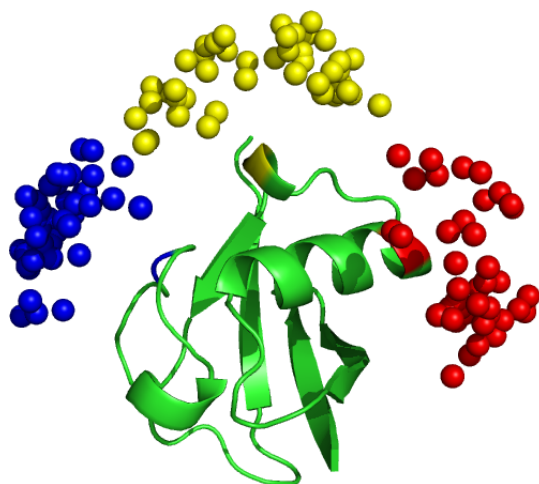


Figure V.35. Cartoon representation of ubiquitin (in green) with the attached pEDTA- Mn^{2+} . To take into account the mobility of the probe, 45 conformers were generated for every mutant. The positions of the Mn^{2+} atoms in the different conformers are represented as colored spheres. Color code: red (mutant D32C), yellow (mutant D39C) and blue (mutant E51C).

The use of intermolecular distance restraints from several pEDTA-Mn²⁺ positions increases the accuracy and precision of the structure calculations, as compared to the case of a single pEDTA-Mn²⁺ probe. In the same way, averaging of the restraints over a number of possible orientations of the paramagnetic probe further enhances the quality of the final solutions (Iwahara J. et al., 2004).

Direct use of the experimental Γ_2 restraints in an ensemble-based, rigid-body simulated annealing structure calculation protocol was pioneered by Clore and co-workers (Iwahara J. et al., 2004; Schwieters C.D. et al., 2006; Tang C. et al., 2006). Following this approach, multiples copies of ubiquitin were docked simultaneously to UBZ1+2 by minimizing the difference between the combination of PREs from all ubiquitin molecules and the experimental Γ_2 values. We used a rigid-body protocol as our experiments did not provide any structural information on the protein side-chains.

We performed multiple structure calculations using an ensemble of ubiquitin molecules, with the ensemble size (N) varying from 1 to 20. To assess the quality of the solutions, we computed a Q factor (Q_f) (Iwahara J. et al., 2003), which is a measure of agreement with the experimental data (the smaller the Q_f , the better the agreement). For the ubiquitin binding to UBZ2, N = 1 was found to be sufficient to accurately reproduce the measured Γ_2 s. The Q_f values were plotted as a function of N for the ubiquitin ensemble interacting with UBZ1 (Fig. V.36D). Q_f diminishes with the increasing ensemble size, reaches a minimum of 0.29 at N = 10, and then increases again, indicating that N = 10 is the optimal size of the ubiquitin ensemble required to satisfy the experimental restraints.

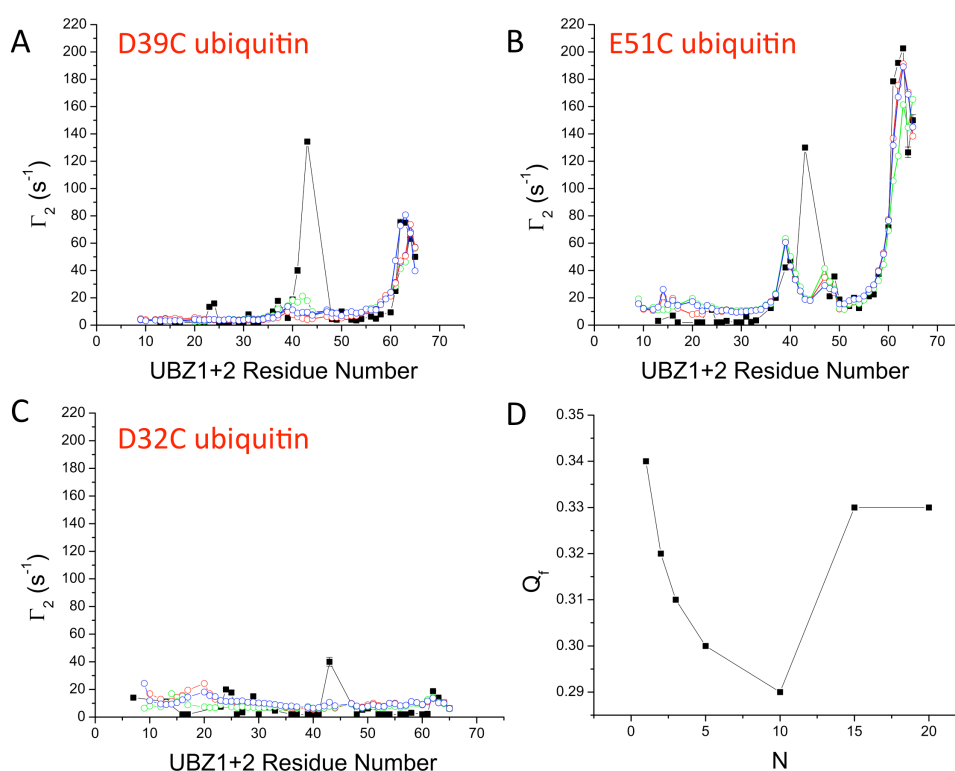


Figure V.36. Experimental Γ_2 s (in black symbols) and Γ_2 s back-calculated from three representative solutions (open symbols) for the UBZ1+2 bound to D39C (A), E51C (B) and D32C (C) pEDTA-Mn²⁺ ubiquitin. D. Q_f plotted as a function of N (size of the ubiquitin ensemble interacting with UBZ1 domain).

As can be seen from the comparison of experimental and back-calculated Γ_{2s} PRE profiles (Fig. V.36), a combination of PREs from the ubiquitin bound at both UBZ1+2 sites in the refined complex ensemble provides a good agreement with the experimental data. Being strongly affected in all cases, the residue Gln43 is consistently violated for reasons that are not immediately apparent at this stage. PREs for the N- and C-terminal UBZ1+2 residues (2 - 8 and 66 - 69) were not used in the ensemble refinement due to the high mobility of the UBZ1+2 termini. Interestingly, when used in the ensemble refinement of the complex with ubiquitin, the NMR structure of UBZ1+2 provides better agreement with the experimental data than its X-ray structure (as evidenced by lower Q_f and fewer PRE violations). In this sense, PRE experiments have validated the UBZ1+2 NMR structure.

To visualize the distribution of ubiquitin molecules in the complex with UBZ1+2, we used a reweighted atomic probability density map (Schwieters C.D. and Clore G.M., 2002; Volkov A.N. et al., 2010), calculated from the 20 best, independently generated ensembles with $N = 10$ (Fig. V.37). The ubiquitin binding to UBZ2 domains is very well defined, as the ubiquitin-UBZ2 solutions in multiple docking runs converge to the same position. In contrast, the ensemble of ubiquitin molecules that interact with the UBZ1 domain is spread over two regions at both sides of UBZ1 domain with similar relative populations (60% in face A and 40% in face B) (Fig. V.37). This finding indicates that the interaction between ubiquitin and UBZ1 is not only weaker than that with UBZ2 (as determined by ITC and NMR), but also more dynamic.

Based on this structural model, we can conclude that UBZ1+2 tandem interacts with ubiquitin using two binding sites, one located in UBZ2 domain, where ubiquitin occupies a single, well-defined orientation, and other in UBZ1 domain, occupied by an ensemble of ubiquitin molecules similarly distributed at both sides of UBZ1.

The binding interfaces of UBZ1+2 and ubiquitin mapped by the chemical shift perturbations fit well the PRE-based model (section 5.3.3.2). It is important to notice that the existence of two populated states at both sides of UBZ1 domain (also suggested by ITC and NMR experiments) might underlie functional promiscuity in the recognition mechanism of polyubiquitin chains.

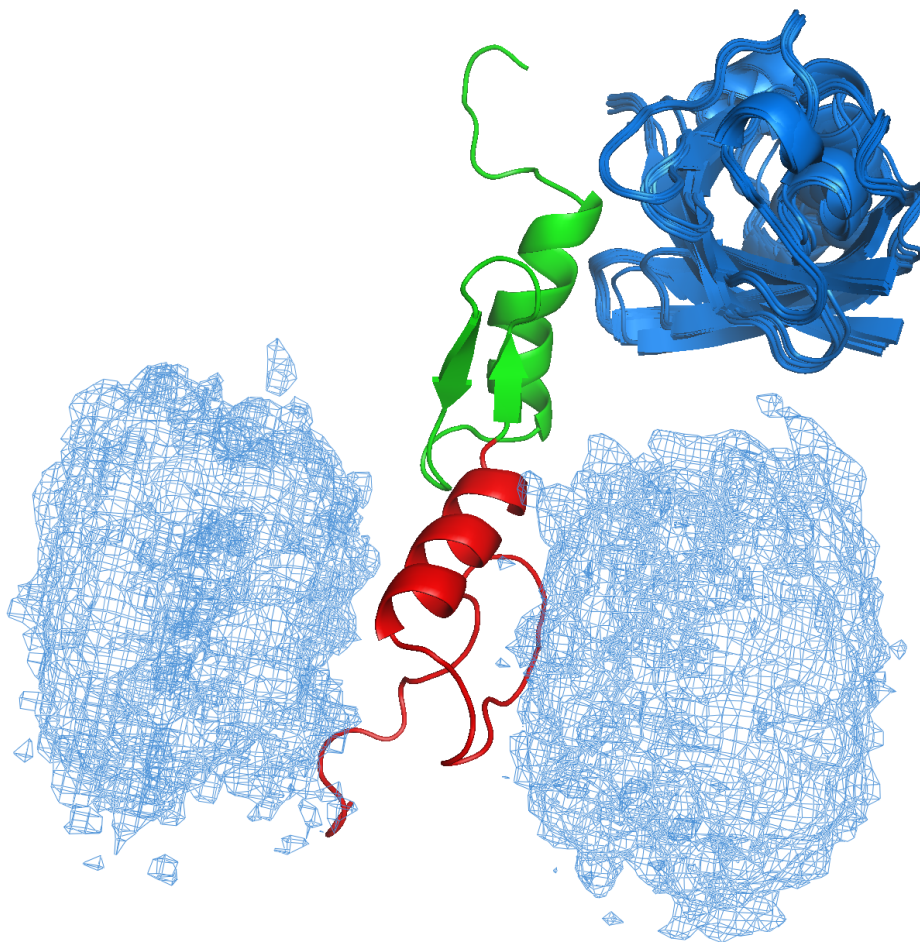


Figure V.37. Cartoon representation of the tandem formed by UBZ1 (red) and UBZ2 (green) domains of TAX1BP1 in complex with ubiquitin. The complex is a superposition of the 20 best PRE-based ensemble calculations ($N = 10$). Ubiquitin molecule (blue cartoon) binds UBZ2 domain in a single, well-defined orientation. Ubiquitin molecules binding to the UBZ1 domain are represented as a reweighted atomic probability density map, plotted as a blue mesh at a threshold of 30% maximum.

5.3.5.3 Experimental controls

To ensure the accuracy of the experimental data, it was necessary to setup some controls.

In the first place, to check whether the probe attachment to the surface of ubiquitin interferes with UBZ1+2 binding, we observed the chemical shift perturbations of UBZ1+2 in the presence of 3 molar equivalents of ubiquitin-Cys-pEDTA-Zn²⁺. In all cases, binding shifts were identical to those of the complex with the ubiquitin WT at the same protein ratio (Fig. V.38A), indicating that the attached pEDTA does not alter the binding.

It was also important to check the stability of the samples and the reproducibility of the measurements. Figure V.38B shows the comparison of the signal intensities in the HSQC spectra of UBZ1+2 in complex with the diamagnetic tag, ubiquitin-E51C-pEDTA-Zn²⁺, acquired immediately after the probe conjugation and 48 hours after.

Judging from the perfect linear correlation, the sample remains stable, enabling highly reproducible NMR experiments.

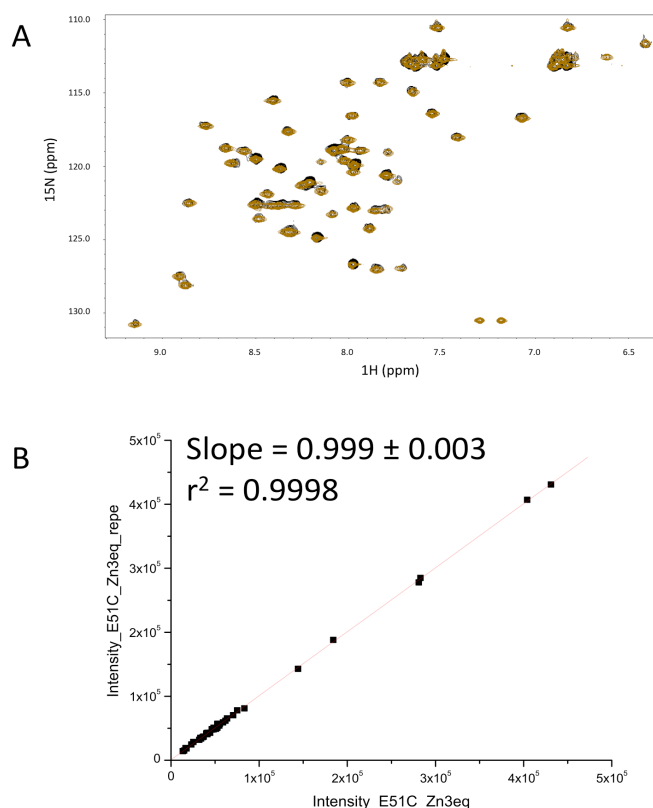


Figure V.38. *A.* Overlay of the ^1H - ^{15}N -HSQC spectra of the ^{15}N -labeled UBZ1+2 in the presence of 3 molar equivalents of WT (black) and E51C-pEDTA-Zn²⁺ ubiquitin (brown). *B.* Comparison of the cross-peak intensities of the sample of ^{15}N -labeled UBZ1+2 in complex with ubiquitin-E51C-pEDTA-Zn²⁺ measured with a delay of 48 hours.

5.3.6 Small Angle X-ray analysis of the free and bound UBZ1+2

SAXS provides information about the shape and the size of molecules samples from a population of randomly oriented molecules in solution (Koch M.H. et al., 2003; Putnam C.D. et al., 2007).

In this way, SAXS was a useful tool to validate the solution structure of UBZ1+2 obtained by NMR. The value of the χ^2 (less than one, Fig. V.39A) indicated the excellent fitting between the experimental scattering data and the one predicted from the NMR model using the program CRY SOL (Fig. V.39). For modeling we used the program DAMMIF (Svergun D.I., 1999). All the models predicted from the scattering curves, were highly reproducible in independent runs and the *ab initio* reconstructed shape fitted perfectly with the UBZ1+2 NMR structure.

Furthermore, the results obtained from the thermodynamic and structural characterization of UBZ1+2 in complex with ubiquitin, strongly support a scenario where UBZ1+2 interacts with ubiquitin through the two UBZ domains, being the interaction to UBZ2 well defined and tighter than the interaction to UBZ1, one order of magnitude weaker and more dynamic. The tighter interaction of UBZ2 domain with ubiquitin has been confirmed by SAXS. The PRE-based model of UBZ2-ubiquitin

complex fits perfectly the experimental scattering data, with very good χ^2 fitting statistics (Fig. V.39B).

On the other hand, the weak interaction of UBZ1 with ubiquitin could not be detected due to the measurement procedure used in the synchrotron (section 5.2.6), that requires to pass the sample through a size exclusion column before to acquire the data, which caused the disruption of the complex.

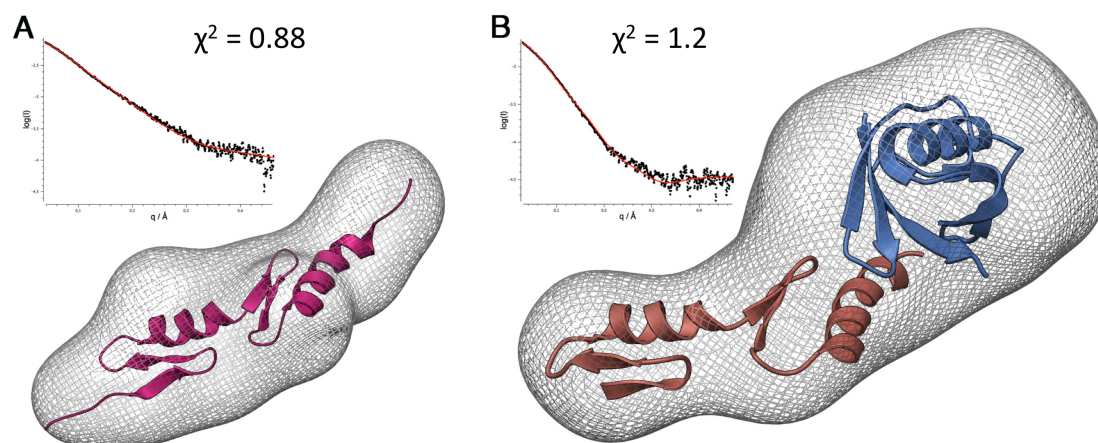


Figure 39. *A. SAXS analysis of the free UBZ1+2. From top to bottom: χ^2 fitting statistics; experimental SAXS scattering curves for the free UBZ1+2 tandem, the fitting of the NMR model to the experimental data is shown as a red solid line; Ab initio reconstructed shape superposed onto the NMR structure of UBZ1+2. B. SAXS analysis of the complex formed between UBZ2 and ubiquitin. From top to bottom: χ^2 fitting statistics; experimental SAXS scattering curves of UBZ1+2 in complex with one ubiquitin, the fitting of the PRE-based model to the experimental data is shown as a red solid line; Ab initio reconstructed shape superposed onto the PRE-based model of UBZ1+2 in complex with one ubiquitin bound to UBZ2.*

5.3.7 Interaction of UBZ1+2 with Lys48-, Lys63- and linear polyubiquitin chains

NMR, ITC and SAXS and especially PRE NMR, have been crucial to determine an accurate model of the binding complex of UBZ1+2 with ubiquitin, showing that the tight single-oriented complex of the UBZ2 domain with ubiquitin enhances the interaction of ubiquitin with the UBZ1 domain, mainly through two populated states at both sides of the domain. This dynamic binding event in UBZ1, together with the structural disposition of the two tightly connected UBZ domains might promote an appropriate conformation to bind different linkages of polyubiquitin chains, like Lys48-, Lys63- and linear linked polyubiquitins.

But can this model be validated with experimental data? Has UBZ1+2 a preference for any type of polyubiquitin chains linkages? This information would be crucial to understand the biological role of the protein TAX1BP1 in the different ubiquitination/deubiquitination pathways in which is involved, especially within the regulation of the nuclear factor NF- κ B.

Nuclear factor NF- κ B signaling is regulated by Lys48-, Lys63- and linear polyubiquitin substrates (Skaug B. et al., 2009; Iwai K. and Tokunaga F., 2009). Until now,

TAX1BP1 has been reported to participate in the negative regulation of the NF- κ B signaling pathway by recognizing Lys63-linked substrates (TRAF6, RIP1, Tax1).

To explore if TAX1BP1 would be able to recognize other than Lys63-linked substrates, as suggested by the structural model presented in this thesis, we have examined the avidity of the tandem UBZ1+2 for the three principal types of polyubiquitin chains (Lys48-, Lys63- and linear polyubiquitin chains) using bio-layer interferometry (BLI) (see Chapter III, section 3.6). Our results indicate that UBZ1+2 can recognize not only the previously known Lys63-linked polyubiquitin chains but also at least two other types of polyubiquitin linkages, respectively linear and Lys48-linked.

We measured the binding avidity of UBZ1+2 toward linear, Lys48- and Lys63-diubiquitin and tetraubiquitin chains. Sensor tips with each type of polyubiquitin chain immobilized were immersed in solutions containing a range of concentrations of UBZ1+2 tandem (from 0 to 300 μ M). We could not measure the interaction of Lys48-diubiquitin chains due to problems with the coupling of this protein to the sensor. Apparent dissociation constants were obtained from the steady-state analysis (response at equilibrium plotted as a function of UBZ1+2 concentration, Fig. V.40 and Table V.5) for each measured polyubiquitin linkage.

The obtained apparent K_{ds} (Table V.5) indicate that UBZ1+2 binds linear polyubiquitin chains with more avidity (9.3 μ M and 7.3 μ M for di- and tetraubiquitin chains, respectively) than monoubiquitin (66.98 μ M). Moreover, the binding to linear diubiquitin chains is almost one order of magnitude higher than the binding to Lys63-diubiquitin chains (74.3 μ M), showing a predisposition of UBZ1+2 for linear linkages. Furthermore, the avidity for tetraubiquitin linkages increases as compared with diubiquitin ones, and the differences between the K_{ds} determined for linear, Lys63- and Lys48-linked chains are less significant. However, UBZ1+2 still shows a slight preference to bind linear tetraubiquitin chains (7.3 μ M), while the avidity for Lys63-polyubiquitin chains (14.8 μ M) is a bit higher than for Lys48-linkages (31.6 μ M). This selectivity for linear polyubiquitin chains is not unexpected, as it can be predicted based on the structural model of the complex between UBZ1+2 and ubiquitin that we have determined (see section 5.3.5). Indeed, linear linkages implicate binding to the same face of UBZ1+2, which is in agreement with the fact that the face A of UBZ1+2 is more affected upon binding with ubiquitin (mapped by chemical shift perturbations, section 5.3.3.2), with the well defined interaction of the UBZ2 domain with ubiquitin (through the face A of the tandem) and also with the probability density map around UBZ1 (section 5.3.5.2) slightly more populated on face A (60%) than on face B (40%).

The binding of Lys63-tetraubiquitin chains is tighter than the binding of Lys63-diubiquitin chains. This might be related to the structural disposition of UBZ1+2, together with the length of the chains, longer for the tetraubiquitin linkage, which gives place to a higher degree of flexibility for possible readjustment. However, the binding to linear diubiquitin and linear tetraubiquitin do almost not differ, due to the structural predisposition of the system for this type of linkages. Another interesting observation is that the binding affinity for tetraubiquitin chains is always tighter than for monoubiquitin. This could be explained by the tighter binding of UBZ2 to one ubiquitin that directly enhances the binding of another ubiquitin molecule of the chain to UBZ1.

Together we can conclude that the tandem UBZ1+2 binds both mono- and polyubiquitin chains, having special avidity for linear polyubiquitin chains. We propose that the cooperation between UBZ1 and UBZ2 domains within the tandem is not only crucial but also beneficial to the binding of TAX1BP1 to polyubiquitin chains, necessary for its function within the NF- κ B signaling pathway.

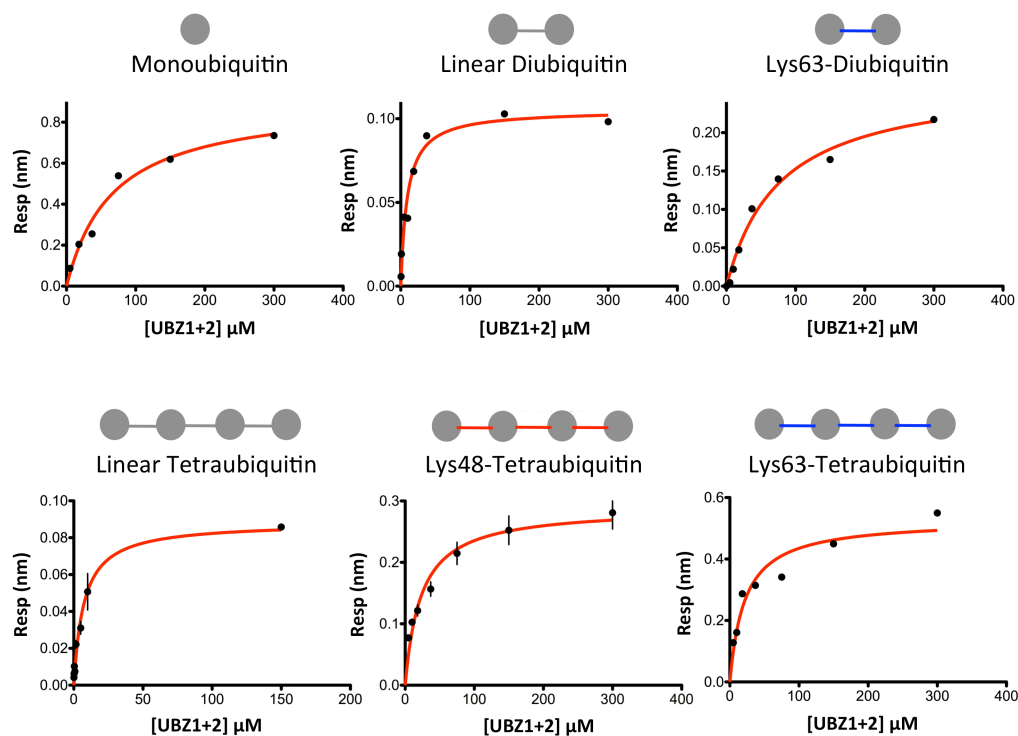


Figure V.40. Steady-state analysis of linkage-specific polyubiquitin recognition by UBZ1+2 to determine overall equilibrium dissociation constants (response at equilibrium plotted against analyte concentration). Experiments were done in 50 mM Tris buffer at pH 7.5, 150 mM NaCl, 0.5 mM DTT and 0.005% (v/v) tween-20, on an OctetRed 96 system (Fortebio) at 25 °C. Above the graphics is shown a schematic representation of monoubiquitin (one grey sphere) and different polyubiquitin chains: diubiquitin (two grey spheres) and tetraubiquitin (four grey spheres), where the grey spheres are connected by linkers with different colors: grey for ubiquitin molecules linked through Met1-Gly76, blue for linkage through Lys63-Gly76 and red for linkage through Lys48-Gly76. The reported values are averaged over at least two experiments and the errors are calculated as the standard deviations from the mean values.

	Monoubi quitin	Linear Diubiquitin	Lys63 Diubiquitin	Linear Tetraubiquitin	Lys48 Tetraubiquitin	Lys63 Tetraubiquitin
K_d (μ M)	67.0 ± 16	9.3 ± 1.8	74.3 ± 13.6	7.3 ± 1.2	31.6 ± 4.0	14.8 ± 1.9

Table V.5. Apparent dissociation constants (K_d s) obtained from the steady-state analysis of the binding data acquired for UBZ1+2 to mono- and linkage-specific polyubiquitin chains using BLI. Uncertainties are the errors of the fit of the experimental data (Fig. V.40).

5.4 Concluding remarks

We have determined for the first time the 3D solution structure of the C-terminal UBZ1 and UBZ2 domains of TAX1BP1 in tandem by NMR. UBZ1+2 is well structured, except for residues located at the flexible termini, with mean pairwise r.m.s deviations of 0.87 and 1.19 Å for the backbone and heavy atoms of residues 8 - 62 (corresponding to the residues 727 - 782 in TAX1BP1 sequence), respectively. The structures of the UBZ1 and UBZ2 domains contain all the features that are typically found in classical Cys₂His₂ zinc finger domains, where the α -helix and two antiparallel β -sheets are packed against each other by tetrahedral coordination of a Zn²⁺ atom.

The solution structure has served to study the interaction with one of the principal targets of TAX1BP1, ubiquitin. The presence of two ubiquitin-binding sites within TAX1BP1 affords opportunities for multiple simultaneous interactions that depend on the intrinsic affinities of each domain and the energetic communication between them, which significantly enhance the efficiency of ubiquitin-binding by cooperation between them, compared with a single domain.

The main characteristics of UBZ1+2 and its interaction with ubiquitin are discussed below.

5.4.1 The role of the Zn²⁺ atom in the structure of the UBZ domains

We have demonstrated by NMR and CD that the coordination of the Zn²⁺ atom is essential to maintain the correct 3D structure of the UBZ1+2 domains in solution. Zn²⁺ is one of the most abundant transition metal in living organisms, however its intracellular concentration is kept very low, in the femto molar range (Suhy D.A. et al., 1999), especially by some zinc fingers proteins known as metallothioneins (Maret W., 1994 and 1995) (Chapter I, Table I.2) that can bind more than seven Zn²⁺ atoms in their two ligand binding domains.

In proteins, the Zn²⁺ binding site can play a catalytic (catalytic Zn²⁺-sites) or a structural (structural Zn²⁺-sites) role (Dudev T. and Lim C., 2003). The family of zinc fingers domains is included in the structural Zn²⁺-site group. The principal difference between these two groups is that structural Zn²⁺-sites used to be fully or partially buried and surrounded by a network of hydrogen bonds (Maynard A.T. and Covell D.G., 2001), like in the case of the UBZ domains of TAX1BP1, where a hydrophobic core enhances the tetrahedral coordination of the Zn²⁺ atom.

Zn²⁺ presents an ionic radius of 0.72 Å and in structural Zn²⁺-sites within proteins, used to adopt a tetrahedral arrangement that is more stable than the other zinc polyhedra (octahedral and penta-coordination) (Dudev T. and Lim C., 2003). On average, the Zn²⁺-ligand distances for tetrahedral coordination with Cys and His are Zn²⁺-N (His): 2.07 - 2.09 Å and Zn²⁺-S (Cys): 2.21 - 2.35 Å (Ablerts I.L. et al., 1998), which fits perfectly with the distances measured in the UBZ1+2 NMR structure, 2.0 and 2.3 Å respectively. In this sense, the short length of the bonds between metal and ligand in the tetrahedral Zn²⁺-binding sites is well suited for stabilizing a given protein fold (Dudev T. and Lim C., 2003).

Many studies have been performed on the determination of the affinity of the Zn^{2+} atom that coordinates Cys and His side chains in zinc finger domains (Krizek B.A. et al., 1993; Hartwig A., 1998; Hartwig A., 2001; Razmiafshari M. et al., 2001). The association constants for different types of zinc finger domains (Cys₂His₂, Cys₂HisCys, Cys₄, and so on) with Zn^{2+} have been measured by spectroscopic titration techniques, giving values in the order of $10^8 - 10^{11} M^{-1}$ for the classical Cys₂His₂ zinc fingers.

Due to its strong affinity it is not easy to displace the Zn^{2+} atom from the Zn^{2+} -finger core. However several studies (Predki P.F. and Sarkar B., 1992; Hartwig A., 2001; Hartwig A. et al., 2002) have considered the existence of a possible poisoning mechanism in cells by heavy metals that can eject the Zn^{2+} from the binding site leading to the inactivation of the zinc finger, since the heavy metal cannot maintain the proper conformation of the protein. In general, classical zinc fingers (like UBZ domains) bind Zn^{2+} tighter than heavier metals like Cd^{2+} , Co^{2+} or Pb^{2+} , and appear to be protected against them (Dudev T. and Lim C., 2003).

All these findings are in agreement with the behavior that we have observed for the UBZ1 and UBZ2 domains of TAX1BP1, where Zn^{2+} plays a crucial role in the maintenance of the domains folding (section 5.3.1.4). In general, the improper folding of the zinc finger domains is associated with the loss of their function (Hartwig A., 2001). This is visualized in the HSQC spectra in figure V.41, where after the formation of the complex with ubiquitin by titrating ubiquitin with UBZ2 (trajectory of the chemical shifts perturbations indicated by a pink arrow), addition of EDTA in small steps till exceeding the equimolar ratio of the UBZ domain, results in a reversion of the ubiquitin resonances trajectories back to the original point (Fig. V.41, blue arrow) corresponding to the free ubiquitin. This effect is caused by the strong chelating effect of EDTA (affinity for Zn^{2+} in the order of $10^{16} M^{-1}$). Thus EDTA is able to remove the coordinated Zn^{2+} atom from the core of the UBZ domain causing the loss of its stable 3D structure, thus losing its ubiquitin binding-competent conformation.

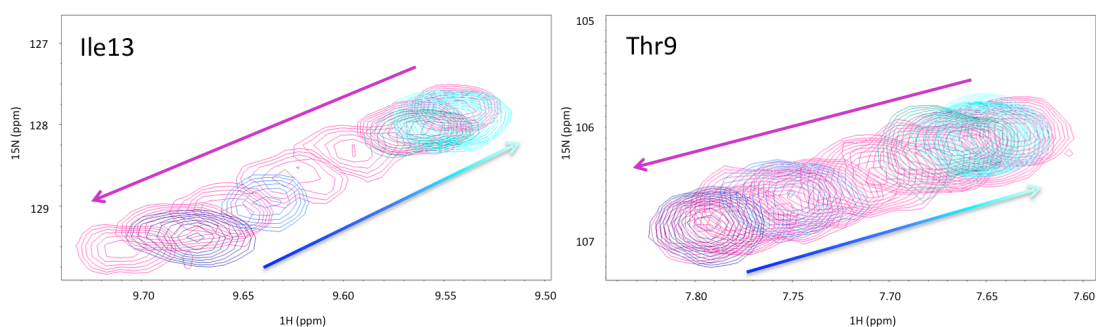


Figure V.41. Regions of the 1H - ^{15}N -HSQC spectra corresponding to the ubiquitin residues Ile13 and Thr9. Pink arrows represent the directionality of the chemical shift perturbation of the ^{15}N -labeled ubiquitin residues during the titration with increasing amounts of unlabeled UBZ2 domain. Blue arrows (from dark blue to light blue) represent the directionality of the titration with increasing amounts of EDTA till a ratio 1:2 (UBZ2:EDTA).

5.4.2 UBZ1 and UBZ2 domains of TAX1BP1 bind to ubiquitin

We carried out the interaction study between ubiquitin and the UBZ1 and UBZ2 domains of TAX1BP1 in tandem but also as isolated domains, by ITC and NMR. Both techniques confirm that the binding affinity of the tandem UBZ1+2 is increased in comparison with the affinity of the isolated domains. This points to the existence of a cooperative mechanism between both UBZ domains, as observed for other tandems where the zinc fingers are connected through short linkers, such as UIM1+2 of AT3 (Song A.X. et al., 2010). The presence of longer linkers between the domains abolishes the cooperative interaction (Fujiwara K. et al., 2004; Wang Q. et al., 2005).

The present work has been crucial to determine that UBZ1 domain of TAX1BP1 is also able to interact with ubiquitin but with lower affinity than UBZ2. The interaction of the UBZ2 domain with ubiquitin was previously detected by GST-pull down experiments (Iha H. et al., 2008), however, nothing was known about the binding affinity of this interaction, the residues involved, and the role of UBZ1 in ubiquitin-binding. The use of NMR has allowed us to map the binding interface of UBZ1+2 and ubiquitin and to dissect the communication within the tandem through the hydrophobic cores of both domains that leads to a conformationally driven cooperativity effect.

The titration of ¹⁵N-labeled ubiquitin with isolated UBZ1 and UBZ2, and the tandem UBZ1+2, has permitted us to directly compare the chemical shift changes of ubiquitin residues during the different titrations. The size of the chemical shift changes has been suggested to correlate with the dynamics of the protein complexes (Worrall J.A. et al., 2002; Volkov A.N. et al., 2005). Small chemical shift perturbations (as observed in the interaction of the isolated UBZ1 with ubiquitin) are indicative of dynamic and weak interactions, while large binding shifts (like those for the interaction of the isolated UBZ2 domain with ubiquitin) are associated with predominantly single orientation complexes (Prudêncio M. and Ubbink M., 2004), totally in agreement with the different affinities of both domains, their thermodynamic signature and the UBZ1+2-ubiquitin complex established by PRE NMR.

5.4.2.1 UBZ1+2 recognition area

Comparison of the binding interfaces of UBZ1 and UBZ2 domains with other UBZ domains involved in ubiquitin is limited to two proteins, the DNA pol eta and WRNIP1. The UBZ domain of DNA pol eta is the only zinc finger with Cys₂His₂ coordination for which the interaction with ubiquitin has been reported, including a model of the complex formed with ubiquitin (Bienko M. et al., 2005; Bomar M.G. et al., 2007). The structure of the ubiquitin-bound UBZ domain of WRNIP1, which corresponds to a variation of the classical UBZ domains presenting a Cys₂HisCys motif, has been determined by X-ray (PDB code 3VHT) (Suzuki N. et al., 2010). The UBZ domain of WRNIP1 interacts with the hydrophobic patch of ubiquitin mainly through the complete face A of the α -helix and the first β -sheet (Fig. V.42A), while the binding of UBZ of DNA pol eta to ubiquitin (Fig. V.42B) is centered exclusively in the C-terminal part of the α -helix.

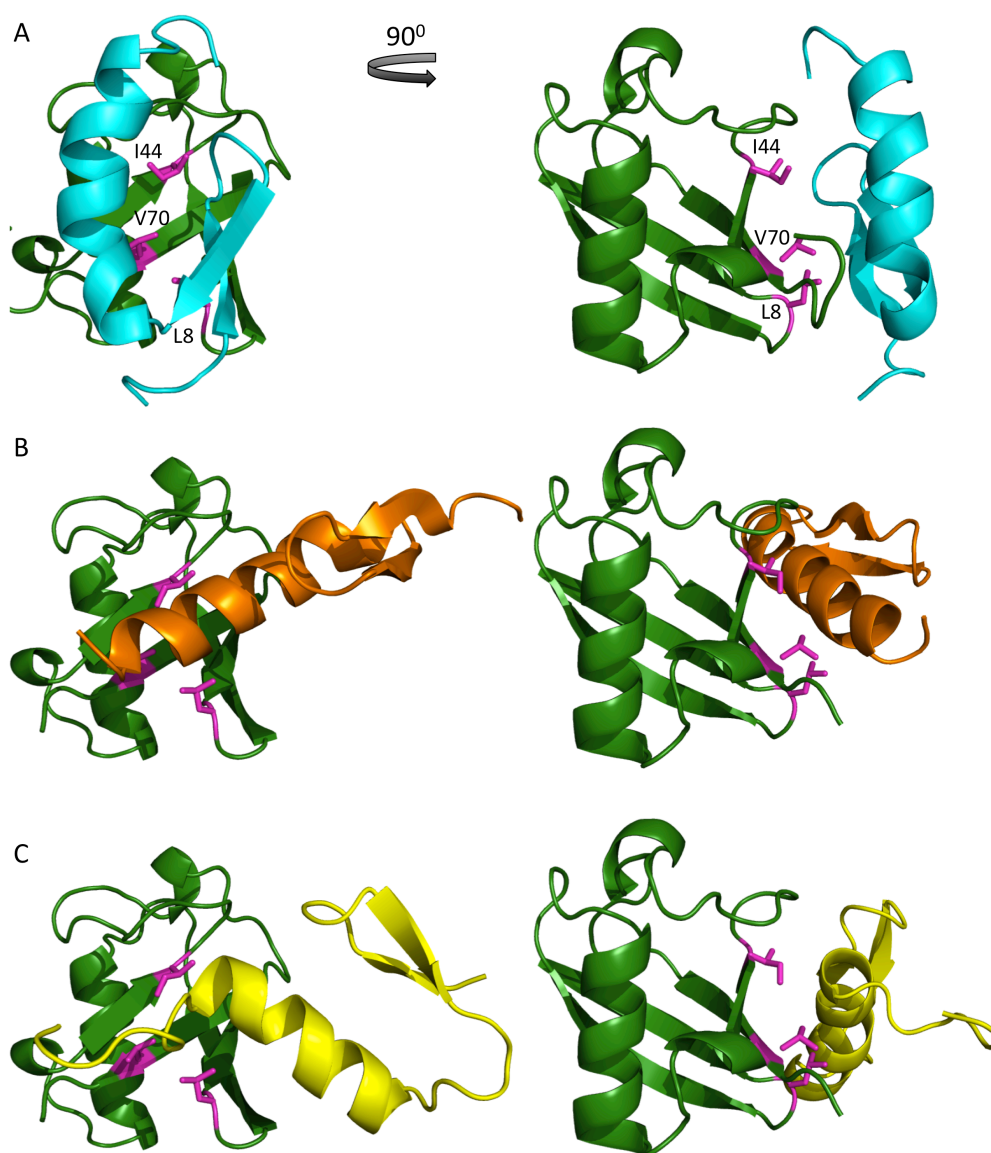


Figure V.42. Differences in ubiquitin binding within the members of the UBZ family. **A.** X-ray structure of UBZ domain of WRNPI (cyan) in complex with ubiquitin (PDB code 3VHT). **B.** Model of the UBZ domain of DNA pol eta (orange) in complex with ubiquitin (Bomar M.G. et al., 2007). **C.** UBZ2 domain of TAX1BP1 (yellow) in complex with ubiquitin determined in this thesis by PRE NMR. In all cases ubiquitin is colored in green, and residues Leu8, Ile44 and Val70 (representing the hydrophobic patch of ubiquitin) are shown as pink sticks.

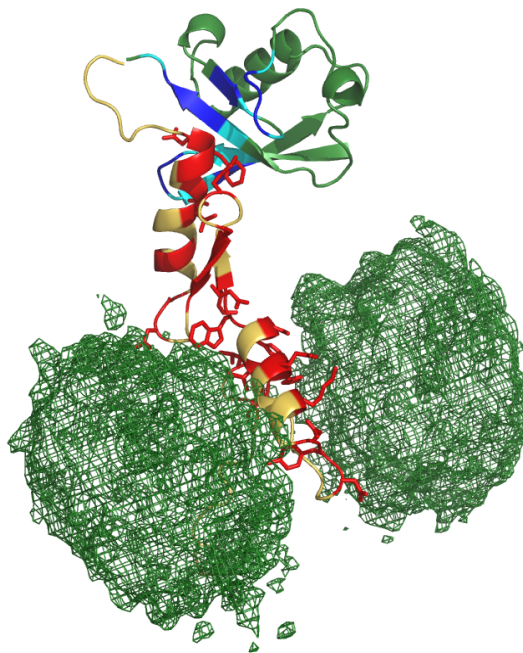


Figure V.43. Complex of UBZ1+2 (yellow) and ubiquitin (green) determined by PRE NMR. Side chains of the most affected residues upon ubiquitin binding (based on chemical shift perturbations study, section 5.3.3.2) are represented in sticks and colored in red to highlight their directionality toward face A and face B. Residues of ubiquitin affected upon binding with UBZ1+2 are colored in dark blue (more affected) and cyan (less affected), based on the chemical shift perturbations (Fig. V.20).

Figure V.43 shows the model of the UBZ1+2-ubiquitin complex that has been determined in this thesis by the use of PRE NMR. This model fits well with the binding interfaces of UBZ1+2 and ubiquitin mapped by chemical shift perturbations. The agreement between these two independent measurements has served to validate the correctness of the complex. In this complex, UBZ2 interacts with ubiquitin in a similar (but not identical) way than the UBZ domain of DNA pol eta (Fig. V.42B and C). UBZ of DNA pol eta binds ubiquitin exclusively through the C-terminal part of the α -helix, while the binding interface of TAX1BP1-UBZ2 includes the C-terminal part of the α -helix and also the Met39 located in the loop that connects the two β -sheet of UBZ2. Moreover, the relative orientation towards ubiquitin of the interacting helix in the diverse UBZ domains is substantially different (Fig. V.42). On the other hand, UBZ1 interacts with ubiquitin in a more dynamic way, totally different from the other UBZ domains, and characterized by a fractional occupancy at both sides of the domain (Fig. V.43).

We can conclude that UBZ1 and UBZ2 of TAX1BP1, UBZ of WRNIP1 and UBZ of DNA pol eta display different ubiquitin-binding modes, even though all of them recognize the canonical hydrophobic patch on the β -sheet of ubiquitin (Fig. V.42).

Different groups have tried to establish, by comparison, a canonical recognition sequence based on the interaction of the UBZ domain of DNA pol eta with ubiquitin, that is localized in the α -helix at the C-terminus of the UBZ domain (Bomar M.G. et al., 2007). This comparison (Fig. V.44A) included other ubiquitin-binding domains that

also bind ubiquitin through their helix, like the UIM (ubiquitin interacting motif, Hofmann K. and Falquet L., 2001; Swanson K.A. et al., 2003), the MIU/IUIM (motif interacting with ubiquitin/inversed ubiquitin interacting motif) of Rabex-5 (the Rab5 exchange factor, Penengo L. et al., 2006) and the zinc finger of NEMO (Cordier F et al., 2008) that is also folded forming a $\beta\beta\alpha$ -unit.

By sequence alignment of the α -helices some conserved residues at positions -7,-4,-3,-1 and +3 and especially the invariant alanine (A656 for UBZ of DNA pol eta and a Met for NEMO ZnF) in the center of the consensus sequence (due to its importance it occupies the position zero) were identified (Cordier F. et al., 2008). This Ala has been reported to be crucial for the interaction with the hydrophobic pocket of ubiquitin (Leu8-Ile44-Val70), since mutation of this residue totally abolishes ubiquitin binding (Bomar M.G. et al., 2007).

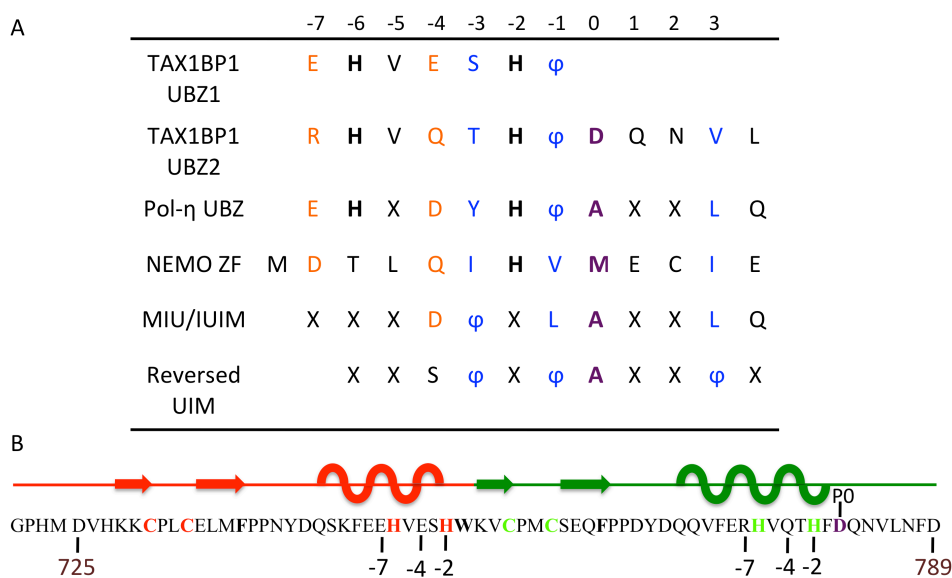


Figure V.44. A. Alignment of the signature motif present in the α -helices of DNA pol eta UBZ domain with the sequences of NEMO ZnF, MIU/IUIM of Rabex-5 and the reversed UIM, and also with the sequence of UBZ1 and UBZ2 of TAX1BP1. On top of the table appears the numbering corresponding to the consensus sequence (Bomar M.G. et al., 2007). The central invariant Ala in purple (referred to as position 0) is replaced by a Met in NEMO ZF (Cordier F et al., 2008). Conserved hydrophobic residues are in blue and acidic residues are in orange. ϕ denotes large hydrophobic (Trp in UBZ1, Phe in UBZ2) residues, and X any residues. This table is adapted from Bomar M.G, et al., 2007, by including the UBZ1 and UBZ2 domains of TAX1BP1. **B.** Schematic representation of the secondary structure of the UBZ1 (in red) and UBZ2 (in green) domains. The numbers at the beginning and at the end correspond to the amino acid sequence in TAX1BP1. The number -7, -4 and -2 correspond to the numbering of the consensus sequence showed above the table, and P0 is the position of the conserved Ala, that in the case of the UBZ1 domain does not exist and in UBZ2 is occupied by an Asp. This aspartate residue shows significant chemical shift perturbations (Fig. V.23), which fit with the complex determined by PRE NMR, where Asp62 (negatively charged) could interact with Lys48 and Arg42 of ubiquitin (positively charged).

Both UBZ domains of TAX1BP1 do not show important similarities with the consensus sequence of the aligned α -helices of other UBD domains (Fig. V.44), for instance lacking the conserved Ala (or Met) in their sequence. This is not an exclusive feature of UBZ1+2 of the human TAX1BP1, because sequence alignment shows that the UBZ domains of TAX1BP1 are very well conserved among different organisms (Fig. V.45), including the high sequence homology between UBZ1 and UBZ2 themselves, with the main difference that the α -helix is longer in the UBZ2 domain.

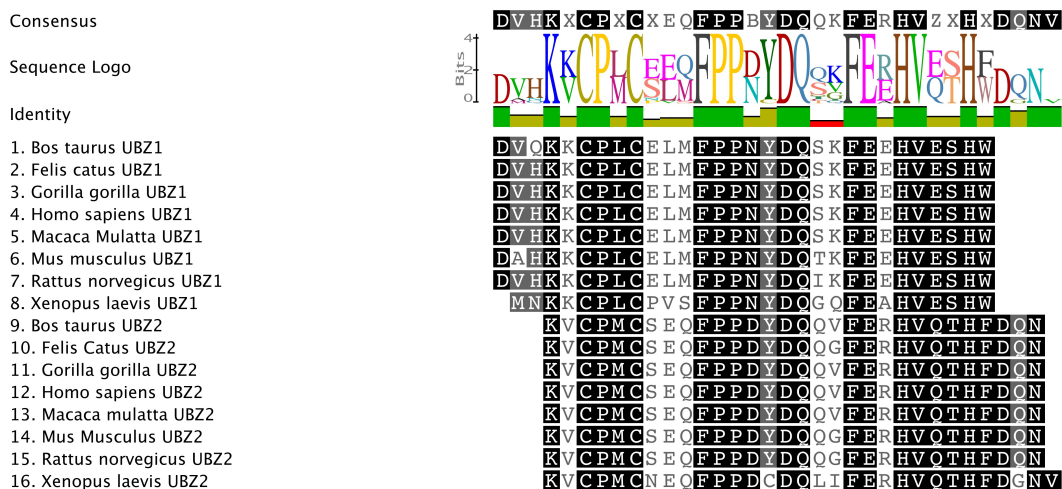


Figure V.45. Sequence alignment of the UBZ1 and UBZ2 domains of TAX1BP1 in different organisms. The search of the sequences was done with the online tools of BLAST and the sequential alignment was carried out with the program Genious 6.1.4 (Kearse M. et al., 2012).

Further comparison of the α -helix sequences of the UBZ1 and UBZ2 domains with the consensus sequence based on DNA pol eta, indicate that the residues located in the positions -7 and -4 are conserved (Fig. V.44). It is important to notice that the binding interface of UBZ2 also involves the C-terminal part of the α -helix despite the lack of the key alanine. Interestingly, the isolated TAX1BP1-UBZ2 domain binds ubiquitin tighter (with a K_d of 36 μ M) than the UBZ domain of DNA pol eta (with a K_d of 73 μ M). On the other hand, TAX1BP1-UBZ1, which α -helix is shorter lacking any residue at key position 0, is also able to bind ubiquitin with an order of magnitude less affinity (K_d of 150 μ M within UBZ1+2 and 360 μ M when isolated).

We can conclude that even though the consensus sequence is not present in the UBZ domains of TAX1BP1, both UBZ1 and UBZ2 are able to interact with ubiquitin in a way that was not described till now.

5.4.2.2 Ubiquitin recognition area

The results of this work clearly indicate that both UBZ domains of TAX1BP1 interact with the hydrophobic patch of ubiquitin (Leu8-Ile44-Val70). In general, most of the ubiquitin-binding domains are α -helical: UBA (ubiquitin associated), UIM (ubiquitin interacting motif), DUIM (double-sided UIM), MIU (motif interacting with ubiquitin), CUE (coupling of ubiquitin conjugation to endoplasmic reticulum degradation), GAT (GGA and TOM) (see Chapter I, section 1.5). And all the helical ubiquitin-binding

domains are known to interact with the hydrophobic patch of ubiquitin (Leu8-Ile44-Val70). For example, the MIU domains present in the region 48-74 of Rabex-5 (Lee S. et al., 2006; Penengo L. et al., 2006) or the GAT domains that appear to bind to the hydrophobic patch through two different sites (Bilodeau P.S. et al., 2004). However, some zinc finger domains can recognize other surfaces on ubiquitin. This is the case of the A20 ZnF domains, that bind predominantly the polar patch of ubiquitin centered around Asp58, as well as a TEK-box motif, consisting of a short region around the ⁹TGK¹¹ sequence in ubiquitin (Bosanac I. et al., 2010). Another binding site of ubiquitin is the diglycine patch including the C-terminal Gly76, which is recognized by the ZnF UBP domain (ubiquitin-binding module found in various deubiquitinating enzymes, the ubiquitin ligase IMP/BRAP2 and the microtubule deacetylase HDAC6) (Reyes-Turcu F.E. et al., 2006).

It is known that the interactions between UBDs and monoubiquitin are predominantly weak, and span a wide range of affinities, with K_d ranging from 100 μ M to 500 μ M (Hurley J.H. et al., 2006). The modular domains of adaptor proteins that bind ubiquitinated membrane proteins, can bind monoubiquitin with affinities around 100 μ M or lower (Hurley J.H. et al., 2006), which is the case of the UBZ1 and UBZ2 domains of TAX1BP1 adaptor protein. Weak interactions of ubiquitin with UBDs are physiologically relevant because these domains are generally involved in transitory protein networks (Hicke L. et al., 2005).

5.4.3 Electrostatic and hydrophobic forces involved in binding of ubiquitin to UBZ1+2 of TAX1BP1

Electrostatic interactions play an important role in determining protein-binding specificity (Kukic P. and Nielsen J.E., 2010). In order to characterize if the binding interfaces of UBZ1+2 domains and ubiquitin are electrostatically predisposed to interact, we have calculated the distribution of the electrostatic potential on the surface of both proteins.

Figure V.46 shows that the UBZ1+2 binding surface is mainly negatively charged, except for a small region at the beginning of the α -helix of UBZ1 that is positively charged. However, the binding interface of ubiquitin (canonical patch) is positively charged (Fig. V.46B). The distribution of electrostatic potential suggests that the ionic interactions are playing an important role in the recognition mechanism of ubiquitin and UBZ1+2.

Consideration of intermolecular electrostatic forces alone is often not enough for the correct prediction of the binding geometry in transient protein-protein interactions (Volkov A.N., Ph.D. thesis, 2007). Complex stability is strongly influenced by other factors than electrostatics, like hydrophobic forces, mainly due to the presence of hydrophobic residues in the binding interfaces as in the case of ubiquitin (hydrophobic Leu8-Ile44-Val70 patch) and UBZ1+2, which show significant chemical shift perturbations. In this sense, it is expected that the hydrophobic interactions are playing a short-distance interaction role, especially because of the entropy gained by removal of hydrophobic surface area from ordered solvating water, which provides most of the binding affinity in a protein complex (Sturtevant J.M., 1977). Other intermolecular contacts, such as van der Waals, dipole-dipole, salt bridges, etc., further stabilize the complex and fine-tune the specificity.

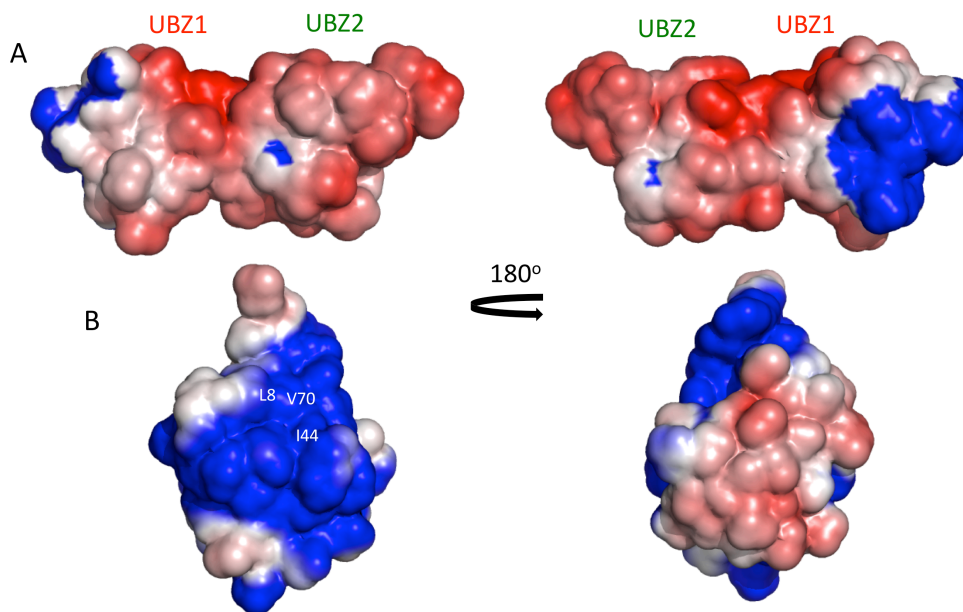


Figure V.46. ± 5 KT/e electrostatic potential of UBZ1+2 (**A**) and ubiquitin (**B**) in PyMOL plotted on the solvent accessible surface. Calculations were done using the Python software PDB2PQR and visualization was done using the computational packages of APBS (run in PyMOL) using the Poisson-Boltzmann electrostatics calculation (Dolinsky T.J. et al., 2004) in conjunction with the solvent-accessible surface area (Kukic P. and Nielsen J.E., 2010). Color code: grey for non-polar residues, blue and red for positively and negatively charged residues, respectively. Positions of the UBZ1 and UBZ2 domains are indicated with labels (**A**), and Leu8, Ile44 and Val70 are shown on the surface of ubiquitin (**B**).

5.4.4 Promiscuous interactions of UBZ1+2 with polyubiquitin chains

BLI experiments show that UBZ1+2 is a promiscuous binding partner for ubiquitin chains, as it is able to interact with Lys63- and Lys48-linked chains, but also with linear chains in a tighter way. These results fit perfectly with the structural model proposed by the use of PRE NMR, where the tight binding of UBZ2 and the weak and dynamic binding of UBZ1 with ubiquitin enable UBZ1+2 to recognize different polyubiquitin linkages. Until now, none UBZ tandem, and only a few UBD tandems have been characterized in its interaction with polyubiquitin chains (Shin D. et al., 2012): the CARD domains of RIG-I recognize Lys63-linkages (Zeng W. et al., 2010), the UIM domains of Ataxin-3 interact with Lys63- and Lys48-linkages (Song A.X. et al., 2010), the UIM domains of RAP80 selectively bind linear and Lys63-linked polyubiquitin (Sato Y. et al., 2009; Cho H.J. et al., 2009; Sekiyama N. et al., 2012) and the tandem of UBDs domains of RABEX-5 recognizes linear, Lys63- and Lys48-linked polyubiquitin chains with a slight preference for linear and Lys63-linked ones (Shin D. et al., 2012). Thus, linkage specific ubiquitin recognition is crucial for the diverse set of functional outcomes associated with ubiquitination.

We can conclude that the work presented in this chapter constitutes the first structural and thermodynamic characterization of the interaction between UBZ1+2 of TAX1BP1 and ubiquitin, providing insight in the signaling pathways in which TAX1BP1 is involved. The NMR assignment and determination of the 3D structure of UBZ1+2 in solution establish the basis for future studies oriented to map the interaction of these domains with different partners like the WW domains of Itch. Upcoming studies focussed on characterizing the regulation of ubiquitin binding by UBZ domains in the context of full-length TAX1BP1 will enrich our understanding of the immune system and inflammatory pathways through a thorough characterization of the innate response from a structural point of view. This may ultimately lead to the design of specific drugs that might be used in treatments of inflammatory diseases.

CHAPTER VI

GENERAL DISCUSSION

Protein-protein recognition... is not mere binding but is, in principle, binding with purpose. This purpose is to discriminate between the reaction partners and other proteins in the cellular milieu. Peter Crowley.

Chapter VI. GENERAL DISCUSSION

Adaptor proteins are key players in signal transduction pathways within the cell. Often they are composed by multiple modular domains of the same kind, as the three SH3 domains located at the N-terminus of the CIN85/CD2AP family and the two UBZ domains at the C-terminal extreme of TAX1BP1. These modular domains are involved in transient interactions with different targets, characterized by moderate to weak affinities. Organization of these versatile modular domains in complexes and clusters improves the efficiency and fidelity of the cellular signaling, regulating diverse biological processes such as immunity, inflammation, endocytosis of cellular receptors and apoptosis.

The identification of the binding interfaces of these modular domains and the determination of the versatile range of interactions in which they are involved, are crucial for the assembly of protein networks and the identification of new targets with biomedical purpose. However, detailed understanding of protein interactions is not an easy aim as the use of a single biophysical technique often does not provide the right and/or sufficient information on the macromolecular complex under study. The work described in this thesis has highlighted the necessity to bridge atomic resolution structures determined by NMR or X-ray, with other techniques that provide information about transient interactions or conformational states that are too flexible, or too difficult to stabilize. Techniques like Isothermal Titration Calorimetry (ITC), Small Angle X-ray Scattering (SAXS) and Paramagnetic Relaxation Enhancement (PRE) NMR, have been extremely important for this purpose. It is important to take into account that conformational changes associated with binding events occur on “slow” (microsecond to millisecond) time scales, however, “fast” (pico to nanosecond) protein motions play important roles in all aspects of the binding event. NMR is a technique that can cover this wide range of time scales (Fig. VI.1). Furthermore, to understand the forces driving the interaction between proteins, a detailed description of the binding thermodynamics is of high importance, because the recognition event might be associated with changes in both the structure and dynamics of each counterpart.

The combined use of ITC and NMR allows the correlation between thermodynamics and structure, and has been very useful to understand the interaction mechanisms of UBZ and SH3 domains with their natural targets. Heteronuclear NMR is able to measure the occupancies of individual binding sites on proteins and therefore determining apparent binding affinities for every residue. Coupling this residue-specific data (e.g. chemical shift mapping and/or relaxation analysis) with the macroscopic binding data from ITC has allowed a complete description of the binding properties of the systems studied in this thesis.

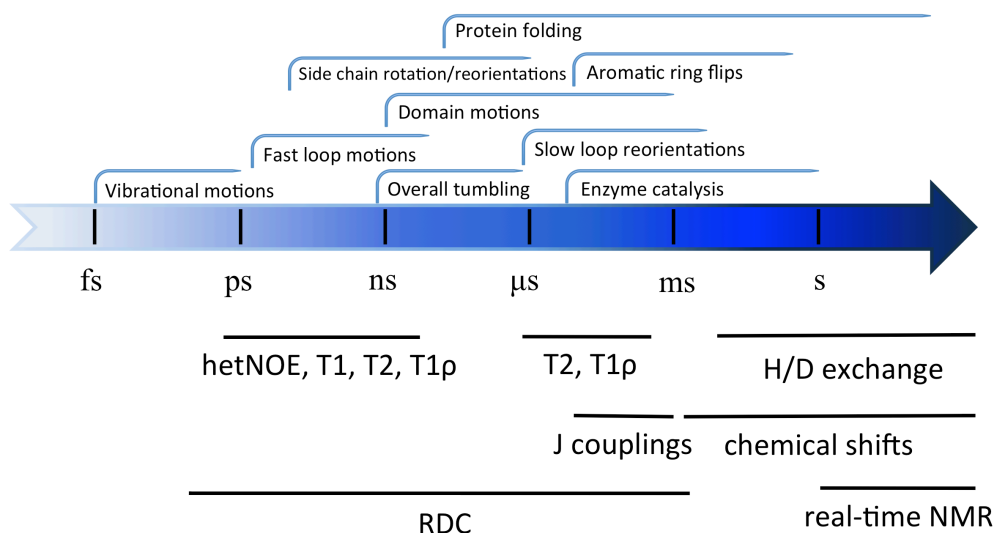


Figure VI.1. Motions detected by NMR measurements. Adapted from Zwier M.C. and Chong L.T., 2010.

One of the objectives of this thesis aimed to understand the versatile interactions involving the N-terminal SH3 (SH3A) domain of the CIN85/CD2AP family of adaptor proteins with atypical proline-rich targets, establishing the molecular basis of the clustering in which these modular domains are involved, crucial for their biological role inside the cell (Chapter IV).

The interaction between the SH3A domains of CIN85 and CD2AP with atypical proline-rich targets displays an enthalpy driven thermodynamic signature, a common feature for the interaction of SH3 domains with proline-rich ligands (Wittekind M. et al., 1994; Arold S. et al., 1998; Ferreón J.C. and Hilser V.J., 2004), and implies that both electrostatic and hydrophobic interactions play a crucial role in the assembly of the complex. ITC has been very useful to determine the oligomeric state of the complex, showing a 1:1 stoichiometry for the interaction of CD2AP-SH3A with Cbl-b and CD2 and for the interaction of CIN85-SH3A with CD2, and a 1:2 stoichiometry for the interaction of CIN85-SH3A with Cbl-b. Moreover, NMR titrations, used to map the interaction between the SH3A domains with CD2 and Cbl-b, confirm the moderate to weak interactions observed by ITC, another characteristic of SH3 domains (Nguyen J.T. et al., 1998). Weak interactions are biologically relevant because they are required for the formation of networks that can undergo rapid assembly and disassembly.

The sensitivity of NMR titrations to detect conformational rearrangements upon binding Cbl-b has allowed the observation of multiple binding events indicated by the unusual curvature in the chemical shift perturbations of the Trp indole (Trp36 in CIN85-SH3A and Trp37 in CD2AP-SH3A) resonances. This Trp is conserved in all SH3 domains as it plays a crucial role in binding polyproline stretches (Fernandez-Ballester G. et al., 2004), serving as an excellent probe for monitoring changes that take place in the binding interface. The mutation of the two key arginines flanking the PxxxP atypical-proline sequence present in Cbl-b (Cbl-bR904G and Cbl-bR911A), have permitted to decompose the curvature in two linear shift patterns. The upfield shift change corresponds to the formation of type I dimers (driven by Arg904) and the downfield shift change indicative for the formation of type II dimers (driven by the Arg911).

These experimental NMR data, together with the stoichiometry obtained by ITC and with the *ab initio* reconstructed shapes determined by SAXS have shown that CD2AP-SH3A forms hetero-dimers type I and type II with Cbl-b in solution, while in the case of CIN85-SH3A, multimeric states in solution are formed, including a hetero-trimer where type I and type II interactions take place at the same time. The presence of the N-terminal arginine in Cbl-b (Arg904) plays a key role, since its absence only results in the formation of type II dimers as it is the case for the interaction of CD2 with both CIN85- and CD2AP-SH3A. It is important to stress that the presence of the N-terminal arginine in the atypical proline-rich sequence does not define the multimerization state of the SH3:peptide complex, but only the ability to interact in a type I orientation.

The results of this work have pointed out the need to extend the atypical proline-rich sequence PxxxPR to RxPxxxPR. In this way, for peptides containing the N-terminal arginine in their sequences, type I interaction is expected and for those ones containing the C-terminal arginine, type II; for proteins containing both arginines, both type I and type II interactions are possible. The PxxxPR sequence is presented in some of the binding partners of the CIN85/CD2AP family of adaptor proteins (Kurakin A.V. et al., 2003; Kowanetz K. et al., 2003) like TAF_{II}70, Pak-2, BLNK, CD2, Cbl, Alix and ZO-2. Sequences alignment (Fig. VI.2) shows that type II interaction is conserved among the natural targets, while the extended RxPxxxPR sequence is only present in Cbl-b and c-Cbl, but also in two regions of ZO-2 (Zonula occlude protein-2, ZO-2/291-303 and ZO-2/961-973). Moreover, we have shown by ITC that the type II interaction is stronger reflected by the higher affinities (K_d for type II 1-3 μ M, for type I around 50 μ M).

Cbl-b (903-913)	A	R	P	P	K	P	R	P	R	R	T
C-Cbl (836-846)	E	R	P	P	K	P	F	P	R	R	I
CD2 (324-334)	K	G	P	P	L	P	R	P	R	V	Q
TAF _{II} 70 (494-404)	S	Q	A	P	Q	P	G	P	R	T	P
Pak-2 (184-194)	E	P	P	P	V	I	A	P	R	P	E
BLNK (239-249)	P	A	A	P	S	P	L	P	R	A	G
Alix (737-747)	P	T	P	P	T	P	A	P	R	T	M
ZO-2 (292-302)	S	R	S	P	S	P	E	P	R	G	R
ZO-2 (962-972)	I	R	K	P	S	P	E	P	R	A	Q

Figure VI.2. Sequence alignment of some interacting natural partners of CIN85/CD2AP family of adaptor proteins that present the atypical proline-rich sequence PxxxPR (based on the targets reported in Kurakin A.V. et al., 2003 and Kowanetz K. et al., 2003). Between brackets the numbering corresponding with the protein sequence presented in the Table. Conserved prolines are in bold and highlighted by a grey panel, arginine residues responsible for the type II interaction are highlighted in green and the position of the N-terminal arginine responsible for the type I interaction in blue.

As mentioned above, the presence of the N-terminal arginine defines the existence of type I interactions, but the formation or not of a hetero-trimeric complex is the result of the absence or presence, respectively, of steric hindrance, specially due to the

differences within the respective n-Src loops of the SH3A domains (Fig. VI.3). In general, differences in the n-Src loop of SH3 domains have been reported to account for binding selectivity (Mayer B.J., 2001). This differential complex formation might yield different biological responses of the proteins involved.

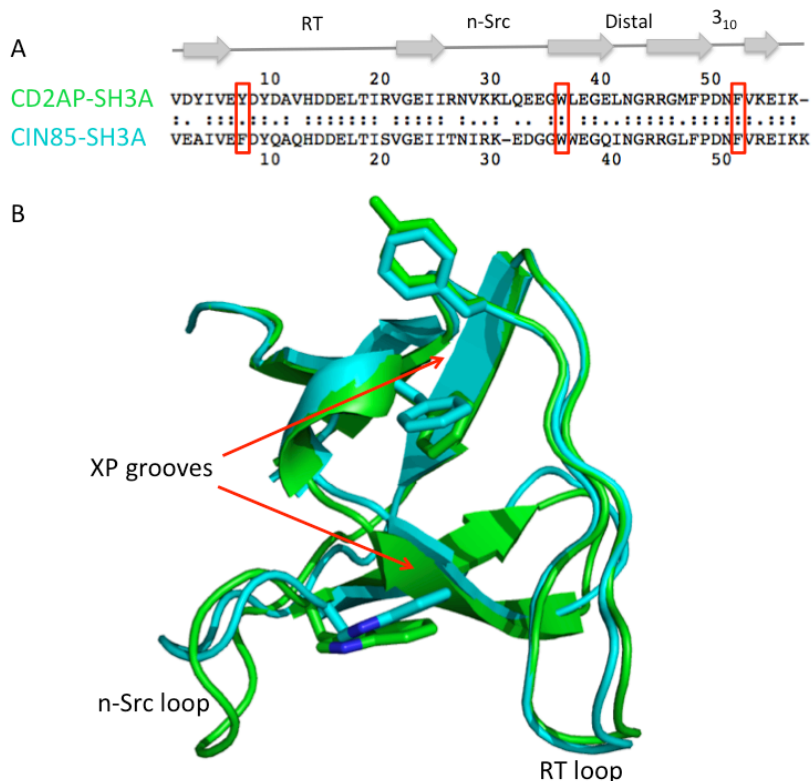


Figure VI.3. Differences in the n-Src loop of the SH3A domains of CIN85 and CD2AP. **A.** Sequence alignment of the SH3A domains; the n-Src loop of CD2AP-SH3A is one residue longer than in CIN85-SH3A. **B.** Superposition of the X-ray structures of the SH3A domains of CD2AP (green, PDB entry 2J6K) and CIN85 (cyan, PDB entry 2BZ8), highlighting the structural differences in the n-Src loops of both proteins. The n-Src and RT loops are indicated, and the residues Y8, W37 and F53 for CD2AP-SH3A and F8, W36 and F52 for CIN85-SH3A, are represented in sticks to indicate the hydrophobic XP-grooves.

Thus the additional type I interaction likely plays an important biological role in the binding of CIN85 and CD2AP with Cbl-b. We know that CIN85/CD2AP is involved in the regulation of Cbl-directed RTK downregulation (Petrelli A. et al., 2002; Soubeyran P. et al., 2002; Haglund K. et al., 2002; Dikic I., 2002) via Cbl/Cbl-b interaction. CIN85 positions endophilins (regulatory components of clathrin-coated vesicles) in proximity of internalizing receptors (Petrelli A. et al., 2002; Soubeyran P. et al., 2002) by binding the atypical proline-rich sequence of Cbl/Cbl-b, after EGF stimulation. But CIN85 also participates in post-internalization processes such as endocytic trafficking and receptor degradation. In this way, CIN85 is involved in the degradation of the EGF receptors via Cbl interaction (Haglund K. et al., 2002) and it has been reported that the mutation of either of the two arginine residues of Cbl (Arg-904 and Arg911) in transfected cells, resulted in a substantial decrease of the degradation of EGF receptors (Jozic D. et al., 2005), indicating that both arginines play an important role in the pathways controlling the EGF receptor downregulation. Furthermore, the results of our work reveal that the interaction mechanism of CD2AP-SH3A and CIN85-SH3A domains with Cbl-b is

different: CD2AP forms mainly hetero-dimers (type I and type II), while CIN85 is able to form hetero-trimers. The finding that the clustering mechanism of CIN85-SH3A is Cbl-b concentration dependent might imply important biological consequences.

The interaction of CIN85-SH3A domain with CD2 has been characterized for the first time, and as in the case of the interaction of CD2AP-SH3A with CD2, leads to the formation of type II hetero-dimers. It is shown that the interaction of CD2AP with CD2 is crucial to mediate the immunological synapse. T-cell activation induces the binding of the SH3A domain of CD2AP with the proline-rich sequence within the last 20 residues of the cytoplasmic domain of CD2, which is necessary for CD2 clustering and T-cell polarization (Dustin M. et al., 1998). These events are critical for the formation of an effective T-cell-antigen presenting cell junction. Based on our results, CIN85 is also able to interact with CD2 in the same way as CD2AP, so CIN85 could compensate CD2AP-deficiency in this pathway.

Recently, a positive role for CIN85 in pre-TCR (T-cell receptor) signaling has been suggested (Kometani K. et al., 2011). The relation of CIN85 with the immune system is not limited to T cells, but it has a physiological function in B cells. In this sense, CIN85 has a positive role in BCR (B-cell receptors)-mediated survival and proliferation by participating in the canonical NF- κ B signaling pathway (Kometani K. et al., 2011). The precise mechanism by which CIN85 activates NF- κ B remains unclear. However, knowing that the SH3 domains of CIN85 can function as ubiquitin-binding domains (Bezsonova I. et al., 2008), CIN85 may exert its function through binding to ubiquitinated proteins as IKK- γ (NEMO) and TRAF6 that have been reported to be important for NF- κ B activation (Sun L. Chen Z.J., 2004). The SH3 domains of CD2AP are also shown to bind ubiquitin (Ortega-Roldan J.L., Phd thesis, 2010; Ortega-Roldan J.L. et al., 2009), and it has been observed that CIN85 and CD2AP can bind to BLNK (signaling molecule used by BCR and TCR; Wienands J. et al., 1998), so it is possible that the association of CIN85/CD2AP with BLNK in the BCR signaling context, allows CIN85 and CD2AP to participate in the activation of the NF- κ B pathway (Kometani K. et al., 2011).

Finally, it was shown that Cbl and Cbl-b mediates monoubiquitination of CIN85 and CD2AP after EGF stimulation (Haglund K. et al., 2002). Monoubiquitination of proteins is associated with protein endocytosis and targeting for proteasomal degradation (Hicke L., 2001). So the monoubiquitination of CIN85/CD2AP might be a signal for their sorting to the lysosome and non-ubiquitinated CIN85/CD2AP are recycled back to the cytoplasm (Haglund K. et al., 2002).

Ubiquitination is one of the most sophisticated and versatile post-translational modifications, which regulates numerous signaling pathways within the cell. Ubiquitin signals are read out in a non-covalently way by ubiquitin-binding domains, which use to show preference for determined type of polyubiquitin chains (with different length and linkages) or monoubiquitin, giving rise to different functional outcomes.

To understand how this non-covalently “read-out” works within the ubiquitination-deubiquitination pathways within the cell, we have studied the interaction of ubiquitin with the zinc finger (ZnF) domains located in the novel ubiquitin-binding adaptor protein TAX1BP1 (Chapter V). Like SH3 domains, ZnF domains are some of the most

versatile intracellular domains found in the proteasome (Matthews M.J. and Sunde M., 2002).

TAX1BP1 interacts with the Lys63-polyubiquitin chains presented in TRAF6 and RIP1, involved in the IL-1 and TNF receptors pathways respectively (Chapter I). It has been demonstrated that TAX1BP1 plays a crucial role in the negative regulation of NF- κ B and JNK (c-jun N-terminal Kinase) signaling pathways from which TNF and IL-1 are potent activators (Shembade N. et al., 2007), opposite to the CIN85/CD2AP family of adaptor proteins, that have been recently linked to the activation of the NF- κ B pathway (Kometani K. et al., 2011).

TAX1BP1 forms part of the A20-ubiquitin-editing complex, composed also by A20, Itch, RNF11 and ABIN-1. However, till now the mechanisms involved in the interaction of TAX1BP1 with its different targets remains unknown. This requires structural information of TAX1BP1 itself and the macromolecular complexes. The determination of the high-resolution structure of UBZ1+2 in solution has been the first step in this challenging task, and was crucial for the study of the interaction with one of the key targets, ubiquitin (Chapter V).

It was reported that binding of TAX1BP1 with the polyubiquitinated substrates TRAF6, RIP1 and A20 involved exclusively the second C-terminal ubiquitin-binding domain (UBZ2) of TAX1BP1 (Iha H. et al., 2008). However, in our study we have shown that both UBZ domains of TAX1BP1 are able to bind ubiquitin with different affinities. The versatile interaction of ubiquitin with the two UBZ domains, UBZ2 in a tight, well-defined complex and UBZ1 in a weak, dynamic way, explains the ability of TAX1BP1 to recognize different polyubiquitin chains.

Based on the long-range intermolecular restraints provided from PRE NMR measurements, a model of the complex formed by the tandem UBZ1+2 and ubiquitin was determined, in agreement with the binding interfaces mapped on the surface of UBZ1+2 and ubiquitin by the use of chemical shift perturbation. The exquisite sensitivity of PRE NMR to determine low population states has been crucial to define the probability distribution map of ubiquitin around the weaker UBZ1 binding site within the tandem. The structural and dynamical features of the ubiquitin-UBZ interactions suggest a predisposition of the tandem for efficient binding to Lys48- and Lys63-polyubiquitin linkages, but especially to linear polyubiquitin chains. The use of bio-layer interferometry (BLI) has confirmed this hypothesis and validated the structural model obtained by PRE NMR.

In the TNF and IL-1 receptors signaling pathways, A20 cleaves Lys63-polyubiquitin chains attached to RIP1 and TRAF6, respectively, and replaces them by Lys48-polyubiquitin chains targeting them for degradation (Wertz I.E. et al., 2004; Newton K. et al., 2008; Habelhah H., 2010). In this mechanism, TAX1BP1 functions as an ubiquitin-binding adaptor protein, via its UBZ1 and UBZ2 domains, recruiting A20 to the polyubiquitinated proteins TRAF6 and RIP1 and leading to the disruption of IL-1 and TNF-induced NF- κ B signaling, respectively (Iha H. et al., 2008). In TAX1BP1-deficient cells, A20 fails to deubiquitinate Lys63-polyubiquitin TRAF6 and RIP1, resulting in the persistence of the TNF and IL-1, triggering NF- κ B and JNK activation (Shembade N. et al., 2007). By the use of BLI we have shown that UBZ1+2 can interact with Lys63- and Lys48-polyubiquitin chains. We can speculate that despite to recruit

Lys63-polyubiquitin substrates to A20, TAX1BP1 can also transport the Lys48-polyubiquitin substrates generated by A20 to the cytoplasm, where they can be recruited for proteasomal degradation.

TAX1BP1 is also involved in the antiviral immune responses by the negative regulation of the interferon regulatory factor (IRF) 3 (Parvatiyar K. et al., 2010). In this pathway TAX1BP1 also acts in concert with A20 and Itch, and they inhibit antiviral signaling by targeting TBK1 (TANK-binding kinase 1) and IKK ϵ kinases, both Lys63-polyubiquitinated. In this case, the IRF3 inhibitory effect of TAX1BP1 also depends on both UBZ domains.

Recently a new type of polyubiquitin linkage, known as linear or head-to-tail ubiquitin polymers, was suggested to be crucial for the NF- κ B signaling (Iwai K. and Tokunaga F., 2009; Tokunaga F. et al., 2009; Rahighi S. et al., 2009). LUBAC (linear ubiquitin chain assembly complex) is the only known E3 ligase that specifically catalyzes linear ubiquitin chains. Till recent, NEMO (NF- κ B essential modulator) was the only substrate known to be modified by linear polyubiquitin chains (Kirisako T. et al., 2006; Iwai K. and Tokunaga F., 2009). However, recently it was shown that LUBAC can also linearly ubiquitinate pre-existing ubiquitin moieties on RIP1 (Gerlach B. et al., 2011; Walczak H., 2011), highlighting the importance of linear ubiquitination in the TNF receptor pathway. In absence of LUBAC the TNF-mediated activation of NF- κ B still can occur by the presence of Lys63-polyubiquitin chains (Gerlach B. et al., 2011) but at significantly reduced levels (Ikeda F. et al., 2011; Tokunaga F. et al., 2011), making cells sensitive to TNF-induced cell death (Haas T.L. et al., 2009; Ikeda F. et al., 2011). Hence, linear ubiquitination is involved in the prevention of chronic inflammation through interference with TNF-induced cell death. In this sense, our results show the versatility of TAX1BP1-UBZ1+2 to recognize different polyubiquitinated substrates, which can have important consequences within the regulation of NF- κ B pathway.

On the other hand, TAX1BP1 is also directly related with a signaling pathway in which the linear polyubiquitinated NEMO is involved. TAX1BP1 was originally identified as binding partner of Tax1, which is the HTLV-I (human T cell leukemia virus type I) viral oncoprotein. This virus causes adult T-cell leukemia/lymphoma (ATL) and Tax plays a central role in the immortalization of infected T-cells and the persistence of infection in a host (Shoji T. et al., 2009; Mizuguchi M. et al., 2009). NF- κ B activation is crucial in HTLV-1 mediated oncogenesis and this process occurs mainly in the cytoplasm where the ubiquitinated form of Tax (composed predominantly of Lys63-linked chains) directly interacts with NEMO and triggers the activation of IKK α and IKK β (Chu Z.L. et al., 1999; Harhaj E.W. and Sun S.C., 1999; Jin D.Y. et al., 1999). It has been observed that TAX1BP1 can participate in the enhancement of Tax1 ubiquitination by interacting with the NRP (NEMO-related protein) and Tax1 forming a ternary complex that cooperate to positively regulate Tax1-induced NF- κ B activation (Journé C. et al., 2009). This mechanism is still unclear and future studies need to be addressed, however, linear polyubiquitinated NEMO plays an important role.

It is worth to notice that the Tax1/NRP/TAX1BP1 and the A20-editing binding complexes, lead to the activation and termination of NF- κ B signaling, respectively. Since, TAX1BP1 participates in both antagonist processes, it is logic to think that the adaptor protein is involved in an auto-regulation mechanism by interacting with different polyubiquitinated substrates and forming part of different complexes. The

different affinities of UBZ1+2 for distinct ubiquitin specific linkages contribute to the diverse set of functional outcomes associated with TAX1BP1.

The presence of multiple ubiquitinated substrates, involved in different signaling pathways within the cell, makes of ubiquitin an attractive target for adaptor proteins. CD2AP and CIN85 are also able to bind to the canonical hydrophobic patch on the β -sheet of ubiquitin through their SH3 domains (Bezsonova I. et al., 2008; Ortega-Roldan J.L., Phd thesis, 2010). Like UBZ1+2 of TAX1BP1 (Fig. V.46 Chapter V), the binding surface of both SH3A domains is negatively charged (Fig. VI.4), favoring the interaction with the positively charged canonical region of ubiquitin. Additionally, SH3-ubiquitin and UBZ-ubiquitin binding is stabilized by hydrophobic interactions involving large hydrophobic residues, including aromatics.

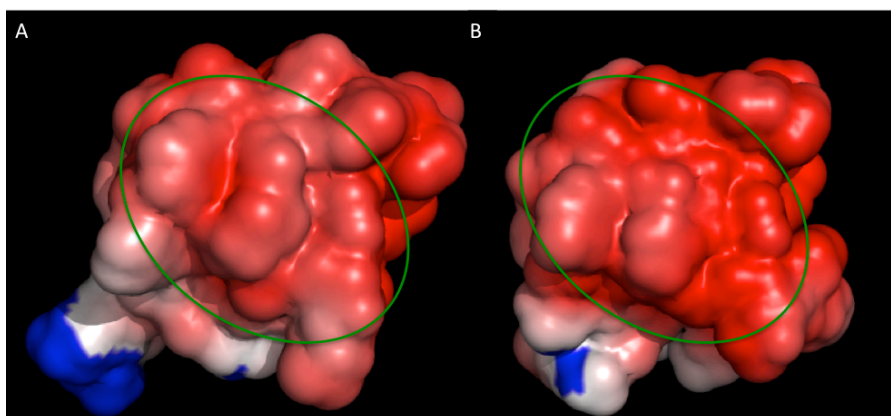


Figure VI.4. ± 5 KT/e electrostatic potential of the CIN85-SH3A (A) CD2AP-SH3A (B) domains in PyMOL plotted on the solvent accessible surface. Calculations were done using the Python software PDB2PQR and visualization was done using the computational packages of APBS (run in PyMOL) using the Poisson-Boltzmann electrostatics calculation (Dolinsky T.J. et al., 2004) in conjunction with the solvent-accessible surface area (Kukic P. and Nielsen J.E., 2010). Color code: grey for non-polar residues, blue and red for positively and negatively charged residues, respectively.

Consequently, electrostatic and hydrophobic forces are playing an important role in the binding mechanism of SH3 and UBZ domains with their natural targets. Electrostatic attraction between oppositely-charged molecules accelerates the association rates of the macromolecules (Elcock A.H. et al., 1999; Sheinerman F.B. et al., 2000; Gabdouliline R.R. and Wade R.C., 2001) and increase the amount of time that the two binding partners stay close to each other, allowing them to rotate into a proper orientation for binding (Sheinerman F.B. et al., 2000). However, consideration of intermolecular electrostatic forces alone is often not enough for the correct prediction of the binding geometry in transient protein-protein complexes (Volkov A.N., Phd thesis, 2007). The hydrophobic effect is one of the driving forces in protein-protein interactions (Chothia C. and Janin J., 1975), associated with binding-induced water release from the interface, which provides most of the binding affinity in a protein complex (Sturtevant J.M., 1977).

Interestingly, the ubiquitin-binding surface of the SH3A domains overlap with the canonical binding surface for proline-rich ligands, despite that ubiquitin does not contain proline residues on its SH3-domain binding interface (Bezsonova I. et al., 2008; Ortega-Roldan J.L., Phd thesis, 2010). This fact suggests an important competition between ubiquitin and proline-rich sequences in processes like clathrin-mediated receptor endocytosis, crucial for the internalization of the EGFR.

Competition between ubiquitin and other targets, such as the WW domains of Itch, is likely the case for UBZ1+2 of TAX1BP1 binding. TAX1BP1 binds WW domains via the conserved PPXY sequences present in both UBZ domains (Shembade N. et al., 2008). However it is not clear how TAX1BP1 can bind polyubiquitin substrates (RIP1 and TRAF6), Itch and A20 at the same time to form the A20-ubiquitin-editing complex. Possible explanations pointed to the formation of TAX1BP1 dimers (Verstrepen L. et al., 2011), since previous studies reported that the interactions with TRAF6, RIP1 and A20 were mediated exclusively by the UBZ2 domain (Iha H. et al., 2008). However, we have shown that UBZ1 also can interact with ubiquitin, opening up a wide range of possibilities. For example, one of the UBZ domains could present more affinity for WW domains than for ubiquitin, enabling binding of both targets at the same time; or both UBZ domains have similar affinities for both targets, implicating competition between the WW domains of Itch and ubiquitin. Given the structural characteristics of the UBZ1+2 tandem, both scenarios are possible. The PPXY motif in UBZ2 is located in the hinge region of UBZ1+2 (Fig. VI.5), which was shown important for the communication between the two UBZ domains. The PPXY sequence in UBZ1, however, is part of the ubiquitin-binding region, but seems more accessible for WW binding. So ubiquitin binding maybe does not interfere with WW binding to the UBZ2 domain, but will compete in the case for binding to UBZ1. One could envisage a heterotrimeric complex of ubiquitin bound to UBZ2 and a WW domain of Itch to UBZ1. This is very speculative and it will therefore be extremely interesting and important to study the interaction of the WW domains of Itch with UBZ1+2, and the possible competition with ubiquitin binding. It is worth to notice that like SH3 domains, WW domains also mediate protein interactions by binding proline-rich ligands (Chan D.C. et al., 1996).

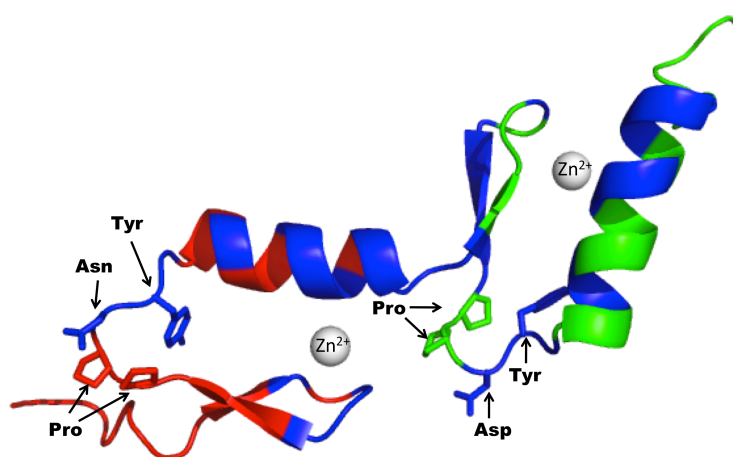


Figure VI.5. NMR structure of the UBZ1 (in red) and UBZ2 (in green) domains of TAX1BP1 in tandem. The residues showing large chemical shift perturbations upon binding ubiquitin are colored in blue and the side chains of the residues corresponding with the conserved PPXY sequences (PPNY in UBZ1 and PPDY in UBZ2) are shown in sticks.

As discussed above, the interactions of UBZ domains of TAX1BP1 and the SH3A domains of CIN85/CD2AP with ubiquitin are driven by similar forces (electrostatic and hydrophobic), and both of them recognize the same canonical binding interface on ubiquitin. However, the secondary structural elements used by UBZ and SH3A to bind ubiquitin are totally different. This is very common in ubiquitin-binding domains (UBDs), that diverge a lot in the structure and recognition mode that they use. Most of UBDs involve α -helices to bind the hydrophobic patch of ubiquitin (Dikic I. et al., 2009). However, comparison of the α -helix sequences of the UBZ1 and UBZ2 domains with some of the most characteristic UBDs presenting helical structures, including the UBZ domain of the human DNA Y-polymerase (DNA pol eta), shows that UBZ1+2 of TAX1BP1 does not contain the previously established canonical binding sequence (Bomar M.G. et al., 2007; Cordier F. et al., 2008). In this way, we have characterized a new ubiquitin-binding mode involving UBD domains that contain an α -helix within their structure, especially in the family of UBZ domains. Consequently, it is very difficult to identify UBDs and to predict ubiquitin-binding surfaces without experimental data. The work described in this thesis contributes to the better understanding of the structure and interaction mechanism of the family of UBZ domains, which function was less understood till now, in comparison with other UBD domains (Hurley J.H. et al., 2006; Grabbe C. and Dikic I., 2009), setting the bases for future studies.

Until now, only two structures of UBZ domains can be found in the PDB, apart from the UBZ1+2 domains of the human TAX1BP1. The first one corresponds to the UBZ domain of the human DNA pol eta (PDB code 2I5O), a classical Cys₂His₂ UBZ domain like UBZ1+2 of TAX1BP1; the second one is the UBZ domain of the human WRNIP1 (Werner helicase-interacting protein 1) (PDB code 3VHS) that is involved in DNA replication (Crosetto N. et al., 2008). The UBZ of WRNIP1 presents a variation of the classical Cys₂His₂ motif, being characterized by a Cys₂HisCys motif. Other UBZ domains like the ones in ubiquitin ligase RAD18 and polymerase kappa are also Cys₂HisCys domains (Crosetto N. et al., 2008). UBZ domains of DNA pol eta, DNA pol kappa, RAD18 and WRNIP1 share a common biochemical function, besides that they are able to interact with ubiquitin, all of them bind DNA and participate in post-replication DNA repair (Bish R.A. and Myers M., 2007). This is one of the main differences with the UBZ domains of TAX1BP1, which have never been reported to bind DNA (Hofmann K., 2009). Other important aspect to take into account is that the UBZ domain of DNA pol eta and both UBZ domains of TAX1BP1 are able to bind monoubiquitin while the UBZ domain of WRNIP1 does not bind monoubiquitin but only polyubiquitin chains with no specificity for a determined chain linkage (Bish R.A. and Myers M., 2007).

Given the degree of structural similarity between the three published structures and knowing their functional differences, it is interesting to identify if these members of the UBZ family of proteins were evolutionary related. In order to answer this question, the program jPRED3 (Cole C. et al., 2008) was used to search sequences similar to the ones of the UBZ1 and UBZ2 domains of TAX1BP1, UBZ domain of pol eta and UBZ domain of WRNIP1, where all the alignments were filtered out for redundancy at 75% sequence identity (Cole C. et al., 2008). The output was used to perform a multiple sequence alignment with the program Genious 6.1.4 (Kearse M. et al., 2012). Figure VI.6 shows the alignment of all the sequences corresponding to different organisms.

As we can see, UBZ1 and UBZ2 of TAX1BP1 are four residues longer than the other two types of UBZ domains in the region corresponding to the loop that links the β -sheet with the α -helix (Fig. VI.6). We have shown by comparison of X-ray and NMR structures that this loop in the UBZ2 domain acts like a hinge in the tandem UBZ1+2, determining the relative orientation of one domain with respect to the other, possibly contributing to regulate the tandem flexibility necessary for target recognition. In this sense, the existence of longer loops in UBZ domains of TAX1BP1 can be related with the fact that they are forming a tandem, while DNA pol eta and WRNIP1 only contain a single UBZ domain. In general Cys and His residues are very well conserved in the three different proteins within the different species, which highlight their essential role in maintaining the 3D structure of the domains by Zn^{2+} -coordination. It is worth to notice, that the Ala (in the C-terminal part of the α -helix) crucial for the binding of the UBZ domain of DNA pol eta with ubiquitin (Bormar M.G. et al., 2007), is always conserved across the organisms inside of the DNA pol eta family but not in the other UBZ domains. For UBZ2 of TAX1BP1 there is an Asp in that position and the consensus sequence centered around the Ala (Bormar M.G. et al., 2007; Cordier F. et al., 2008) is not even present in the UBZ1 domain of TAX1BP1, which α -helix is shorter. The UBZ domain of WRNIP1 presents an α -helix whose length is comparable to TAX1BP1 UBZ1 and does neither contain this consensus region. However, it contains a conserved Asp residue at position 37 in the middle of the helix that is crucial for binding to ubiquitin as a single point mutation of Asp37 into an Ala completely abolished ubiquitin-binding (Bish R.A. and Myers M., 2007). Despite that the conserved Ala of DNA pol eta is not present in any of the other UBZ domains, they are still able to bind ubiquitin, showing that the consensus sequence proposed by the sequence alignment of DNA pol eta and other zinc finger domains like NEMO and IUM (Bormar M.G. et al., 2007; Cordier F. et al., 2008) is not a general feature to explain the binding to ubiquitin of UBD domains containing an α -helix.

Other conserved residues are the prolines positioned after the first conserved Cys, and in the fifth position after the second conserved Cys; this last Pro corresponds to the first Pro of the PPXY sequence exclusively conserved in the UBZ1 and UBZ2 of TAX1BP1 but not in the UBZ domains of the DNA pol eta or WRNIP1. This Pro is always located at the beginning of the loop that connects the β -sheet with the α -helix, and its conservation is another characteristic structural feature of this family of domains. In UBZ1+2 of TAX1BP1 and the UBZ domain of DNA pol eta, a Val or a Met after the first consensus His is also conserved, and an aromatic residue after the second conserved His (a Trp in UBZ1 and a Phe in UBZ2 of TAX1BP1 and in UBZ of DNA pol eta) which is missing in UBZ of WRNIP1 that contains a conserved Leu in that position. These aromatic residues in the UBZ domains of TAX1BP1 and DNA pol eta are not part of the hydrophobic core of the zinc fingers but are on the exposed surface of the α -helices and are involved in binding to ubiquitin (Chapter V).

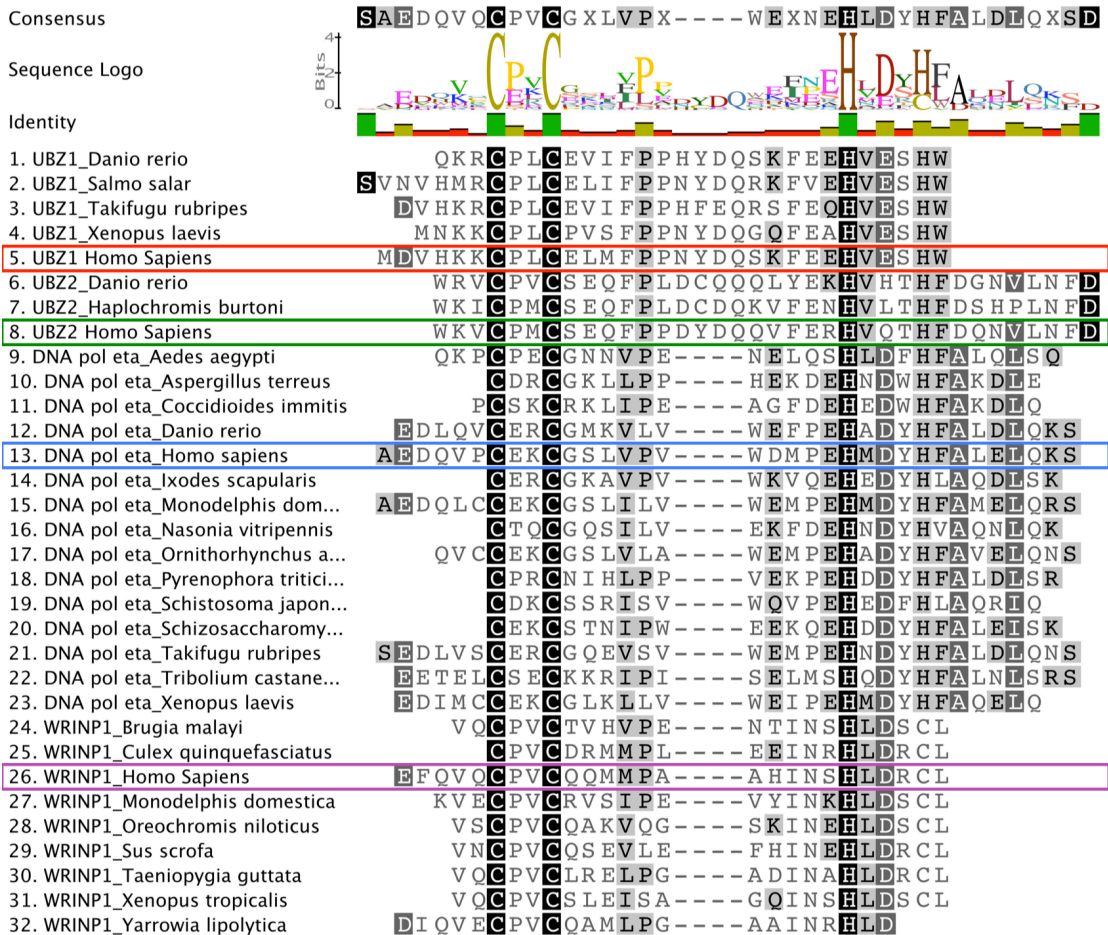


Figure VI.6. Multiple sequence alignment performed with the program *Genious 6.1.4* (Kearse M. et al., 2012) for the UBZ domains of *TAX1BP1* (human UBZ1 surrounded by a red box and UBZ2 by a green one), *DNA pol eta* (human UBZ surrounded by a blue box) and *WRINP1* (human UBZ surrounded by a pink box) present in different organisms.

Based on the sequence alignment and using the program *Genius 6.1.4* (Kearse M. et al., 2012) we created a phylogenetic tree (see Fig. VI.7). This diagram represents the interferences of the evolutionary relationship among organisms, where every intersection corresponds to the last common ancestor. As we can see in figure VI.7, despite their structural and sequence similarity, the UBZ domains of *TAX1BP1*, *DNA pol eta* and *WRINP1* have evolved in independent ways.

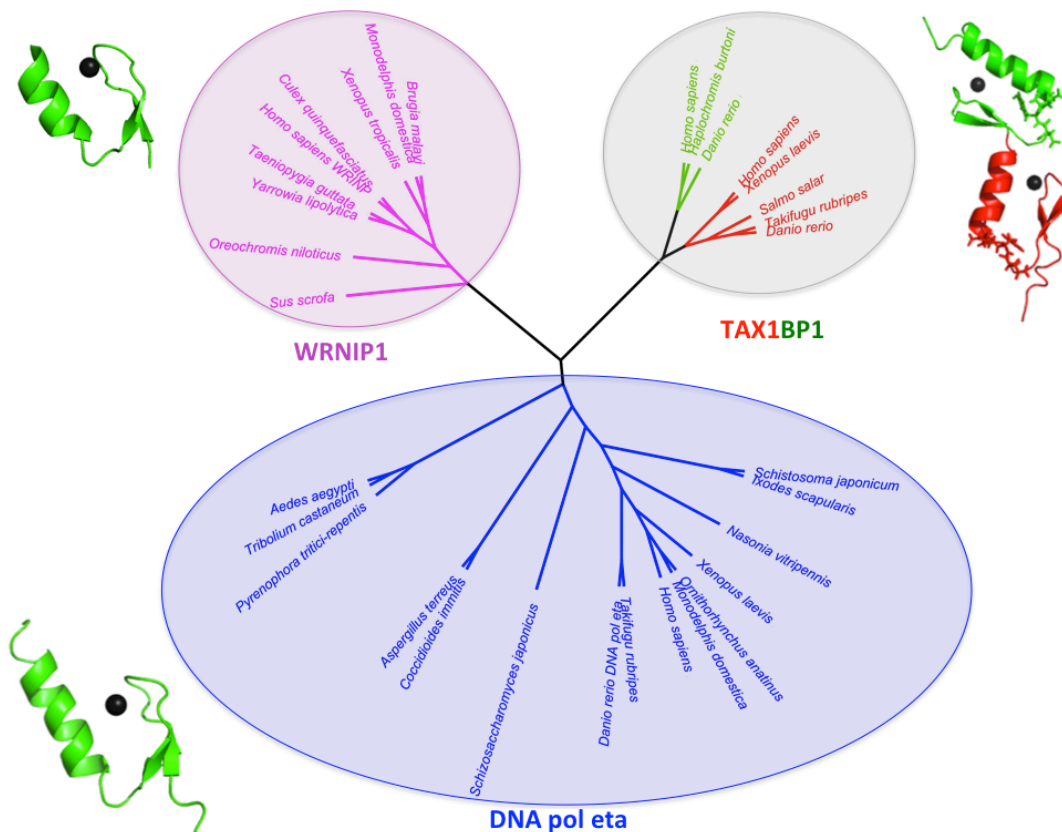


Figure VI.7. Phylogenetic tree (made with Genious 6.1.4) based on the multiple sequence alignment (Fig.VI.6) of the UBZ domains of TAX1BP1, WRNIP1 and DNA pol eta in different organisms, where it is shown the divergent evolution of the UBZ domains present in the three proteins. Color code of the tree in agreement with figure VI.6. For every group a cartoon representation of the published structure of the representative human UBZ domain is shown (UBZ domain of DNA pol eta, PDB code 2I5O; UBZ domain of WRNIP1, PDB code 3VHS and UBZ1+2 tandem of TAX1BP1, PDB code 2M7Q). The four residues that increase the length of the loops of the UBZ domains of TAX1BP1 have been represented in sticks for clarity. Zn²⁺ atoms are represented as black spheres.

The divergent evolution followed by the UBZ domains of these three proteins can be related with their distinct functions within the cell: DNA pol eta is implicated in DNA repair by translesion synthesis, WRNIP1 is implicated in DNA replication and repair, and TAX1BP1 is implicated in the regulation of transcription factors, such as NF- κ B and JNK, but not directly involved in DNA binding. The phylogenetic differences based on sequence alignment point to the existence of distinct binding mechanisms within these UBZ domains to its common partner, ubiquitin, opening the door for new studies oriented to deeply characterize this versatile family of domains.

The phylogenetic tree also reveals an important detail about the UBZ domains of TAX1BP1: they diverged during evolution, presenting a common ancestor from which they differentiate probably due to a genetic mutation that might result energetically more favorable, and because of that evolved till the present. It would be very interesting to perform a detail phylogenetic study about these domains to determine their common ancestor and follow their trajectories of evolution till the present.

We can conclude that the versatile and selective interactions of modular domains present in adaptor proteins such as CD2AP, CIN85 and TAX1BP1 with their natural targets provide a crucial mechanism to organize cell regulatory pathways, necessary for the dynamic responses to external and internal signals. Based on that, small modular domains like SH3 and UBZ present a potential therapeutic value, being key targets to the design of specific drugs that might be used in treatments of inflammatory and autoimmune diseases, nephritic syndrome and cancer, among others.

CHAPTER VII

RESUMEN Y CONCLUSIONES *SUMMARY AND CONCLUSIONS*

▪

The scientist has in common with the artist only this: that he can find no better retreat from the world than his work and no stronger link with the world than his work. Max Delbruck

Chapter VII. RESUMEN Y CONCLUSIONES

7.1 Resumen

El correcto funcionamiento celular tiene importantes implicaciones en la salud de los seres vivos. Dentro de las células hay un continuo flujo de información y energía, crucial para mantener las funciones celulares esenciales, como proliferación, migración y homeostasis.

La unión de ligandos externos a los receptores situados en la superficie de las células desencadena una cascada de interacciones entre proteínas que da lugar a la amplificación de las señales externas, lo que finalmente se traduce en una respuesta celular.

Para asegurar los correctos niveles de señalización, la naturaleza ha desarrollado inteligentes mecanismos de control negativo que permiten la terminación de dichas cascadas de señalización celular activadas por la unión de ligandos a receptores celulares.

Uno de esos ligandos es el factor de crecimiento epidérmico (EGF), que se puede unir a su correspondiente receptor celular EGFR (receptor del factor de crecimiento epidérmico), perteneciente a la familia de los receptores tirosina kinasa (RTKs) y está implicado en diferentes procesos biológicos como proliferación, diferenciación, migración, adhesión y apoptosis celular. La regulación negativa de EGFR está mediada por la formación de un complejo capaz de eliminar receptores activados de la membrana plasmática y que implica la interacción de distintas macromoléculas, como endofilinas, la ubiquitina ligasa Cbl y la proteína adaptadora CIN85.

Las proteínas adaptadoras juegan un papel muy importante en los mecanismos de señalización celular, y se caracterizan por poseer en su estructura pequeños dominios modulares en tándem, a través de los cuales interactúan con dianas moleculares dando lugar a complejos multi-proteicos. Su especificidad y afinidad por determinados ligandos va a determinar la respuesta y la eficiencia de un determinado proceso de señalización. CIN85 y su homólogo CD2AP, forman parte de una familia de proteínas adaptadoras que contienen tres dominios SH3 (SH3A, SH3B y SH3C) en su extremo amino terminal. Estos dominios SH3 son capaces de interactuar con secuencias “atípicas” ricas en prolina, a las que pueden reconocer en dos posibles orientaciones, conocidas como tipo I y/o tipo II. Algunas proteínas como Cbl/Cbl-b y CD2 presentan estas secuencias atípicas de prolina. La interacción de CD2AP con CD2 cumple un papel importante en la polarización de células T y en la sinapsis inmunológica. También se cree que CIN85 está involucrado en este proceso debido a su presencia en células T y a que es capaz de interactuar con CD2. Sin embargo la composición y estequiometría de los complejos que los dominios SH3 de estas proteínas adaptadoras forman con sus ligandos no están totalmente caracterizadas, pudiendo variar entre la formación de dímeros, trímeros o incluso mayores oligómeros.

Mediante el uso de RMN, ITC y SAXS, hemos podido caracterizar los diferentes complejos formados cuando el dominio amino terminal (SH3A) de las proteínas adaptadoras CD2AP y CIN85 interacciona con las secuencias atípicas de prolina presentes en las proteínas CD2 y Cbl-b, y los hemos comparado con otros estudios realizados al respecto y diferentes estructuras cristalinas presentes en la base de datos de proteínas (PDB). Nuestros experimentos muestran que el dominio SH3A de CD2AP puede formar un complejo dimérico tipo II con CD2 y complejos diméricos tipo I y tipo II con Cbl-b. Por su parte, el dominio SH3A de CIN85 también forma un complejo dimérico tipo II con CD2, pero con Cbl-b es capaz de formar un trímero donde las interacciones tipo I y tipo II tienen lugar al mismo tiempo. Estos resultados podrían explicar porqué la interacción de similares dominios SH3 con sus ligandos naturales puede regular un amplio rango de mecanismos de señalización celular involucrados en la endocitosis de tirosinas kinasas, adhesión celular y señalización en células T.

TAX1BP1 es una proteína adaptadora involucrada en el mecanismo de regulación negativo del factor de transcripción NF- κ B, que es esencial para el control de procesos inflamatorios. Para ello TAX1BP1 forma parte de un complejo proteico junto con otras proteínas (A20, Itch, RNF11 y ABIN), conocido como complejo-editor de ubiquitinas, donde TAX1BP1 actúa como una proteína adaptadora capaz de reclutar substratos poliubiquitinados para la enzima A20. El extremo carboxílico terminal de TAX1BP1 contiene dos pequeños dominios modulares en tándem, conocidos como dedos de zinc, que son capaces de interactuar con ubiquitina (UBZ, ubiquitin-binding zinc fingers: UBZ1 y UBZ2). Estos dominios son esenciales para entender el funcionamiento de TAX1BP1, sin embargo, hasta ahora, su estructura en solución y su mecanismo de interacción con ubiquitina no eran conocidos.

Para poder caracterizar la interacción de los dominios UBZ1+2 de TAX1BP1 con ubiquitina, primero ha sido necesario llevar a cabo la determinación estructural del tándem UBZ1+2 en solución mediante el uso de RMN. La estructura tridimensional de UBZ1+2 presenta las características comunes de los dedos de zinc clásicos (Cys₂His₂), incluyendo la coordinación tetraédrica a los átomos de zinc. Esta estructura representa el primer tándem de dominios UBZ en solución que ha sido publicado hasta el momento.

El uso de ITC, RMN, RMN paramagnético (PRE NMR) y SAXS ha sido imprescindible para la caracterización termodinámica y estructural de la interacción del tándem UBZ1+2 con ubiquitina, dando lugar a la obtención de un modelo estructural del complejo que podría explicar la capacidad de TAX1BP1 de interactuar con distintas cadenas de poliubiquitina, lo que podría tener importantes consecuencias en la regulación del factor NF- κ B. La promiscuidad del tándem UBZ1+2, capaz de interactuar con distintos tipos de cadenas de poliubiquitina (Lys48-, Lys63- y lineales) ha sido demostrada mediante el uso de BLI. Estos resultados han servido para aumentar nuestro conocimiento sobre el mecanismo de interacción de TAX1BP1 con una de sus principales dianas naturales, ubiquitina.

7.2 Conclusiones

A continuación se enumeran las principales conclusiones que se pueden extraer del trabajo presentado en esta tesis doctoral.

Conclusiones: dominios SH3A de la familia de proteínas adaptadoras CIN85/CD2AP

- Los dominios SH3A presentes en la familia de proteínas adaptadoras CD2AP y CIN85 interaccionan con péptidos que presentan la secuencia atípica rica en prolina PxxxPR (que no contiene un residuo con carga positiva en el extremo amino terminal) formando dímeros tipo II, como en el caso de CD2.
- Los dominios SH3A presentes en la familia de proteínas adaptadoras CD2AP y CIN85 interaccionan en orientación tipo I y tipo II con los péptidos que presentan una arginina amino terminal en la secuencia atípica rica en prolinas, como es el caso de Cbl-b. Es por tanto necesario extender la secuencia atípica rica en prolinas de PxxxPR a RxPxxxPR para incluir la arginina amino terminal. La ausencia de dicha arginina resulta en una interacción prevalentemente tipo II.
- El dominio SH3A de CD2AP puede interaccionar con Cbl-b formando dímeros tipo I y tipo II. Sin embargo, el dominio SH3A de CIN85 es capaz de interaccionar con Cbl-b formando un trímero donde las interacciones tipo I y tipo II tienen lugar simultáneamente.
- La interacción de los dominios SH3A de CD2AP y CIN85 en una orientación tipo I o tipo II esta determinada por la existencia o no de una arginina amino o carboxilo terminal en la secuencia rica en prolinas, sin embargo, la posibilidad de formar un trímero o un dímero está relacionada con la existencia de impedimentos estéricos entre los dominios SH3, especialmente debido a las diferencias de longitud y orientación existentes en el lazo n-Src del dominio SH3A de CIN85 y CD2AP.
- La direccionalidad de los desplazamientos químicos observados durante las titulaciones de RMN en el protón indol del Trp presente en la superficie de interacción de los dominios SH3A, es informativa. Los desplazamientos químicos a campo bajo son indicativos de la interacción tipo II y los desplazamientos químicos a campo alto son indicativos de la interacción tipo I. Sin embargo, la presencia de curvatura en la trayectoria de dichos desplazamientos es indicativa de ambos tipos de interacciones, tipo I y tipo II.
- La proporción trímero/dímero que el dominio SH3A de CIN85 puede formar con Cbl-b está directamente relacionada con la concentración de Cbl-b en el medio. Este hecho implica que los mecanismos de multimerización de CIN85 y CD2AP son distintos, lo que podría tener importantes implicaciones biológicas.
- El uso complementario de tres diferentes técnicas en solución (ITC, NMR y SAXS), ha sido decisivo para clarificar la composición y estequiometría de los complejos moleculares en los que los dominios SH3A de las proteínas adaptadoras CIN85/CD2AP están involucrados.

Conclusiones: dominios UBZ1+2 de la proteína adaptadora TAX1BP1

- Se ha determinado la estructura tridimensional en solución del tándem UBZ1+2 de TAX1BP1 mediante el uso de RMN. El tándem UBZ1+2 presenta todas las características estructurales comunes en dedos de zinc clásicos (Cys₂His₂). El alto grado de concordancia con la reconstrucción *ab initio* de la estructura obtenida mediante SAXS ha servido para validar la estructura en solución. La coordinación tetraédrica de átomos de Zn²⁺ es crucial para el mantenimiento de la estructura tridimensional de los dominios UBZ1+2, necesaria para interactuar con ubiquitina.
- Mediante la comparación de la estructura cristalina y en solución del tándem UBZ1+2, se ha podido observar que el lazo que conecta la lámina β con la hélice α del dominio UBZ2 funciona como “articulación” del tándem, determinando la orientación relativa de un dominio con respecto a otro e incrementando la flexibilidad del tándem, probablemente necesaria para el reconocimiento de ligandos. El análisis cuantitativo de los parámetros de relajación (RMN) ha confirmado la flexibilidad de este lazo en la forma libre del tándem UBZ1+2.
- Los dos dominios UBZ de TAX1BP1 son capaces de interactuar con ubiquitina, tanto como dominios aislados como formando parte del tándem UBZ1+2. En cualquier caso, la unión de UBZ2 a ubiquitina es siempre un orden de magnitud más favorable que la interacción de UBZ1.
- Los dominios UBZ1 y UBZ2 experimentan un incremento de la afinidad por ubiquitina cuando están juntos formando parte del tándem UBZ1+2, en comparación con las afinidades de los dominios independientes. Estos resultados apuntan a la posible existencia de un mecanismo de cooperación positiva entre ambos dominios que ha sido sugerido por los experimentos de ITC y RMN.
- Los dominios UBZ1 y UBZ2 en tándem y como dominios independientes, son capaces de reconocer selectivamente una única interfaz en ubiquitina, la cavidad hidrofóbica conocida como motivo canónico o Leu8-Ile44-Val70.
- La distribución de potenciales electrostáticos en la superficie de UBZ1+2 y ubiquitina sugiere que las interacciones iónicas están jugando un papel muy importante en el mecanismo de unión. Sin embargo, la estabilidad del complejo está fuertemente influenciada por fuerzas hidrofóbicas, principalmente debido a la presencia de residuos hidrofóbicos en las interfaces de unión de ubiquitina (cavidad hidrofóbica) y UBZ1+2.
- El uso de RMN paramagnético ha sido crucial para obtener un modelo estructural del complejo formado entre UBZ1+2 y ubiquitina. En este complejo, ubiquitina se une a UBZ2 a través de una posición única y bien definida, mientras que la unión de ubiquitina a UBZ1 está caracterizada por la flexibilidad espacial de ubiquitina, que puede ser representada por un mapa de densidad de probabilidad, donde dos áreas a ambos lados del dominio UBZ1 están principalmente pobladas. Este modelo ha sido validado por su alto grado de

concordancia con las interfaces de unión determinadas mediante el uso de desplazamientos químicos por RMN.

- Los datos paramagnéticos calculados en el proceso de docking usando la estructura en solución del tándem UBZ1+2 determinada mediante RMN, se ajustan mejor a los resultados experimentales que los datos calculados mediante el uso de la estructura cristalina de UBZ1+2. De este modo, RMN paramagnético ha permitido validar la estructura de la forma libre de UBZ1+2 obtenida por RMN.
- Alineamiento de secuencias y estudios filogenéticos revelan que la secuencia consenso determinada para la unión a ubiquitina de dominios UBD que contienen una hélice α no es representativa de toda la familia de dominios UBZ, no estando presente en los dominios UBZ1+2 de TAX1BP1. Por otra parte, a pesar de las semejanzas en estructura y secuencia dentro de la familia de dominios UBZ, los dominios UBZ1 y UBZ2 de TAX1BP1 han evolucionado de una manera independiente al resto de UBZ cuya estructura se conoce por el momento.
- BLI ha servido para estudiar la avidez de UBZ1+2 por distintas cadenas de poliubiquitina (promiscuidad). UBZ1+2 presenta preferencia por cadenas de poliubiquitina lineales, siendo la interacción con cadenas Lys63- un poco más afín que con cadenas Lys48-. Este versátil rango de interacciones puede tener importantes consecuencias en los mecanismos de señalización relacionados con substratos poliubiquitinados en los que TAX1BP1 está involucrada.

7.3 Summary

The proper functioning of cells determines whether a human being is healthy or not. Inside of the cells there is a continuous flow of information and energy, which is crucial for proliferation, migration, cell survival and homeostasis.

The binding of external ligands to the receptors located in the cellular surface triggers a cascade of protein-protein interactions that amplifies the signals, which finally causes a cellular response.

To ensure proper control of the signalization levels, the nature has developed a clever feedback, based on mechanisms of negative regulation that leads to the termination of the activate signaling pathways.

One of this external ligand is the epidermal growth factor (EGF) that binds EGF receptors (that belong to the family of receptor tyrosine kinases, RTKs), and is implicated in different biological processes like proliferation, differentiation, migration, adhesion or apoptosis. The down-regulation of the EGFR is mediated by the formation of a complex involving endophilins, Cbl ubiquitin ligase and CIN85 adaptor protein, capable of removing activated receptors form the plasma membrane.

Adaptor proteins play an important role in signalization pathways; they contain small modular domains in tandem crucial for the assembly of multi-protein complexes. Their specificity and affinity for particular targets determine the efficiency and outcome of cellular signaling. CIN85 and its homolog CD2AP, form part of a family of adaptor proteins that contain three SH3 domains that binds atypical proline-rich sequences, which can be recognized in two different orientations, type I and/or type II. Cbl/Cbl-b and CD2 enclose a region with such atypical proline-rich sequence. Direct molecular interaction of CD2AP with CD2 is crucial for CD2 clustering as well as T-cell polarization and recruitment of CD2 during the formation of the immunological synapse. CIN85 has also been proposed to participate in this process since it is highly expressed in T-cells and can bind CD2. However, the composition and stoichiometry of protein clusters involving such adaptor molecules was not fully established and might vary from dimeric to trimeric or larger oligomeric complexes.

Using NMR, ITC and SAXS methods, we have characterized several binding modes of the N-terminal SH3 domain (SH3A) of CD2AP and CIN85 with two natural atypical proline-rich ligands present in CD2 and Cbl-b; our data were compared and contrasted with previous studies and published crystal structures. Our experiments show that the CD2AP-SH3A domain forms type II dimer with CD2, and both type I and type II dimeric complexes with Cbl-b. Like CD2AP, the CIN85-SH3A domain forms a type II complex with CD2, but a trimeric complex with Cbl-b where type I and type II interactions take place at the same time. Together, these results explain how multiple interactions among similar SH3 domains and ligands can produce a high degree of diversity in tyrosine kinase, cell adhesion or T-cell signaling pathways.

TAX1BP1 is an adaptor protein, involved in the negative regulation of the NF- κ B signaling pathway by forming the TAX1BP1-Itch-A20 ubiquitin-editing complex, essential in limiting inflammation. TAX1BP1 functions as a crucial ubiquitin-binding adaptor protein that recruits A20 to specific ubiquitinated substrates. The C-terminal

extreme of this adaptor contains two small modular domains (UBZ1 and UBZ2) in tandem, known like zinc fingers domains, which can bind ubiquitin. These domains are crucial for the understanding of TAX1BP1 function.

To characterize the network of interactions involving the UBZ domains of TAX1BP1, we have determined the solution structure of the tandem UBZ1+2 by NMR. The three-dimensional structure of UBZ1+2 shows all characteristics typically found in classical zinc finger domains, including the tetrahedral Cys₂His₂ Zn²⁺-coordination. Noteworthy, this is the first tandem of classical UBZ domains whose structure in solution is published till present.

The structural and thermodynamic characterization of the interaction between UBZ1+2 domains of TAX1BP1 and ubiquitin by the use of NMR (including paramagnetic NMR), ITC and SAXS have been crucial to determine that both UBZ domains bind ubiquitin with different affinities and to obtain a solution structural model of the complex between ubiquitin and the tandem UBZ1+2. This model underlies functional promiscuity in the TAX1BP1 recognition mechanism of polyubiquitin chains, which has been experimentally tested using BLI. These results have enriched our understanding of the interaction mechanisms of TAX1BP1 with one of its principal natural targets, ubiquitin.

7.4 Conclusions

The principal conclusions extracted from the work presented in this thesis are enumerated below.

Conclusions: *SH3A domain of CIN85/CD2AP family of adaptor proteins*

- The SH3A domains of CIN85/CD2AP family of adaptor proteins, interact forming hetero-dimers in a type II mode with peptides that present the atypical proline-rich sequence PxxxPR, and lack a positive charge in the N-terminal core, like is the case of CD2 peptide.
- The SH3A domains of CIN85/CD2AP family of adaptor proteins, interact in a type I and type II mode with peptides that present an Arg in the N-terminal core, like is the case of Cbl-b peptide. The earlier defined atypical proline recognition sequence, PxxxPR, has to be extended to RXPxxxPR, including the N-terminal Arg. The absence of such residue results in prevalent type II dimers.
- The SH3A domains of CD2AP can interact with the peptide Cbl-b forming hetero-dimers type I and type II. However the SH3A domain of CIN85 is able to interact with Cbl-b peptide forming a hetero-trimeric complex where the interactions type I and type II take place at the same time.
- The formation of type I or type II interaction involving SH3A of the CIN85/CD2AP family is related to the polyproline sequence of the target molecules, while the formation of a dimer or a trimer is related to steric impediments between the two SH3 domains, especially due to differences in the n-Src loop of the SH3A domains of CIN85 and CD2AP.
- The directionality of the NMR shift change of the indole proton of the conserved tryptophan in the binding interface of SH3A domains is informative. Downfield change is indicative of type II interaction, the upfield shift change is indicative of type I binding, while the curvature in the trajectory it is indicative of type I and type II interactions.
- The formation of the trimer of CIN85-SH3A with Cbl-b is concentration dependent. The fact that the mechanism of clustering in the case of CIN85 and CD2AP will be different at high or at low concentration of ligand might have important biological implications.
- The use of three different and complementary techniques in solution, like NMR, ITC and SAXS, has been crucial to clarify the composition and stoichiometry of the protein clusters in which the SH3A domains of CIN85/CD2AP family of adaptor proteins are involved.

Conclusions: *UBZ1+2 domains of TAX1BP1 adaptor protein*

- We have determined the 3D structure of UBZ1+2 tandem of TAX1BP1 in solution by NMR, which presents all the typical features of a classical Cys₂His₂ zinc finger domain. The *ab initio* reconstructed shape obtained by SAXS, has been very useful to validate the NMR structure. The tetrahedral coordination of the Zn²⁺ atoms is crucial to maintain the 3D structure of the UBZ1+2 domains, necessary for interacting with ubiquitin.
- We have shown by comparison of the X-ray and NMR structures that the loop connecting the β -strand with the α -helix in the UBZ2 domain acts like a hinge region in the tandem UBZ1+2, determining the relative orientation of one domain with respect to the other, possibly enhancing the tandem flexibility necessary for target recognition. The quantitative analysis of the relaxation parameters confirms the flexibility of the loop region in the free UBZ1+2 tandem.
- Both UBZ domains of TAX1BP1 are able to interact with ubiquitin, either as single UBZ domains or in tandem. The binding of UBZ2 domain to ubiquitin is always one order of magnitude tighter than binding of UBZ1.
- UBZ1 and UBZ2 domains in tandem experience an enhancement in the binding affinity to ubiquitin as compared with the single domains, indicating the possibility of the existence of a mechanism of positive cooperativity between both UBZ, as suggested by ITC and NMR titrations.
- UBZ1 and UBZ2 domains in tandem and isolated selectively recognize one binding interface of ubiquitin, the Leu8-Ile44-Val70 hydrophobic patch of ubiquitin, also known as canonical interface.
- The distribution of the electrostatic potential on the surface of UBZ1+2 and ubiquitin suggests that ionic interactions are playing an important role, however, complex stability is strongly influenced by hydrophobic forces, mainly due to the presence of hydrophobic residues in the binding interfaces of ubiquitin (hydrophobic patch) and UBZ1+2.
- The use of PRE NMR has been crucial to obtain a model of the complex between UBZ1+2 and ubiquitin, where one ubiquitin binds to UBZ2 domain in a well defined relative position, while the binding to UBZ1 is characterized by a spatial distribution of ubiquitin molecules, where two areas at both sides of the UBZ1 domain are mainly populated. This model was validated by the agreement with the NMR chemical shift perturbation distribution.
- The PRE rates back-calculated from the UBZ1+2 NMR ensemble in the docking process fit the experimental PRE better than the ones from the X-ray structure. In this way, PRE NMR has served to validate the NMR ensemble.
- Sequence alignment and phylogenetic studies reveal that the canonical binding sequence present in other α -helical UBD domains is not a common feature, being not present in the UBZ1 and UBZ2 domains of TAX1BP1. Moreover,

despite structural and sequence similarity between members of the UBZ family, the UBZ domains of TAX1BP1 have evolved in independent ways to the rest of UBZ whose structures have been published till the present.

- BLI experiments show the promiscuity in the interaction of UBZ1+2 of TAX1BP1 with mono and polyubiquitin chains. UBZ1+2 has a preference for binding to linear polyubiquitin chains, and binds Ly63-linkages a bit tighter than Lys48-linkages. This versatile range of interactions might have important biological consequences in the regulation of the ubiquitin-related pathways in which TAX1BP1 is involved.

ABBREVIATIONS

ABBREVIATIONS

1D	1 dimensional
2D	2 dimensional
2xTY	Two times Tryptone-Yeast extract medium
3D	3 dimensional
A20	The ubiquitin-editing enzyme
Amp	Ampicillin
AT3	Ataxin-3
ATL	T-cell leukemia/lymphoma
BLI	Bio-layer interferometry
BLNK	B-cell linker protein
C-terminal	Carboxylic terminal extreme of the protein
Cbl	Casitas B-lineage lymphoma
Cbl-b	Casitas B-lineage lymphoma proto-oncogen b
CC	Coiled-coil
CD	Circular Dichroism
CD2	Cluster of differentiation 2
CD2AP	CD2 associated protein
CD4	Cluster of differentiation 4
CIN85	Cbl interacting protein of 85 kDa
CMS	Cas ligand with multiple SH3 domains
C_p	Heat capacity
CSP	Chemical shift perturbation
CUE	Coupling of ubiquitin conjugation to endoplasmic reticulum degradation
D_{max}	Maximum distance within the scattering particle
DNA	Deoxyribonucleic acid
DTDP	4,4'-dithiodipyridine
DTT	1,4-Dithiothreitol
DUIM	Double-sided UIM
EDC	N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride
EFG	Epidermal growth factor
e.g.	For example

ABBREVIATIONS

FID	Free induction decay
FOG	Friend of GATA
GAT	GGA and TOM
GATA	Transcription factors able to bind the DNA sequence “GATA”
GGA	Golgi-localized, gammear-containing, ADP-ribosylation-factor-binding protein
hetNOE	Hetero-nuclear Overhauser effect
HSQC	Hetero-nuclear single quantum correlation
HTLV-1	Human T-cell leukemia virus type-1
i.e.	That is
IL-1R	Interleukin-1 Receptor
IPTG	Isopropyl-B-D-Thiogalactopyranoside
ITC	Isothermal titration calorimetry
Itch	Itchy E3 ubiquitin protein ligase
IUIM	Inversed Ubiquitin Interacting Motif
JNK	c-jun N-terminal Kinase
Kan	Kanamycin
K_b	Binding constant
K_d	Dissociation constant
K_ds	Dissociation constants
K_{ex}	Exchange rate
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
LB	Lysogeny broth
L_T	Total ligand concentration
LUBAC	Linear ubiquitin chain assembly complex
MIU	Motif interacting with ubiquitin
M_T	Total protein concentration
Myb	Myeloblastosis
n	Number of binding sites
N-terminal	Amino terminal extreme of the protein
NEMO	NF-kB essential modulator
NHS	N-Hydroxysuccinimide
NMR	Nuclear magnetic resonance
NRP	NEMO-related protein

OD₆₀₀	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
pEDTA	N-[S-(2-pyridylthio)cysteaminy]ethylendiamine-N,N,N,N'-tetraacetic acid monoamine
PRE	Paramagnetic relaxation enhancement
Q_f	Q factor parameter
R	Ideal gases constant
R1	Longitudinal relaxation rate
R2	Transverse relaxation rate
R_g	Radius of gyration
RIP1	Kinase receptor interacting protein-1
r.m.s.d	Root-mean-square deviation
RNF11	Ring finger protein 11
rpm	Revolutions per minute
RTKs	Receptor tyrosine kinases
S²	Order parameter
S5a	Subunit 5a of the 26S proteasome
SAXS	Small angle x-ray scattering
SDS	Sodium dodecyl sulfate
SH3	Src homology 3
T	Temperature
t	Time
Tax1	HTLV-1 viral oncoprotein
TAX1BP1	Tax-1 binding protein 1
TB	Terrific broth
TCEP	Tris-(2-Carboxyethyl)posphine hydrochloride
TEK-box	⁹ TGK ¹¹ sequence in ubiquitin
TNF	Human necrosis factor
TOM	Target of Myb
TRAF	TNF-receptor associated factors
UBA	Ubiquitin-associated domain
UBD	Ubiquitin-binding domain
UBZ	Ubiquitin-binding zinc finger
UBZ1	First UBZ domain of TAX1BP1

ABBREVIATIONS

UBZ1+2	Tandem of UBZ1 and UBZ2 domains of TAX1BP1
UBZ2	Second UBZ domain of TAX1BP1
UGR	Universidad de Granada
UIM	Ubiquitin interacting motif
UV	Ultraviolet
V_c	Total cell volume
vdW	van der Waals repulsion term
VIS	Visible
Vps27p	Vacuolar protein sorting-associated protein 27
WT	Wild type
Z	Partition function
ZnF	Zinc finger domain
Γ_2	Transverse paramagnetic relaxation enhancement
ΔC_p	Heat capacity change
ΔG_b	Gibbs energy change of binding
ΔH_b	Enthalpy change of binding
ΔS_b	Entropy change of binding
$\Delta\delta$	Chemical shift difference
$\Delta\nu$	Frequency difference
ε	Molar extinction coefficient
τ_c	Overall rotational correlation time

REFERENCES

REFERENCES

- Ababou A, Pfuhl M, Ladbury JE. (2008) The binding stoichiometry of CIN85 SH3 domain A and cbl-b. *Nat Struct Mol Biol.* 15, 890-891; author reply 891-892.
- Ababou A, Pfuhl M, Ladbury JE. (2009) Novel Insights into the Mechanisms of CIN85 SH3 Domains Binding to Cbl Proteins: Solution-based investigations and in vivo implications. *J Mol Biol.* 387, 1120-36
- Adzhubei AA, Sternberg MJ. (1993) Left-handed polyproline II helices commonly occur in globular proteins. *J Mol Biol.* 229(2):472-93.
- Aghazadeh B, Rosen MK. (1999) Ligand recognition by SH3 and WW domains: the role of N-alkylation in PPII helices. *Chem Biol.* 6(9):R241-6. Review.
- Ahl IM, Jonsson BH, Tibell LA. (2009) Thermodynamic characterization of the interaction between the C-terminal domain of extracellular superoxide dismutase and heparin by isothermal titration calorimetry. *Biochemistry.* 48(41):9932-40.
- Alam SL, Sun J, Payne M, Welch BD, Blake BK, Davis DR, Meyer HH, Emr SD, Sundquist WI. (2004) Ubiquitin interactions of NZF zinc fingers. *EMBO J.* 7;23(7):1411-21.
- Alberts IL, Nadassy K, Wodak SJ. (1998) Analysis of zinc binding sites in protein crystal structures. *Protein Sci.* 7(8):1700-16.
- Aloy P, Russell RB. (2006) Structural systems biology: modelling protein interactions. *Nat Rev Mol Cell Biol.* 7(3):188-97.
- Argos P. (1988) An investigation of protein subunit and domain interfaces. *Protein Eng.* Jul;2(2):101-13
- Arold S, O'Brien R, Franken P, Strub MP, Hoh F, Dumas C, Ladbury JE. (1998) RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. *Biochemistry.* 37(42):14683-91.
- Auguin D, Barthe P, Royer C, Stern MH, Noguchi M, Arold ST, Roumestand C. (2004) Structural basis for the co-activation of protein kinase B by T-cell leukemia-1 (TCL1) family proto-oncoproteins. *J Biol Chem.* 279(34):35890-902.
- Bache KG, Slagsvold T, Stenmark H. (2004) Defective downregulation of receptor tyrosine kinases in cancer. *EMBO J.* 23(14):2707-12. Review.
- Bader S, Kühner S, Gavin AC. (2008) Interaction networks for systems biology. *FEBS Lett.* 582(8):1220-4. Review.
- Balbo A, Minor KH, Velikovsky CA, Mariuzza RA, Peterson CB, Schuck P. (2005) Studying multiprotein complexes by multisignal sedimentation velocity analytical ultracentrifugation. *Proc Natl Acad Sci U S A.* 102(1):81-6.
- Barbato G, Ikura M, Kay LE, Pastor RW, Bax A. (1992) Backbone dynamics of calmodulin studied by ¹⁵N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry.* 31(23):5269-78.
- Battiste JL, Wagner G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry.* 39(18):5355-65.
- Berg JM. (1988) Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. *Proc Natl Acad Sci U S A.* 85(1):99-102.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. (2000) The Protein Data Bank. *Nucleic Acids Res.* 28(1):235-42.
- Berry DM, Nash P, Liu SK, Pawson T, McGlade CJ. (2002) A high-affinity Arg-X-X-Lys SH3 binding motif confers specificity for the interaction between Gads and SLP-76 in T cell signaling. *Curr Biol.* 12(15):1336-41.
- Bertini I, Donaire A, Jiménez B, Luchinat C, Parigi G, Piccioli M, Poggi L. (2001) Paramagnetism-based versus classical constraints: an analysis of the solution structure of Ca Ln calbindin D9k. *J Biomol NMR.* 21(2):85-98.

- Bertini I, Luchinat C, Rosato A. (1996) The solution structure of paramagnetic metalloproteins. *Prog Biophys Mol Biol.* 66(1):43-80. Review.
- Bezsonova I, Bruce MC, Wiesner S, Lin H, Rotin D, Forman-Kay JD. (2008) Interactions between the three CIN85 SH3 domains and ubiquitin: implications for CIN85 ubiquitination. *Biochemistry.* 47(34):8937-49.
- Bhattacharyya RP, Reményi A, Good MC, Bashor CJ, Falick AM, Lim WA. (2006) The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science.* 311(5762):822-6.
- Bian M, Yu M, Yang S, Gao H, Huang Y, Deng C, Gao Y, Sun F, Huang F. (2008) Expression of Cbl-interacting protein of 85 kDa in MPTP mouse model of Parkinson's disease and 1-methyl-4-phenyl-pyridinium ion-treated dopaminergic SH-SY5Y cells. *Acta Biochim Biophys Sin (Shanghai).* 40(6):505-12.
- Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I. (2005) Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science.* 310(5755):1821-4.
- Bilodeau PS, Winistorfer SC, Allaman MM, Surendhran K, Kearney WR, Robertson AD, Piper RC. (2004) The GAT domains of clathrin-associated GGA proteins have two ubiquitin binding motifs. *J Biol Chem.* 279(52):54808-16.
- Bish RA, Myers MP. (2007) Werner helicase-interacting protein 1 binds polyubiquitin via its zinc finger domain. *J Biol Chem.* 282(32):23184-93.
- Bivona TG, Philips MR. (2003) Ras pathway signaling on endomembranes. *Curr Opin Cell Biol.* 15(2):136-42. Review.
- Blake-Hall J, Walker O, Fushman D. (2004) Characterization of the overall rotational diffusion of a protein from ¹⁵N relaxation measurements and hydrodynamic calculations. *Methods Mol Biol.* 278:139-60.
- Bögler O, Furnari FB, Kindler-Roehrborn A, Sykes VW, Yung R, Huang HJ, Cavenee WK. (2000) SETA: a novel SH3 domain-containing adapter molecule associated with malignancy in astrocytes. *Neuro Oncol.* 2(1):6-15.
- Bogler O. (2000) The glioma-associated protein SETA interacts with AIP1/Alix and ALG-2 and modulates apoptosis in astrocytes. *J. Biol. Chem.* 275 : 19275 – 192.
- Bomar MG, Pai MT, Tzeng SR, Li SS, Zhou P. (2007) Structure of the ubiquitin-binding zinc finger domain of human DNA Y-polymerase eta. *EMBO Rep.* 8(3):247-51.
- Bonvin AM, Boelens R, Kaptein R. (2005) NMR analysis of protein interactions. *Curr Opin Chem Biol.* 9(5):501-8. Review.
- Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C, Hurley P, Chien M, Chai S, Hitotsumatsu O, McNally E, Pickart C, Ma A. (2004) The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat Immunol.* 5(10):1052-60.
- Bosanac I, Wertz IE, Pan B, Yu C, Kusam S, Lam C, Phu L, Phung Q, Maurer B, Arnott D, Kirkpatrick DS, Dixit VM, Hymowitz SG. (2010) Ubiquitin binding to A20 ZnF4 is required for modulation of NF- κ B signaling. *Mol Cell.* 40(4):548-57.
- Boxus M, Twizere JC, Legros S, Dewulf JF, Kettmann R, Willems L. (2008) The HTLV-1 Tax interactome. *Retrovirology* 5: 76. Review.
- Brandman O, Meyer T. (2008) Feedback loops shape cellular signals in space and time. *Science.* 322(5900):390-5. Review.
- Brangwynne CP, Koenderink GH, MacKintosh FC, Weitz DA. (2008) Cytoplasmic diffusion: molecular motors mix it up. *J Cell Biol.* 183(4):583-7. Review.
- Braun W, Bösch C, Brown LR, Go N, Wüthrich K. (1981) Combined use of proton-proton Overhauser enhancements and a distance geometry algorithm for determination of polypeptide conformations. Application to micelle-bound glucagon. *Biochim Biophys Acta.* 667(2):377-96.
- Bray D, Levin MD, Morton-Firth CJ. (1998) Receptor clustering as a cellular mechanism to control sensitivity. *Nature.* 393(6680):85-8.
- Breslauer KJ, Freire E, Straume M. (1992) Calorimetry: a tool for DNA and ligand-DNA studies. *Methods Enzymol.* 211:533-67. Review.

REFERENCES

- Brett TJ, Traub LM, Fremont DH. (2002) Accessory protein recruitment motifs in clathrin-mediated endocytosis. *Structure*. 10(6):797-809.
- Bronowska AK, (2011) Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design (Book chapter), *Thermodynamics - Interaction Studies - Solids, Liquids and Gases*. InTech, ISBN 978-953-307-563-1.
- Brown A. (2009) Analysis of cooperativity by isothermal titration calorimetry. *Int J Mol Sci*. 10(8): 3457-3477.
- Bruck S, Huber TB, Ingham RJ, Kim K, Niederstrasser H, Allen PM, Pawson T, Cooper JA, Shaw AS. (2006) Identification of a novel inhibitory actin-capping protein binding motif in CD2-associated protein. *J Biol Chem*. 281(28):19196-203.
- Brulet A, Boue F, Cotton JP. (1996) About the experimental determination of the persistence length of wormlike chains of polystyrene. *Journal of Physics II (France)*.6, 855-891.
- Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 54(Pt 5):905-21.
- Buchman VL, Luke C, Borthwick EB, Gout I, Ninkina N. (2002) Organization of the mouse Ruk locus and expression of isoforms in mouse tissues. *Gene*. 295(1):13-17.
- Cárdenes N. (2007) Phd thesis: Structural basis for Cbl- induced SH3 oligomerization. Departamento de Biología Molecular, Universidad Autónoma de Madrid, Madrid. Spain.
- Carducci M, Perfetto L, Briganti L, Paoluzi S, Costa S, Zerweck J, Schutkowski M, Castagnoli L, Cesareni G. (2012) The protein interaction network mediated by human SH3 domains. *Biotechnol Adv*. 30(1):4-15.
- Casares S, Ab E, Eshuis H, Lopez-Mayorga O, van Nuland NA, Conejero-Lara F. (2007) The high-resolution NMR structure of the R21A Spc-SH3:P41 complex: understanding the determinants of binding affinity by comparison with Abl-SH3. *BMC Struct Biol*.7:22.
- Casares S, López-Mayorga O, Vega MC, Cámara-Artigas A, Conejero-Lara F. (2007) Cooperative propagation of local stability changes from low-stability and high-stability regions in a SH3 domain. *Proteins*. 67(3):531-47.
- Casares S. (2003) Phd thesis: Análisis de las rutas cooperativas de propagación de estabilidad en el dominio SH3. Aplicación al estudio de las interacciones proteína-ligando. Departamento de Química Física, Universidad de Granada. Granada. Spain.
- Ceregido MA, Garcia-Pino A, Ortega-Roldan JL, Casares S, López Mayorga O, Bravo J, van Nuland NA, Azuaga AI. (2013) Multimeric and differential binding of CIN85/CD2AP with two atypical proline-rich sequences from CD2 and Cbl-b. *FEBS J*. May 11. doi: 10.1111/febs.12333.
- Ceregido MA, Spinola M, Buts L, Garcia-Pino A, Rivera-Torres J, Bravo J, van Nuland NAJ. (2013) High-resolution structures of the UBZ1+2 tandem of the ubiquitin-binding adaptor protein TAX1BP1. Manuscript in preparation.
- Cesareni G, Panni S, Nardelli G, Castagnoli L. (2002) Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett*. 513(1):38-44.
- Chacón P, Díaz JF, Morán F, Andreu JM. (2000) Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. *J Mol Biol*. 299(5):1289-302.
- Chacón P, Morán F, Díaz JF, Pantos E, Andreu JM. (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys J*. 74(6):2760-75.
- Chan DC, Bedford MT, Leder P. (1996) Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J*. 15(5):1045-54.
- Chang CE, Chen W, Gilson MK. (2007) Ligand configurational entropy and protein binding. *Proc Natl Acad Sci U S A*. 104(5):1534-9.
- Chen B, Borinstein SC, Gillis J, Sykes VW, Bogler O. (2000) The glioma-associated protein SETA interacts with AIP1/Alix and ALG-2 and modulates apoptosis in astrocytes. *J Biol Chem*. 275(25):19275-81.

- Chen F, Castranova V, Shi X, Demers LM. (1999) New insights into the role of nuclear factor- κ B, a ubiquitous transcription factor in the initiation of diseases. *Clin Chem.* 45(1):7-17. Review.
- Chen ZJ, Sun LJ. (2009) Nonproteolytic functions of ubiquitin in cell signaling. *Mol Cell.* 33(3):275-86. Review.
- Chin KT, Chun AC, Ching YP, Jeang KT, Jin DY. (2007) Human T-cell leukemia virus oncoprotein tax represses nuclear receptor-dependent transcription by targeting coactivator TAX1BP1. *Cancer Res.* 67(3):1072-81.
- Chiu YH, Zhao M, Chen ZJ. (2009) Ubiquitin in NF- κ B signaling. *Chem Rev.* 109(4):1549-60. Review.
- Cho HJ, Lee S, Kim H. (2009) The linker connecting the tandem ubiquitin binding domains of RAP80 is critical for lysine 63-linked polyubiquitin-dependent binding activity. *BMB Rep.* 42(11):764-8.
- Choi JY, Ryu JH, Kim HS, Park SG, Bae KH, Kang S, Myung PK, Cho S, Park BC, Lee do H. (2007) Co-chaperone CHIP promotes aggregation of ataxin-1. *Mol Cell Neurosci.* 34(1):69-79.
- Chothia C, Janin J. (1975) Principles of protein-protein recognition. *Nature.* 256(5520):705-8.
- Chu ZL, Shin YA, Yang JM, DiDonato JA, Ballard DW. (1999) IKK γ mediates the interaction of cellular I κ B kinases with the tax transforming protein of human T cell leukemia virus type 1. *J Biol Chem.* 274(22):15297-300.
- Claudio E, Brown K, Park S, Wang H, Siebenlist U. (2002) BAFF-induced NEMO-independent processing of NF- κ B2 in maturing B cells. *Nat Immunol.* 3(10):958-653.
- Clore GM. (2013) Generating accurate contact maps of transient long-range interactions in intrinsically disordered proteins by paramagnetic relaxation enhancement. *Biophys J.* 104(8):1635-6.
- Clore GM, Iwahara J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem Rev.* 109(9):4108-39. Review.
- Clore GM, Tang C, Iwahara J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol.* 17(5):603-16. Review.
- Clore GM, Gronenborn AM. (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* 16(1):22-34. Review.
- Clore GM, Gronenborn AM. (1991) Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science.* 252(5011):1390-9.
- Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM. (1990) Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear ^{15}N - ^1H NMR spectroscopy. *Biochemistry.* 29(32):7387-401.
- Cole C, Barber JD, Barton GJ. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* 36(Web Server issue):W197-201.
- Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, Bertoni F, Ponzoni M, Scandurra M, Califano A, Bhagat G, Chadburn A, Dalla-Favera R, Pasqualucci L. (2009) Mutations of multiple genes cause deregulation of NF- κ B in diffuse large B-cell lymphoma. *Nature.* 459(7247):717-21.
- Compton LA, Johnson WC Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem.* 155(1):155-67.
- Coope HJ, Atkinson PG, Huhse B, Belich M, Janzen J, Holman MJ, Klaus GG, Johnston LH, Ley SC. (2002) CD40 regulates the processing of NF- κ B2 p100 to p52. *EMBO J.* 21(20):5375-85.
- Coornaert B, Carpentier I, Beyaert R. (2009) A20: central gatekeeper in inflammation and immunity. *J Biol Chem.* 284(13):8217-21. Review.
- Cordier F, Grubisha O, Traincard F, Véron M, Delepierre M, Agou F. (2009) The zinc finger of NEMO is a functional ubiquitin-binding domain. *J Biol Chem.* 284(5):2902-7.

REFERENCES

- Cordier F, Vinolo E, Véron M, Delepierre M, Agou F. (2008) Solution structure of NEMO zinc finger and impact of an anhidrotic ectodermal dysplasia with immunodeficiency-related point mutation. *J Mol Biol.* 11;377(5):1419-32
- Cormont M, Metón I, Mari M, Monzo P, Keslair F, Gaskin C, McGraw TE, Le Marchand-Brustel Y. (2003) CD2AP/CMS regulates endosome morphology and traffic to the degradative pathway through its interaction with Rab4 and c-Cbl. *Traffic.* 4(2):97-112.
- Cox EH, McLendon GL. (2000) Zinc-dependent protein folding. *Curr Opin Chem Biol.* 4(2):162-5. Review.
- Crosetto N, Bienko M, Hibbert RG, Perica T, Ambrogio C, Kensche T, Hofmann K, Sixma TK, Dikic I. (2008) Human Wrnip1 is localized in replication factories in a ubiquitin-binding zinc finger-dependent manner. *J Biol Chem.* 283(50):35173-85.
- Csiszár A. (2006) Structural and functional diversity of adaptor proteins involved in tyrosine kinase signalling. *Bioessays.* 28(5):465-79. Review.
- Dam J, Schuck P. (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: sedimentation coefficient distributions $c(s)$ and asymptotic boundary profiles from Gilbert-Jenkins theory. *Biophys J.* 89(1):651-66.
- Dam J, Velikovskiy CA, Mariuzza RA, Urbanke C, Schuck P. (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions $c(s)$. *Biophys J.* 89(1):619-34.
- Dam TK, Roy R, Das SK, Oscarson S, Brewer CF. (2000) Binding of multivalent carbohydrates to concanavalin A and Dioclea grandiflora lectin. Thermodynamic analysis of the "multivalency effect". *J Biol Chem.* 275(19):14223-30.
- De Valck D, Jin DY, Heyninck K, Van de Craen M, Contreras R, Fiers W, Jeang KT, Beyaert R. (1999) The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. *Oncogene.* 18(29):4182-90.
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277-293.
- Di Fiore PP, Polo S, Hofmann K. (2003) When ubiquitin meets ubiquitin receptors: a signalling connection. *Nat Rev Mol Cell Biol.* 4(6):491-7. Review.
- Dikic I, Wakatsuki S, Walters KJ. (2009) Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol.* 10(10):659-71. Review.
- Dikic I. (2002) CIN85/CMS family of adaptor molecules. *FEBS Lett.* 529(1):110-5. Review.
- Dix JA, Verkman AS. (2008) Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys.* 37:247-63. Review.
- Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32(Web Server issue):W665-7.
- Dosset P, Hus JC, Blackledge M, Marion D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *Journal of biomolecular NMR* 16 (1):23-28
- Dudev T, Lim C. (2003) Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem Rev.* 103(3):773-88. Review.
- Dueber JE, Yeh BJ, Chak K, Lim WA. (2003) Reprogramming control of an allosteric signaling switch through modular recombination. *Science.* 301(5641):1904-8.
- Dunitz JD. (1995) Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol.* 2(11):709-12. Review.
- Dunitz JD. (1994) The entropic cost of bound water in crystals and biomolecules. *Science.* 264(5159):670.
- Dustin ML, Olszowy MW, Holdorf AD, Li J, Bromley S, Desai N, Widder P, Rosenberger F, van der Merwe PA, Allen PM, Shaw AS. (1998) A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. *Cell.* 94(5):667-77.
- Dyson HJ, Wright PE. (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev.* 104(8):3607-22. Review.

- Elcock AH, Gabdoulline RR, Wade RC, McCammon JA. (1999) Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. *J Mol Biol.* 291(1):149-62.
- Elia AE, Cantley LC, Yaffe MB. (2003) Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates. *Science.* 299(5610):1228-31.
- Espejo A, Côté J, Bednarek A, Richard S, Bedford MT. (2002) A protein-domain microarray identifies novel protein-protein interactions. *Biochem J.* 367(Pt 3):697-702.
- Evans RM, Hollenberg SM. (1988) Zinc fingers: guilt by association. *Cell.* 52(1):1-3. Review.
- Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* 33 (19):5984-6003
- Feller SM, Tuchscherer G, Voss J. (2003) High affinity molecules disrupting GRB2 protein complexes as a therapeutic strategy for chronic myelogenous leukaemia. *Leuk Lymphoma.* 44(3):411-27. Review.
- Feller SM. (2001) Crk family adaptors-signalling complex formation and biological roles. *Oncogene.* 20(44):6348-71. Review.
- Feng S, Chen JK, Yu H, Simon JA, Schreiber SL. (1994) Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science.* 266(5188):1241-7.
- Fernandez-Ballester G, Blanes-Mira C, Serrano L. (2004) The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains. *J Mol Biol.* 335(2):619-29.
- Ferraro E, Peluso D, Via A, Ausiello G, Helmer-Citterich M. (2007) SH3-Hunter: discovery of SH3 domain interaction sites in proteins. *Nucleic Acids Res.* 35(Web Server issue):W451-4.
- Ferrell JE Jr. (2009) Q&A: Cooperativity. *J Biol.* 8(6):53.
- Ferreon JC, Hilser VJ. (2004) Thermodynamics of binding to SH3 domains: the energetic impact of polyproline II (PII) helix formation. *Biochemistry.* 43(24):7787-97.
- Ferreon JC, Hilser VJ. (2003) Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. *Protein Sci.* 12(5):982-96.
- Feshchenko EA, Langdon WY, Tsygankov AY. (1998) Fyn, Yes, and Syk phosphorylation sites in c-Cbl map to the same tyrosine residues that become phosphorylated in activated T cells. *J Biol Chem.* 273(14):8323-31.
- Fesik SW, Gampe RT Jr, Eaton HL, Gemmecker G, Olejniczak ET, Neri P, Holzman TF, Egan DA, Edalji R, Simmer R, et al. (1991) NMR studies of [U-¹³C]cyclosporin A bound to cyclophilin: bound conformation and portions of cyclosporin involved in binding. *Biochemistry.* 30(26):6574-83.
- Filimonov VV, Azuaga AI, Viguera AR, Serrano L, Mateo PL. (1999) A thermodynamic analysis of a family of small globular proteins: SH3 domains. *Biophys Chem.* 77(2-3):195-208.
- Finniss S, Movsisyan A, Billecke C, Schmidt M, Randazzo L, Chen B, Böglér O. (2004) Studying protein isoforms of the adaptor SETA/CIN85/Ruk with monoclonal antibodies. *Biochem Biophys Res Commun.* 325(1):174-82.
- Fogh RH, Schipper D, Boelens R, Kaptein R. (1994) (1)H, (13)C and (15)N NMR backbone assignments of the 269-residue serine protease PB92 from *Bacillus alcalophilus*. *J Biomol NMR.* 4(1):123-8.
- Ford DM. (2005) Enthalpy-entropy compensation is not a general feature of weak association. *J Am Chem Soc.* 127(46):16167-70.
- Fox AH, Liew C, Holmes M, Kowalski K, Mackay J, Crossley M. (1999) Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *EMBO J.* 18(10):2812-22.
- Freire E, Mayorga OL, Straume M. (1990) Isothermal titration calorimetry. *Anal Chem.* 62:950A-959A.

REFERENCES

- Fujiwara K, Tenno T, Sugasawa K, Jee JG, Ohki I, Kojima C, Tochio H, Hiroaki H, Hanaoka F, Shirakawa M. (2004) Structure of the ubiquitin-interacting motif of S5a bound to the ubiquitin-like domain of HR23B. *J Biol Chem.* 79(6):4760-7.
- Gabdoulline RR, Wade RC. (2001) Protein-protein association: investigation of factors influencing association rates by brownian dynamics simulations. *J Mol Biol.* 306(5):1139-55.
- Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. (2007) Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci.* 32(2):63-70. Review.
- Garrett DS, Seok YJ, Peterkofsky A, Clore GM, Gronenborn AM. (1997) Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. *Biochemistry.* 36(15):4393-8.
- Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, Wang HE, Wettstein DA, Stray KM, Côté M, Rich RL, Myszka DG, Sundquist WI. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell.* 107(1):55-65.
- Gerlach B, Cordier SM, Schmukle AC, Emmerich CH, Rieser E, Haas TL, Webb AI, Rickard JA, Anderton H, Wong WW, Nachbur U, Gangoda L, Warnken U, Purcell AW, Silke J, Walczak H. (2011) Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature.* 471(7340):591-6.
- Gibson TJ, Postma JP, Brown RS, Argos P. (1988) A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. *Protein Eng.* 2(3):209-18.
- Gill SJ, Murphy KP, Robed CH. (1990) Partition-Function Formalism for Analyzing Calorimetric Experiments. *J.Chem.Educ.* 67(11):928-931.
- Gmeiner WH, Horita DA. (2001) Implications of SH3 domain structure and dynamics for protein regulation and drug design. *Cell Biochem Biophys.* 35(2):127-40. Review.
- Gómez J, Hilser VJ, Xie D, Freire E. (1995) The heat capacity of proteins. *Proteins.* 22(4):404-12.
- Gordon, SL and Wüthrich, K. (1978) Transient proton-proton Overhauser effects in horse ferrocyanochrome *c*. *J. Am. Chem. Soc.* **100**, 7094–7096.
- Gout I, Middleton G, Adu J, Ninkina NN, Drobot LB, Filonenko V, Matsuka G, Davies AM, Waterfield M, Buchman VL. (2000) Negative regulation of PI 3-kinase by Ruk, a novel adaptor protein. *EMBO J.* 19(15):4015-25.
- Grabbe C, Dikic I. (2009) Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins. *Chem Rev.* 109(4):1481-94. Review.
- Graham TR, Zhou HE, Otero-Marah VA, Osunkoya AO, Kimbro KS, Tighiouart M, Liu T, Simons JW, O'Regan RM. (2008) Insulin-like growth factor-I-dependent up-regulation of ZEB1 drives epithelial-to-mesenchymal transition in human prostate cancer cells. *Cancer Res.* 68(7):2479-88.
- Grishaev A, Wu J, Trewthella J, Bax A. (2005) Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. *J Am Chem Soc.* 127(47):16621-8.
- Grivennikov SI, Greten FR, Karin M. (2010) Immunity, inflammation, and cancer. *Cell.* 140(6):883-99. Review.
- Grunkemeyer JA, Kwok C, Huber TB, Shaw AS. (2005) CD2-associated protein (CD2AP) expression in podocytes rescues lethality of CD2AP deficiency. *J Biol Chem.* 280(33):29677-81.
- Guijarro JI, Morton CJ, Plaxco KW, Campbell ID, Dobson CM. (1998) Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy. *J Mol Biol.* 276(3):657-67.
- Guillièrre F, Danioux C, Jaubert C, Desnoues N, Delepierre M, Prangishvili D, Sezonov G, Guijarro JI. (2013) Solution structure of an archaeal DNA binding protein with an eukaryotic zinc finger fold. *PLoS One.* 8(1):e52908.
- Guinier A, Fournet F. (1955) *Small Angle Scattering of X-Rays.* Wiley Interscience, New York.

- Guntert P, Mumenthaler C, Wuthrich K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol.* 273(1):283-298.
- Guntert P. (1997) Calculating protein structures from NMR data. *Methods Mol Biol* 60:157-194.
- Gurung R, Tan A, Ooms LM, McGrath MJ, Huysmans RD, Munday AD, Prescott M, Whisstock JC, Mitchell CA. (2003) Identification of a novel domain in two mammalian inositol-polyphosphate 5-phosphatases that mediates membrane ruffle localization. The inositol 5-phosphatase skip localizes to the endoplasmic reticulum and translocates to membrane ruffles following epidermal growth factor stimulation. *J Biol Chem.* 278(13):11376-85.
- Haas TL, Emmerich CH, Gerlach B, Schmukle AC, Cordier SM, Rieser E, Feltham R, Vince J, Warnken U, Wenger T, Koschny R, Komander D, Silke J, Walczak H. (2009) Recruitment of the linear ubiquitin chain assembly complex stabilizes the TNF-R1 signaling complex and is required for TNF-mediated gene induction. *Mol Cell.* 36(5):831-44.
- Habelhah H. (2010) Emerging complexity of protein ubiquitination in the NF- κ B pathway. *Genes Cancer.* 1(7):735-747.
- Häcker H, Karin M. (2006) Regulation and function of IKK and IKK-related kinases. *Sci STKE.* (357):re13. Review.
- Haglund K, Dikic I. (2005) Ubiquitylation and cell signaling. *EMBO J.* 24(19):3353-9. Review.
- Haglund K, Shimokawa N, Szymkiewicz I, Dikic I. (2002) Cbl-directed monoubiquitination of CIN85 is involved in regulation of ligand-induced degradation of EGF receptors. *Proc Natl Acad Sci U S A.* 99(19):12191-6.
- Hansen AP, Petros AM, Mazar AP, Pederson TM, Rueter A, Fesik SW. (1992) A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. *Biochemistry.* 31(51):12713-8.
- Harding MM. (2001) Geometry of metal-ligand interactions in proteins. *Acta Crystallogr D Biol Crystallogr.* 57(Pt 3):401-11.
- Harhaj EW, Sun SC. (1999) IKK γ serves as a docking subunit of the I κ B kinase (IKK) and mediates interaction of IKK with the human T-cell leukemia virus Tax protein. *J Biol Chem.* 274(33):22911-4.
- Harper JW, Schulman BA. (2006) Structural complexity in ubiquitin recognition. *Cell.* 124(6):1133-6. Review.
- Harris BZ, Lim WA. (2001) Mechanism and role of PDZ domains in signaling complex assembly. *J Cell Sci.* 114(Pt 18):3219-31. Review.
- Hartwig A, Asmuss M, Blessing H, Hoffmann S, Jahnke G, Khandelwal S, Pelzer A, Bürkle A. (2002) Interference by toxic metal ions with zinc-dependent proteins involved in maintaining genomic stability. *Food Chem Toxicol.* 40(8):1179-84. Review.
- Hartwig A. (2001) Zinc finger proteins as potential targets for toxic metal ions: differential effects on structure and function. *Antioxid Redox Signal.* 3(4):625-34. Review.
- Hartwig A. (1998) Carcinogenicity of metal compounds: possible role of DNA repair inhibition. *Toxicol Lett.* 102-103:235-9. Review.
- Hashimoto S, Hirose M, Hashimoto A, Morishige M, Yamada A, Hosaka H, Akagi K, Ogawa E, Oneyama C, Agatsuma T, Okada M, Kobayashi H, Wada H, Nakano H, Ikegami T, Nakagawa A, Sabe H. (2006) Targeting AMAP1 and cortactin binding bearing an atypical src homology 3/proline interface for prevention of breast cancer invasion and metastasis. *Proc Natl Acad Sci U S A.* 103(18):7036-41.
- Havrylov S, Redowicz MJ, Buchman VL. (2010) Emerging roles of Ruk/CIN85 in vesicle-mediated transport, adhesion, migration and malignancy. *Traffic.* 11(6):721-31. Review.
- Hayden MS, Ghosh S. (2004) Signaling to NF- κ B. *Genes Dev.* 18(18):2195-224. Review.
- He Y, Hicke L, Radhakrishnan I. (2007) Structural basis for ubiquitin recognition by SH3 domains. *J Mol Biol.* 373(1):190-6.

REFERENCES

- Hernandez-Valladares M, Kim T, Kannan B, Tung A, Aguda AH, Larsson M, Cooper JA, Robinson RC. (2010) Structural characterization of a capping protein interaction motif defines a family of actin filament regulators. *Nat Struct Mol Biol.* 17(4):497-503.
- Herrmann T, Guntert P, Wuthrich K (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of biomolecular NMR* 24 (3):171-189.
- Hershko A, Ciechanover A. (1998) The ubiquitin system. *Annu Rev Biochem.* 67:425-79. Review.
- Hicke L, Schubert HL, Hill CP. (2005) Ubiquitin-binding domains. *Nat Rev Mol Cell Biol.* 6(8):610-21. Review.
- Hicke L, Dunn R. (2003) Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol.* 19:141-72. Review.
- Hicke L. (2001) Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol.* 2(3):195-201. Review.
- Hitotsumatsu O, Ahmad RC, Tavares R, Wang M, Philpott D, Turer EE, Lee BL, Shiffin N, Advincula R, Malynn BA, Werts C, Ma A. (2008) The ubiquitin-editing enzyme A20 restricts nucleotide-binding oligomerization domain containing 2-triggered signals. *Immunity.* 28(3):381-90.
- Hoffman DW, Cameron CS, Davies C, White SW, Ramakrishnan V. (1996) Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy. *J Mol Biol.* 264(5):1058-71.
- Hofmann K. (2009) Ubiquitin-binding domains and their role in the DNA damage response. *DNA Repair (Amst).* 8(4):544-56. Review.
- Hofmann K, Falquet L. (2001) A ubiquitin-interacting motif conserved in components of the proteasomal and lysosomal protein degradation systems. *Trends Biochem Sci.* 26(6):347-50.
- Honma I, Torigoe T, Hirohashi Y, Kitamura H, Sato E, Masumori N, Tamura Y, Tsukamoto T, Sato N. (2009) Aberrant expression and potency as a cancer immunotherapy target of alpha-methylacyl-coenzyme A racemase in prostate cancer. *J Transl Med.* 7:103.
- Houben K, Marion D, Tarbouriech N, Ruigrok RW, Blanchard L. (2007) Interaction of the C-terminal domains of sendai virus N and P proteins: comparison of polymerase-nucleocapsid interactions within the paramyxovirus family. *J Virol.* 81(13):6807-16.
- Houtman JC, Brown PH, Bowden B, Yamaguchi H, Appella E, Samelson LE, Schuck P. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.* 16(1):30-42.
- Huang TT, D'Andrea AD. (2006) Regulation of DNA repair by ubiquitylation. *Nat Rev Mol Cell Biol.* May;7(5):323-34. Review.
- Huber TB, Hartleben B, Kim J, Schmidts M, Schermer B, Keil A, Egger L, Lecha RL, Borner C, Pavenstädt H, Shaw AS, Walz G, Benzing T. (2003) Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling. *Mol Cell Biol.* 23(14):4917-28.
- Hurley JH, Lee S, Prag G. (2006) Ubiquitin-binding domains. *Biochem J.* 399(3):361-72. Review.
- Hutchings NJ, Clarkson N, Chalkley R, Barclay AN, Brown MH. (2003) Linking the T cell surface protein CD2 to the actin-capping protein CAPZ via CMS and CIN85. *J Biol Chem.* 278(25):22396-403.
- Iha H, Peloponese JM, Verstrepen L, Zapart G, Ikeda F, Smith CD, Starost MF, Yedavalli V, Heyninck K, Dikic I, Beyaert R, Jeang KT. (2008) Inflammatory cardiac valvulitis in TAX1BP1-deficient mice through selective NF-kappaB activation. *EMBO J.* 27(4):629-41.
- Iha H, Kasai T, Kibler KV, Iwanaga Y, Tsurugi K, Jeang KT. (2000) Pleiotropic effects of HTLV type 1 Tax protein on cellular metabolism: mitotic checkpoint abrogation and NF-kappaB activation. *AIDS Res Hum Retroviruses.* 16(16):1633-8.

- Ikeda F, Deribe YL, Skånland SS, Stieglitz B, Grabbe C, Franz-Wachtel M, van Wijk SJ, Goswami P, Nagy V, Terzic J, Tokunaga F, Androulidaki A, Nakagawa T, Pasparakis M, Iwai K, Sundberg JP, Schaefer L, Rittinger K, Macek B, Dikic I. (2011) SHARPIN forms a linear ubiquitin ligase complex regulating NF- κ B activity and apoptosis. *Nature*. 471(7340):637-41.
- Ikura M, Barbato G, Klee CB, Bax A. (1992) Solution structure of calmodulin and its complex with a myosin light chain kinase fragment. *Cell Calcium*. 13(6-7):391-400.
- Israël A. (2010) The IKK complex, a central regulator of NF- κ B activation. *Cold Spring Harb Perspect Biol*. 2(3):a000158. Review.
- Iuchi S. (2001) Three classes of C2H2 zinc finger proteins. *Cell Mol Life Sci*. 58(4):625-35. Review.
- Iwahara J, Clore GM. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature*. 440(7088):1227-30.
- Iwahara J, Schwieters CD, Clore GM. (2004) Ensemble approach for NMR structure refinement against $(1)H$ paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J Am Chem Soc*. 126(18):5879-96.
- Iwahara J, Anderson DE, Murphy EC, Clore GM. (2003) EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J Am Chem Soc*. 125(22):6634-5.
- Iwai K, Tokunaga F. (2009) Linear polyubiquitination: a new regulator of NF- κ B activation. *EMBO Rep*. 10(7):706-13. Review.
- Iwasaki A, Medzhitov R. (2010) Regulation of adaptive immunity by the innate immune system. *Science*. 327(5963):291-5. Review.
- Jarymowycz VA, Stone MJ. (2006) Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chem Rev*. 106(5):1624-71. Review.
- Jelesarov I, Bosshard HR. (1999) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit*. 12(1):3-18.
- Jia CY, Nie J, Wu C, Li C, Li SS. (2005) Novel Src homology 3 domain-binding motifs identified from proteomic screen of a Pro-rich region. *Mol Cell Proteomics*. 4(8):1155-66.
- Jin DY, Giordano V, Kibler KV, Nakano H, Jeang KT. (1999) Role of adapter function in oncoprotein-mediated activation of NF- κ B. Human T-cell leukemia virus type I Tax interacts directly with IkappaB kinase gamma. *J Biol Chem*. 274(25):17402-5.
- Johnson BA. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol*. 278:313-52.
- Journo C, Filipe J, About F, Chevalier SA, Afonso PV, Brady JN, Flynn D, Tangy F, Israël A, Vidalain PO, Mahieux R, Weil R. (2009) NRP/Optineurin Cooperates with TAX1BP1 to potentiate the activation of NF- κ B by human T-lymphotropic virus type 1 tax protein. *PLoS Pathog*. 5(7):e1000521.
- Jozic D, Cárdenes N, Deribe YL, Moncalián G, Hoeller D, Groemping Y, Dikic I, Rittinger K, Bravo J. (2008) Reply to "The binding stoichiometry of CIN85 SH3 domain A and Cbl-b". *Nat Struct Mol Biol*. 15(9):891-2.
- Jozic D, Cárdenes N, Deribe YL, Moncalián G, Hoeller D, Groemping Y, Dikic I, Rittinger K, Bravo J. (2005) Cbl promotes clustering of endocytic adaptor proteins. *Nat Struct Mol Biol*. 12(11):972-9.
- Kaneko T, Li L, Li SS. (2008) The SH3 domain--a family of versatile peptide- and protein-recognition module. *Front Biosci*. 13:4938-52. Review.
- Kardinal C, Konkol B, Schulz A, Posern G, Lin H, Adermann K, Eulitz M, Estrov Z, Talpaz M, Arlinghaus RB, Feller SM. (2000) Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. *FASEB J*. 14(11):1529-38.
- Kardinal C, Posern G, Zheng J, Knudsen BS, Moarefi I, Feller SM. (1999) Rational development of cell-penetrating high affinity SH3 domain binding peptides that

REFERENCES

- selectively disrupt the signal transduction of Crk family adapters. Amgen Peptide Technology Group. *Ann N Y Acad Sci.* 886:289-92. Review.
- Karin M, Greten FR. (2005) NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol.* 5(10):749-59. Review.
- Karin M, Ben-Neriah Y. (2000) Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu Rev Immunol.* 18:621-63. Review.
- Kato M, Ohshima K, Mizuno M, Kyogoku M, Hashikawa K, Tokura Y, Miyachi Y, Kabashima K. (2009) Analysis of CXCL9 and CXCR3 expression in a case of intravascular large B-cell lymphoma. *J Am Acad Dermatol.* 61(5):888-91.
- Kay BK, Kasanov J, Yamabhai M. (2001) Screening phage-displayed combinatorial peptide libraries. *Methods.* 24(3):240-6. Review.
- Kay BK, Kasanov J, Knight S, Kurakin A. (2000) Convergent evolution with combinatorial peptides. *FEBS Lett.* 480(1):55-62. Review.
- Kay LE. (1998) Protein dynamics from NMR. *Biochem Cell Biol.* 76(2-3):145-52. Review.
- Kay LE, Torchia DA, Bax A. (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry.* 28(23):8972-9.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28(12):1647-9.
- Keizers PH, Ubbink M. (2011) Paramagnetic tagging for protein structure and dynamics analysis. *Prog Nucl Magn Reson Spectrosc.* 58(1-2):88-96. Review.
- Kim JH, Li H, Stallcup MR. (2003) CoCoA, a nuclear receptor coactivator which acts through an N-terminal activation domain of p160 coactivators. *Mol Cell.* 12(6):1537-49.
- Kirisako T, Kamei K, Murata S, Kato M, Fukumoto H, Kanie M, Sano S, Tokunaga F, Tanaka K, Iwai K. (2006) A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J.* 25(20):4877-87. Epub 2006 Sep 28.
- Kirsch KH, Georgescu MM, Shishido T, Langdon WY, Birge RB, Hanafusa H. (2001) The adapter type protein CMS/CD2AP binds to the proto-oncogenic protein c-Cbl through a tyrosine phosphorylation-regulated Src homology 3 domain interaction. *J Biol Chem.* 276(7):4957-63.
- Kirsch KH, Georgescu MM, Ishimaru S, and Hanafusa H. (1999) CMS: an adapter molecule involved in cytoskeletal rearrangements. *Proc. Natl. Acad. Sci. USA* 96, 6211-6216.
- Kishan KV, Newcomer ME, Rhodes TH, Guillot SD. (2001) Effect of pH and salt bridges on structural assembly: molecular structures of the monomer and intertwined dimer of the Eps8 SH3 domain. *Protein Sci.* 10(5):1046-55.
- Kiyokawa E, Mochizuki N, Kurata T, Matsuda M. (1997) Role of Crk oncogene product in physiologic signaling. *Crit Rev Oncog.* 8(4):329-42. Review.
- Klejman A, Schreiner SJ, Nieborowska-Skorska M, Slupianek A, Wilson M, Smithgall TE, Skorski T. (2002) The Src family kinase Hck couples BCR/ABL to STAT5 activation in myeloid leukemia cells. *EMBO J.* 21(21):5766-74.
- Kleckner IR, Foster MP. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta.* 1814(8):942-68. Review.
- Klug A. (1999) Zinc finger peptides for the regulation of gene expression. *J Mol Biol.* 293(2):215-8. Review.
- Klug A, Rhodes D. (1987) Zinc fingers: a novel protein fold for nucleic acid recognition. *Cold Spring Harb Symp Quant Biol.* 52:473-82.
- Kobayashi S, Sawano A, Nojima Y, Shibuya M, Maru Y. (2004) The c-Cbl/CD2AP complex regulates VEGF-induced endocytosis and degradation of Flt-1 (VEGFR-1). *FASEB J.* 18(7):929-31.
- Koch MH, Vachette P, Svergun DI. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev Biophys.* 36(2):147-227. Review.

- Kometani K, Yamada T, Sasaki Y, Yokosuka T, Saito T, Rajewsky K, Ishiai M, Hikida M, Kurosaki T. (2011) CIN85 drives B cell responses by linking BCR signals to the canonical NF-kappaB pathway. *J Exp Med.* 208(7):1447-57.
- Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Cryst.* 39, 277–286
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst.* 36, 1277–1282
- Konev SV. (1967) Fluorescence and Phosphorescence of Proteins and Nucleic Acids. Plenum Press, New York (USA).
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph.* 14(1): 51-5, 29-32.
- Korzhev DM, Bocharov EV, Zhuravlyova AV, Orekhov VY, Ovchinnikova TV, Billeter M, Arseniev AS. (2001) Backbone dynamics of the channel-forming antibiotic zervamicin IIB studied by ¹⁵N NMR relaxation. *FEBS Lett.* 495(1-2):52-5.
- Kowanetz K, Szymkiewicz I, Haglund K, Kowanetz M, Husnjak K, Taylor JD, Soubeyran P, Engstrom U, Ladbury JE, Dikic I. (2003) Identification of a novel proline-arginine motif involved in CIN85-dependent clustering of Cbl and down-regulation of epidermal growth factor receptors. *J Biol Chem.* 278(41):39735-46.
- Krishna SS, Majumdar I, Grishin NV. (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res.* 31(2):532-50.
- Krizek BA, Zawadzke LE, Berg JM. (1993) Independence of metal binding between tandem Cys2His2 zinc finger domains. *Protein Sci.* 2(8):1313-9.
- Kukić P, Nielsen JE. (2010) Electrostatics in proteins and protein-ligand complexes. *Future Med Chem. Review.*
- Kurakin AV, Wu S, Bredesen DE. (2003) Atypical recognition consensus of CIN85/SETA/Ruk SH3 domains revealed by target-assisted iterative screening. *J Biol Chem.* 278(36):34102-9.
- Kwan AH, Mobli M, Gooley PR, King GF, Mackay JP. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J.* 278(5):687-703. Review.
- L'Espérance S, Popa I, Bachvarova M, Plante M, Patten N, Wu L, Têtu B, Bachvarov D. (2006) Gene expression profiling of paired ovarian tumors obtained prior to and following adjuvant chemotherapy: molecular signatures of chemoresistant tumors. *Int J Oncol.* 29(1):5-24.
- Laitly JH, Lee BM, Wright PE. (2001) Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol.* 11(1):39-46. Review.
- Larson SM, Davidson AR. (2000) The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. *Protein Sci.* 9(11):2170-80.
- Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8: 477-486.
- Leavitt S, Freire E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr Opin Struct Biol.* 11, 560-66
- Leckband D. (2000) Measuring the forces that control protein interactions. *Annu Rev Biophys Biomol Struct.* 29:1-26. Review.
- Lee KY, Hyeok Yoon JH, Kim M, Roh S, Lee YS, Seong BL, Kim K. (2002) A dipalmitoyl peptide that binds SH3 domain, disturbs intracellular signal transduction, and inhibits tumor growth in vivo. *Biochem Biophys Res Commun.* 296(2):434-42.
- Lee MS, Cavanagh J, Wright PE. (1989) Complete assignment of the ¹H NMR spectrum of a synthetic zinc finger from Xfin. Sequential resonance assignments and secondary structure. *FEBS Lett.* 254(1-2):159-64.
- Lee MS, Kim KA. (2006) NF-kB Pathway in Metabolic/Endocrine Diseases. *J Korean Endocr Soc.* 21(5):352-363.
- Lee S, Tsai YC, Mattera R, Smith WJ, Kostelansky MS, Weissman AM, Bonifacino JS, Hurley JH. (2006) Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5. *Nat Struct Mol Biol.* 13(3):264-71.

REFERENCES

- Levy ED, Pereira-Leal JB. (2008) Evolution and dynamics of protein interactions and networks. *Curr Opin Struct Biol.* 18(3):349-57. Review.
- Li SS. (2005) Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J.* 390(Pt 3):641-53. Review.
- Liew CK, Simpson RJ, Kwan AH, Crofts LA, Loughlin FE, Matthews JM, Crossley M, Mackay JP. (2005) Zinc fingers as protein recognition motifs: structural basis for the GATA-1/friend of GATA interaction. *Proc Natl Acad Sci U S A.* 102(3):583-8.
- Liew CK, Kowalski K, Fox AH, Newton A, Sharpe BK, Crossley M, Mackay JP. (2000) Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that mediate protein-protein interactions. *Structure.* 8(11):1157-66.
- Lim WA, Richards FM, Fox RO. (1994) Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature.* 372(6504):375-9. Erratum in: *Nature* 1995 Mar 2;374(6517):94.
- Ling L, and Goeddel DV. (2000) T6BP, a TRAF6-interacting protein involved in IL-1 signaling. *Proc. Natl. Acad. Sci. USA.* 97, 9567-9572
- Lipari G, and Szabo A. (1982a). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules I. Theory and range of validity. *J. Am. Chem. Soc.*, 104(17), 4546-4559.
- Lipari G, and Szabo A. (1982b). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules II. Analysis of experimental results. *J. Am. Chem. Soc.*, 104(17), 4559-4570.
- Liu S, Chen Y, Li J, Huang T, Tarasov S, King A, Weissman AM, Byrd RA, Das R. Promiscuous interactions of gp78 E3 ligase CUE domain with polyubiquitin chains. *Structure.* 2012 Dec 5;20(12):2138-50. doi: 10.1016/j.str.2012.09.020. Epub 2012 Nov 1.
- Lustbader JW, Birken S, Pollak S, Pound A, Chait BT, Mirza UA, Ramnarain S, Canfield RE, Brown JM. (1996) Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: an advance toward rapid determination of glycoprotein structures. *J Biomol NMR.* 7(4):295-304.
- Lynch DK, Winata SC, Lyons RJ, Hughes WE, Lehrbach GM, Wasinger V, Corthals G, Cordwell S, Daly RJ. (2003) A Cortactin-CD2-associated protein (CD2AP) complex provides a novel link between epidermal growth factor receptor endocytosis and the actin cytoskeleton. *J Biol Chem.* 278(24):21805-13.
- Mackay JP, Crossley M. (1998) Zinc fingers are sticking together. *Trends Biochem Sci.* 23(1):1-4. Review.
- Malmendal A, Evenäs J, Forsén S, Akke M. (1999) Structural dynamics in the C-terminal domain of calmodulin at low calcium levels. *J Mol Biol.* 293(4):883-99.
- Mandal PK, Majumdar A. (2004) A comprehensive discussion of HSQC and HMQC pulse sequence. *Concepts in Magn. Reson.* 20A, 1-23.
- Maret W. (1995) Metallothionein/disulfide interactions, oxidative stress, and the mobilization of cellular zinc. *Neurochem Int.* 27(1):111-7.
- Maret W. (1994) Oxidative metal release from metallothionein via zinc-thiol/disulfide interchange. *Proc Natl Acad Sci U S A.* 91(1):237-41.
- Marois L, Vaillancourt M, Paré G, Gagné V, Fernandes MJ, Rollet-Labelle E, Naccache PH. (2011) CIN85 modulates the down-regulation of Fc gammaRIIa expression and function by c-Cbl in a PKC-dependent manner in human neutrophils. *J Biol Chem.* 286(17):15073-84.
- Martin-Garcia JM, Luque I, Ruiz-Sanz J, Camara-Artigas A. (2012) The promiscuous binding of the Fyn SH3 domain to a peptide from the NS5A protein. *Acta Crystallogr D Biol Crystallogr.* 68(Pt 8):1030-40.
- Matsuoka M, Jeang KT. (2005) Human T-cell leukemia virus type I at age 25: a progress report. *Cancer Res.* 65(11):4467-70. Review.
- Matsuoka M, Jeang KT. (2007) Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer.* 7(4):270-80. Review.

- Matthews JM, Sunde M. (2002) Zinc fingers--folds for many occasions. *IUBMB Life*. 54(6):351-5. Review.
- Matthews JM, Kowalski K, Liew CK, Sharpe BK, Fox AH, Crossley M, MacKay JP. (2000) A class of zinc fingers involved in protein-protein interactions biophysical characterization of CCHC fingers from fog and U-shaped. *Eur J Biochem*. 267(4):1030-8.
- Mauro C, Pacifico F, Lavorgna A, Mellone S, Iannetti A, Acquaviva R, Formisano S, Vito P, Leonardi A. (2006) ABIN-1 binds to NEMO/IKKgamma and co-operates with A20 in inhibiting NF-kappaB. *J Biol Chem*. 281(27):18482-8.
- Mayer BJ. (2001) SH3 domains: complexity in moderation. *J Cell Sci*. 114(Pt 7):1253-63. Review.
- Mayer BJ. (1999) Endocytosis: EH domains lend a hand. *Curr Biol*. 9(2):R70-3. Review.
- Mayer BJ, Hamaguchi M, Hanafusa H. (1988) A novel viral oncogene with structural similarity to phospholipase C. *Nature*. 332(6161):272-5.
- Maynard AT, Covell DG. (2001) Reactivity of zinc finger cores: analysis of protein packing and electrostatic screening. *J Am Chem Soc*. 123(6):1047-58.
- McColl DJ, Honchell CD, Frankel AD. (1999) Structure-based design of an RNA-binding zinc finger. *Proc Natl Acad Sci U S A*. 96(17):9521-6.
- McDonald CB, Seldeen KL, Deegan BJ, Farooq A. (2009) SH3 domains of Grb2 adaptor bind to PXpsiPXR motifs within the Sos1 nucleotide exchange factor in a discriminate manner. *Biochemistry*. 48(19):4074-85.
- McIntosh LP, Dahlquist FW. (1990) Biosynthetic incorporation of ¹⁵N and ¹³C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q Rev Biophys*. 23(1):1-38. Review.
- McIntosh LP, Griffey RH, Muchmore DC, Nielson CP, Redfield AG, Dahlquist FW. (1987) Proton NMR measurements of bacteriophage T4 lysozyme aided by ¹⁵N isotopic labeling: structural and dynamic studies of larger proteins. *Proc Natl Acad Sci U S A*. 84(5):1244-8.
- McKinnon IR, Fall L, Parody-Morreale A, Gill SJ. (1984) A twin titration microcalorimeter for the study of biochemical reactions. *Anal Biochem*. 139(1):134-9.
- McPherson PS. (1999) Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. *Cell Signal*. 11(4):229-38. Review.
- Metropolis N, Ulam S. (1949) The Monte Carlo method. *J Am Stat Assoc*. 44(247):335-41.
- Miller J, McLachlan AD, Klug A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J*. 4(6):1609-14.
- Mittermaier AK, Kay LE. (2009) Observing biological dynamics at atomic resolution using NMR. *Trends Biochem Sci*. 34(12):601-11. Review.
- Mizuguchi M, Asao H, Hara T, Higuchi M, Fujii M, Nakamura M. (2009) Transcriptional activation of the interleukin-21 gene and its receptor gene by human T-cell leukemia virus type 1 Tax in human T-cells. *J Biol Chem*. 284(38):25501-11.
- Mizuguchi M, Kroon GJ, Wright PE, Dyson HJ. (2003) Folding of a beta-sheet protein monitored by real-time NMR spectroscopy. *J Mol Biol*. 328(5):1161-71.
- Moncalián G, Cárdenes N, Deribe YL, Spínola-Amilibia M, Dikic I, Bravo J. (2006) Atypical polyproline recognition by the CMS N-terminal Src homology 3 domain. *J Biol Chem*. 281(50):38845-53.
- Mongiovi AM, Romano PR, Panni S, Mendoza M, Wong WT, Musacchio A, Cesareni G, Di Fiore PP. (1999) A novel peptide-SH3 interaction. *EMBO J*. 18(19):5300-9.
- Montaville P, Coudeville N, Radhakrishnan A, Leonov A, Zweckstetter M, Becker S. (2008) The PIP2 binding mode of the C2 domains of rabphilin-3A. *Protein Sci*. 17(6):1025-34.
- Monzo P, Cormont M. (2005) CD2AP, a molecule adaptor between endocytosis and actin cytoskeleton in the cell in interphase and during cytokinesis?. *Med Sci (Paris)*. 21(12):1036-7.
- Morriswood B, Ryzhakov G, Puri C, Arden SD, Roberts R, Dendrou C, Kendrick-Jones J, Buss F. (2007) T6BP and NDP52 are myosin VI binding partners with potential roles in cytokine signalling and cell adhesion. *J Cell Sci*. 120(Pt 15):2574-85.

REFERENCES

- Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol.* 177:44-73.
- Mueller TD, Kamionka M, Feigon J. (2004) Specificity of the interaction between ubiquitin-associated domains and ubiquitin. *J Biol Chem.* 279(12):11926-36.
- Murphy KP, Freire E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv Protein Chem.* 43:313-61. Review.
- Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA. (2008) Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat Genet.* 40(9):1062-4.
- Myszka DG, Abdiche YN, Arisaka F, Byron O, Eisenstein E, Hensley P, Thomson JA, Lombardo CR, Schwarz F, Stafford W, Doyle ML. (2003) The ABRF-MIRG'02 study: assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction. *J Biomol Tech.* 14(4):247-69.
- Nagaraja GM, Kandpal RP. (2004) Chromosome 13q12 encoded Rho GTPase activating protein suppresses growth of breast carcinoma cells, and yeast two-hybrid screen shows its interaction with several proteins. *Biochem Biophys Res Commun.* 313(3):654-65.
- Narayan VA, Kriwacki RW, Caradonna JP. (1997) Structures of zinc finger domains from transcription factor Sp1. Insights into sequence-specific protein-DNA recognition. *J Biol Chem.* 272(12):7801-9.
- Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, Nabuurs SB, Guntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AM (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins.* 59 (4):662-672.
- Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, Dugger D, Gordon N, Sidhu SS, Fellouse FA, Komuves L, French DM, Ferrando RE, Lam C, Compaan D, Yu C, Bosanac I, Hymowitz SG, Kelley RF, Dixit VM. (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell.* 134(4):668-78.
- Nguyen JT, Turck CW, Cohen FE, Zuckermann RN, Lim WA. (1998) Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science.* 282(5396):2088-92.
- Nguyen TH, Birraux J, Wildhaber B, Myara A, Trivin F, Le Coultre C, Trono D, Chardot C. (2006) Ex vivo lentivirus transduction and immediate transplantation of uncultured hepatocytes for treating hyperbilirubinemic Gunn rat. *Transplantation.* 82(6):794-803.
- Nickerson NN, Tosi T, Dessen A, Baron B, Raynal B, England P, Pugsley AP. (2011) Outer membrane targeting of secretin PulD protein relies on disordered domain recognition by a dedicated chaperone. *J Biol Chem.* 286(45):38833-43.
- Niebuhr K, Ebel F, Frank R, Reinhard M, Domann E, Carl UD, Walter U, Gertler FB, Wehland J, Chakraborty T. (1997) A novel proline-rich motif present in ActA of *Listeria monocytogenes* and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *EMBO J.* 16(17):5433-44.
- Niuro H, Jabbarzadeh-Tabrizi S, Kikushige Y, Shima T, Noda K, Ota S, Tsuzuki H, Inoue Y, Arinobu Y, Iwasaki H, Shimoda S, Baba E, Tsukamoto H, Horiuchi T, Taniyama T, Akashi K. (2012) CIN85 is required for Cbl-mediated regulation of antigen receptor signaling in human B cells. *Blood.* 119(10):2263-73.
- Nishida M, Nagata K, Hachimori Y, Horiuchi M, Ogura K, Mandiyan V, Schlessinger J, Inagaki F. (2001) Novel recognition mode between Vav and Grb2 SH3 domains. *EMBO J.* 20(12):2995-3007.
- Nooren IM, Thornton JM. (2003) Diversity of protein-protein interactions. *EMBO J.* 22(14):3486-92. Review.
- Nooren IM, Thornton JM. (2003) Structural characterisation and functional significance of transient protein-protein interactions. *J Mol Biol.* 325(5):991-1018.
- Novak AJ, Slager SL, Fredericksen ZS, Wang AH, Manske MM, Ziesmer S, Liebow M, Macon WR, Dillon SR, Witzig TE, Cerhan JR, Ansell SM. (2009) Genetic variation in B-cell-

- activating factor is associated with an increased risk of developing B-cell non-Hodgkin lymphoma. *Cancer Res.* 69(10):4217-24.
- Ohno A, Jee J, Fujiwara K, Tenno T, Goda N, Tochio H, Kobayashi H, Hiroaki H, Shirakawa M. (2005) Structure of the UBA domain of Dsk2p in complex with ubiquitin molecular determinants for ubiquitin recognition. *Structure.* 13(4):521-32.
- Ortega-Roldan JL. (2010) Phd thesis: Structure and themodynamics of the CD2AP-SH3 domains and their interaction with ubiquitin. Departamento de Química Física, Universidad de Granada. Granada. Spain.
- Ortega-Roldan JL, Jensen MR, Brutscher B, Azuaga AI, Blackledge M, van Nuland NA. (2009) Accurate characterization of weak macromolecular interactions by titration of NMR residual dipolar couplings: application to the CD2AP SH3-C:ubiquitin complex. *Nucleic Acids Res.* 37(9):e70.
- Ortega Roldan JL, Romero Romero ML, Ora A, Ab E, Lopez Mayorga O, Azuaga AI, van Nuland NA. (2007) The high resolution NMR structure of the third SH3 domain of CD2AP. *J Biomol NMR.* 39(4):331-6.
- Otting G. (2010) Protein NMR using paramagnetic ions. *Annu Rev Biophys.* 39:387-405. Review.
- Ottinger EA, Botfield MC, Shoelson SE. (1998) Tandem SH2 domains confer high specificity in tyrosine kinase signaling. *J Biol Chem.* 273(2):729-35.
- Palencia A, Cobos ES, Mateo PL, Martínez JC, Luque I. (2004) Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: implications for rational ligand design. *J Mol Biol.* 336(2):527-37.
- Palmer AG, Williams JC, McDermott AE. (1996) Nuclear Magnetic Resonance Studies of Biopolymer Dynamics, *J. Phys. Chem.* 100:13293-13310.
- Palmer AG, Rance M, Wright PE. (1991) Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance C-12 heteronuclear NMR-spectroscopy. *J Am Chem Soc* 113(12):4371–4380.
- Parker MH, Lunney EA, Ortwine DF, Pavlovsky AG, Humblet C, Brouillette CG. (1999) Analysis of the binding of hydroxamic acid and carboxylic acid inhibitors to the stromelysin-1 (matrix metalloproteinase-3) catalytic domain by isothermal titration calorimetry. *Biochemistry.* 38(41):13592-601.
- Parvatiyar K, Barber GN, Harhaj EW. (2010) TAX1BP1 and A20 inhibit antiviral signaling by targeting TBK1-IKKi kinases. *J Biol Chem.* 285(20):14999-5009.
- Pascal JM, Tsodikov OV, Hura GL, Song W, Cotner EA, Classen S, Tomkinson AE, Tainer JA, Ellenberger T. (2006) A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol Cell.* 24(2):279-91.
- Pavletich NP, Pabo CO. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science.* 252(5007):809-17.
- Pawson T. (2003) Organization of cell-regulatory systems through modular-protein-interaction domains. *Philos Transact A Math Phys Eng Sci.* 361(1807):1251-62.
- Pawson T. (2002) Regulation and targets of receptor tyrosine kinases. *Eur J Cancer.* 38 Suppl 5:S3-10. Review.
- Pawson T, Raina M and Nash P. (2002) Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* 513, 2-10
- Pawson T, Scott JD. (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science.* 278(5346):2075-80. Review.
- Payre F, Vincent A. (1988) Finger proteins and DNA-specific recognition: distinct patterns of conserved amino acids suggest different evolutionary modes. *FEBS Lett.* 234(2):245-50. Review.
- Pelzer C, Thome M. (2011) IKK α takes control of canonical NF- κ B activation. *Nat Immunol.* 12(9):815-6.
- Penengo L, Mapelli M, Murachelli AG, Confalonieri S, Magri L, Musacchio A, Di Fiore PP, Polo S, Schneider TR. (2006) Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin. *Cell.* 124(6):1183-95.

REFERENCES

- Perozzo R, Folkers G, Scapozza L. (2004) Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *J Recept Signal Transduct Res.* 24(1-2):1-52. Review.
- Petoukhov MV, Svergun DI. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J.* 89(2):1237-50.
- Petrelli A, Gilestro GF, Lanzardo S, Comoglio PM, Migone N, Giordano S. (2002) The endophilin-CIN85-Cbl complex mediates ligand-dependent downregulation of c-Met. *Nature.* 416(6877):187-90.
- Pickart CM, Eddins MJ. (2004) Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta.* 1695(1-3):55-72. Review.
- Pickart CM, Fushman D. (2004) Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol.* 8(6):610-6. Review.
- Plaxco KW, Guijarro JI, Morton CJ, Pitkeathly M, Campbell ID, Dobson CM. (1998) The folding kinetics and thermodynamics of the Fyn-SH3 domain. *Biochemistry.* 37(8):2529-37.
- Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR, Li W, Tan AK, Bonnard C, Ong RT, Thalamuthu A, Pettersson S, Liu C, Tian C, Chen WV, Carulli JP, Beckman EM, Altshuler D, Alfredsson L, Criswell LA, Amos CI, Seldin MF, Kastner DL, Klareskog L, Gregersen PK. (2007) TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. *N Engl J Med.* 357(12):1199-209.
- Polekhina G, House CM, Traficante N, Mackay JP, Relaix F, Sassoon DA, Parker MW, Bowtell DD. (2002) Siah ubiquitin ligase is structurally related to TRAF and modulates TNF-alpha signaling. *Nat Struct Biol.* 9(1):68-75.
- Poornima CS, Dean PM. (1995) Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. *J Comput Aided Mol Des.* 9(6):500-12.
- Poornima CS, Dean PM. (1995) Hydration in drug design. 2. Influence of local site surface shape on water binding. *J Comput Aided Mol Des.* 9(6):513-20.
- Poornima CS, Dean PM. (1995) Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J Comput Aided Mol Des.* (6):521-31.
- Porod G. (1951) X-ray small angle scattering of close packed colloidal systems. *Kolloid Zeitschrift.* 124, 83-114.
- Poulsen FM (2002) A brief introduction to NMR spectroscopy of proteins. <http://www.cs.duke.edu/brd/Teaching/Bio/asmb/current/2papers/Introreviews/flemming.pdf>.
- Predki PF, Sarkar B. (1992) Effect of replacement of "zinc finger" zinc on estrogen receptor DNA interactions. *J Biol Chem.* 267(9):5842-6.
- Prompers JJ, Groenewegen A, Hilbers CW, Pepermans HA. (1999) Backbone dynamics of *Fusarium solani* pisi cutinase probed by nuclear magnetic resonance: the lack of interfacial activation revisited. *Biochemistry.* 38(17):5315-27.
- Prudêncio M, Ubbink M. (2004) Transient complexes of redox proteins: structural and dynamic details from NMR studies. *J Mol Recognit.* 17(6):524-39. Review.
- Putnam CD, Hammel M, Hura GL, Tainer JA. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys.* 40(3):191-285. Review.
- Rahighi S, Ikeda F, Kawasaki M, Akutsu M, Suzuki N, Kato R, Kensche T, Uejima T, Bloor S, Komander D, Randow F, Wakatsuki S, Dikic I. (2009) Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell.* 136(6):1098-109.
- Raiborg C, Slagsvold T, Stenmark H. A new side to ubiquitin. (2006) *Trends Biochem Sci.* 31(10):541-4. Review.
- Raiborg C. (2006) Hrs makes receptors silent: a key to endosomal protein sorting. *Crit Rev Oncol.* 12(3-4):295, discussion 295-6.

- Ramelot TA, Raman S, Kuzin AP, Xiao R, Ma LC, Acton TB, Hunt JF, Montelione GT, Baker D, Kennedy MA. (2009) Improving NMR protein structure quality by Rosetta refinement: a molecular replacement study. *Proteins*. 75(1):147-67.
- Razmiafshari M, Kao J, d'Avignon A, Zawia NH. (2001) NMR identification of heavy metal-binding sites in a synthetic zinc finger peptide: toxicological implications for the interactions of xenobiotic metals with zinc finger proteins. *Toxicol Appl Pharmacol*. 172(1):1-10.
- Remerowski ML, Domke T, Groenewegen A, Pepermans HA, Hilbers CW, van de Ven FJ. (1994) ¹H, ¹³C and ¹⁵N NMR backbone assignments and secondary structure of the 269-residue protease subtilisin 309 from *Bacillus lentus*. *J Biomol NMR*. 4(2):257-78.
- Ren R, Mayer BJ, Cicchetti P, Baltimore D. (1993) Identification of a ten-amino acid proline-rich SH3 binding site. *Science*. 259(5098):1157-61.
- Renzoni DA, Pugh DJ, Siligardi G, Das P, Morton CJ, Rossi C, Waterfield MD, Campbell ID, Ladbury JE. (1996) Structural and thermodynamic characterization of the interaction of the SH3 domain from Fyn with the proline-rich binding site on the p85 subunit of PI3-kinase. *Biochemistry*. 35(49):15646-53.
- Reyes-Turcu FE, Horton JR, Mullally JE, Heroux A, Cheng X, Wilkinson KD. (2006) The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell*. 124(6):1197-208.
- Richter PH, Eigen M. (1974) Diffusion controlled reaction rates in spheroidal geometry. Application to repressor--operator association and membrane bound enzymes. *Biophys Chem*. 2(3):255-63.
- Rickert M, Boulanger MJ, Goriatcheva N, Garcia KC. (2004) Compensatory energetic mechanisms mediating the assembly of signaling complexes between interleukin-2 and its alpha, beta, and gamma(c) receptors. *J Mol Biol*. 339(5):1115-28.
- Roldan JL, Blackledge M, van Nuland, NA & Azuaga AI (2011) Solution structure, dynamics and thermodynamics of the three SH3 domains of CD2AP. *J Biomol NMR*. 50, 103-117
- Rozakis-Adcock M, Fernley R, Wade J, Pawson T, Bowtell D. (1993) The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Nature*. 363(6424):83-5.
- Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Hariharan IK, Fortini ME, Li PW, Apweiler R, Fleischmann W, Cherry JM, Henikoff S, Skupski MP, Misra S, Ashburner M, Birney E, Boguski MS, Brody T, Brokstein P, Celniker SE, Chervitz SA, Coates D, Cravchik A, Gabrielian A, Galle RF, Gelbart WM, George RA, Goldstein LS, Gong F, Guan P, Harris NL, Hay BA, Hoskins RA, Li J, Li Z, Hynes RO, Jones SJ, Kuehl PM, Lemaitre B, Littleton JT, Morrison DK, Mungall C, O'Farrell PH, Pickeral OK, Shue C, Vossball LB, Zhang J, Zhao Q, Zheng XH, Lewis S. (2000) Comparative genomics of the eukaryotes. *Science*. 287(5461):2204-15.
- Sahu SC, Bhuyan AK, Udgaonkar JB, Hosur RV. (2000) Backbone dynamics of free barnase and its complex with barstar determined by ¹⁵N NMR relaxation study. *J Biomol NMR*. 18(2):107-18.
- Salcini AE, Confalonieri S, Doria M, Santolini E, Tassi E, Minenkova O, Cesareni G, Pelicci PG, Di Fiore PP. (1997) Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. *Genes Dev*. 11(17):2239-49.
- Sato Y, Yoshikawa A, Yamashita M, Yamagata A, Fukai S. (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF domains of TAB2 and TAB3. *EMBO J*. 28(24):3903-9.
- Sattler M, Schleucher J, Griesinger C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc*. 34, 93-158
- Schafer DA, D'Souza-Schorey C, Cooper JA. (2000) Actin assembly at membranes controlled by ARF6. *Traffic*. 1(11):892-903.
- Schiffner CA, Huber R, Wüthrich K, van Gunsteren WF. (1994) Simultaneous refinement of the structure of BPTI against NMR data measured in solution and X-ray diffraction data measured in single crystals. *J Mol Biol*. 241(4):588-99.

REFERENCES

- Schlessinger J. (2000) Cell signaling by receptor tyrosine kinases. *Cell*. 103(2):211-25. Review.
- Schmidt MH, Dikic I. (2005) The Cbl interactome and its functions. *Nat Rev Mol Cell Biol*. 6(12):907-18. Review.
- Schmidt MH, Chen B, Randazzo LM, Bogler O. (2003) SETA/CIN85/Ruk and its binding partner AIP1 associate with diverse cytoskeletal elements, including FAKs, and modulate cell adhesion. *J Cell Sci*. 116(Pt 14):2845-55.
- Schmitz N, Sureda A, Dini G, Claviez A. (2009) Lymphoma and Pediatric Diseases Working Parties. Allogeneic transplantation for children and adolescents with Hodgkin lymphoma. *Blood*. 114(20):4605-6.
- Schroeder B, Weller SG, Chen J, Billadeau D, McNiven MA. (2010) A Dyn2-CIN85 complex mediates degradative traffic of the EGFR by regulation of late endosomal budding. *EMBO J*. 29(18):3039-53.
- Schwieters CD, Kuszewski JJ, Clore GM. (2006) Using Xplor-NIH for NMR molecular structure determination. *Progr. NMR Spectroscopy* 48, 47-62.
- Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson*. 160(1):65-73.
- Schwieters CD, Clore GM. (2002) Reweighted atomic densities to represent ensembles of NMR structures. *J Biomol NMR*. 23(3):221-5.
- Sekiyama N, Jee J, Isogai S, Akagi K, Huang TH, Ariyoshi M, Tochio H, Shirakawa M. (2012) NMR analysis of Lys63-linked polyubiquitin recognition by the tandem ubiquitin-interacting motifs of Rap80. *J Biomol NMR*. 52(4):339-50.
- Senftleben U, Cao Y, Xiao G, Greten FR, Krähn G, Bonizzi G, Chen Y, Hu Y, Fong A, Sun SC, Karin M. (2001) Activation by IKK α of a second, evolutionary conserved, NF- κ B signaling pathway. *Science*. Aug 24;293(5534):1495-9.
- Shaanan B, Gronenborn AM, Cohen GH, Gilliland GL, Veerapandian B, Davies DR, Clore GM. (1992) Combining experimental information from crystal and solution studies: joint X-ray and NMR refinement. *Science*. 257(5072):961-4.
- Sharp R. (2008) The mechanism of paramagnetic NMR relaxation produced by Mn(II): role of orthorhombic and fourth-order zero field splitting terms. *J Chem Phys*. 129(14):144307.
- Sheinerman FB, Norel R, Honig B. (2000) Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol*. Apr;10(2):153-9. Review.
- Shembade N, Pujari R, Harhaj NS, Abbott DW, Harhaj EW. (2011) The kinase IKK α inhibits activation of the transcription factor NF- κ B by phosphorylating the regulatory molecule TAX1BP1. *Nat Immunol*. 12(9):834-43.
- Shembade N, Ma A, Harhaj EW. (2010) Inhibition of NF- κ B signaling by A20 through disruption of ubiquitin enzyme complexes. *Science*. Feb 26;327(5969):1135-9.
- Shembade N, Parvatiyar K, Harhaj NS, Harhaj EW. (2009) The ubiquitin-editing enzyme A20 requires RNF11 to downregulate NF- κ B signalling. *EMBO J*.28(5):513-22.
- Shembade N, Harhaj NS, Parvatiyar K, Copeland NG, Jenkins NA, Matesic LE, Harhaj EW. (2008) The E3 ligase Itch negatively regulates inflammatory signaling pathways by controlling the function of the ubiquitin-editing enzyme A20. *Nat Immunol*. 9(3):254-62.
- Shembade N, Harhaj NS, Liebl DJ, Harhaj EW. (2007) Essential role for TAX1BP1 in the termination of TNF- α -, IL-1- and LPS-mediated NF- κ B and JNK signaling. *EMBO J*. 26(17):3910-22.
- Shen Y, Delaglio F, Cornilescu G, Bax A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 44(4):213-23.
- Shih NY, Li J, Karpitskii V, Nguyen A, Dustin ML, Kanagawa O, Miner JH, Shaw AS. (1999) Congenital nephrotic syndrome in mice lacking CD2-associated protein. *Science*. 286(5438):312-5.
- Shilatifard A. (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem*. 75:243-69. Review.

- Shin D, Lee SY, Han S, Ren S, Kim S, Aikawa Y, Lee S. (2012) Differential polyubiquitin recognition by tandem ubiquitin binding domains of Rabex-5. *Biochem Biophys Res Commun.* 423(4):757-62.
- Shoji T, Higuchi M, Kondo R, Takahashi M, Oie M, Tanaka Y, Aoyagi Y, Fujii M. (2009) Identification of a novel motif responsible for the distinctive transforming activity of human T-cell leukemia virus (HTLV) type 1 Tax1 protein from HTLV-2 Tax2. *Retrovirology.* 6:83.
- Simpson RJ, Yi Lee SH, Bartle N, Sum EY, Visvader JE, Matthews JM, Mackay JP, Crossley M. (2004) A classic zinc finger from friend of GATA mediates an interaction with the coiled-coil of transforming acidic coiled-coil 3. *J Biol Chem.* 279(38):39789-97.
- Sims JJ, Haririnia A, Dickinson BC, Fushman D, Cohen RE. (2009) Avid interactions underlie the Lys63-linked polyubiquitin binding specificities observed for UBA domains. *Nat Struct Mol Biol.* 16(8):883-9.
- Singer KH, Haynes BF. (1987) Epithelial-thymocyte interactions in human thymus. *Hum Immunol.* 20(2):127-44.
- Skaug B, Jiang X, Chen ZJ. (2009) The role of ubiquitin in NF-kappaB regulatory pathways. *Annu Rev Biochem.* 78:769-96. Review.
- Skorski T, Nieborowska-Skorska M, Wlodarski P, Wasik M, Trotta R, Kanakaraj P, Salomoni P, Antonyak M, Martinez R, Majewski M, Wong A, Perussia B, Calabretta B. (1998) The SH3 domain contributes to BCR/ABL-dependent leukemogenesis in vivo: role in adhesion, invasion, and homing. *Blood.* 91(2):406-18
- Slagsvold T, Pattni K, Malerød L, Stenmark H. (2006) Endosomal and non-endosomal functions of ESCRT proteins. *Trends Cell Biol.* 16(6):317-26.
- Snyder DA, Chen Y, Denissova NG, Acton T, Aramini JM, Ciano M, Karlin R, Liu J, Manor P, Rajan PA, Rossi P, Swapna GV, Xiao R, Rost B, Hunt J, Montelione GT. (2005) Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination. *J Am Chem Soc.* 127(47):16505-11.
- Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, Xia B, Livingston DM, Greenberg RA. (2007) RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science.* 316(5828):1198-202.
- Song AX, Zhou CJ, Peng Y, Gao XC, Zhou ZR, Fu QS, Hong J, Lin DH, Hu HY. (2010) Structural transformation of the tandem ubiquitin-interacting motifs in ataxin-3 and their cooperative interactions with ubiquitin chains. *PLoS One.* 5(10):e13202.
- Songyang Z, Fanning AS, Fu C, Xu J, Marfatia SM, Chishti AH, Crompton A, Chan AC, Anderson JM, Cantley LC. (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science.* 275(5296):73-7.
- Soubeyran P, Kowanzet K, Szymkiewicz I, Langdon WY, Dikic I. (2002) Cbl-CIN85-endophilin complex mediates ligand-induced downregulation of EGF receptors. *Nature.* 416(6877):183-7.
- Spinola M. (2009) Phd thesis. Estudios estructurales de BRMS1 y Tax1BP1, proteínas involucradas en la vía de activación de NFkB. Universidad Autónoma de Madrid. Madrid. Spain.
- Stamenova SD, French ME, He Y, Francis SA, Kramer ZB, Hicke L. (2007) Ubiquitin binds to and regulates a subset of SH3 domains. *Mol Cell.* 25(2):273-84.
- Stapley BJ, Creamer TP. (1999) A survey of left-handed polyproline II helices. *Protein Sci.* 8(3):587-95.
- Staub O, Rotin D. (2006) Role of ubiquitylation in cellular membrane transport. *Physiol Rev.* 86(2):669-707. Review.
- Stauffer TP, Martenson CH, Rider JE, Kay BK, Meyer T. (1997) Inhibition of Lyn function in mast cell activation by SH3 domain binding peptides. *Biochemistry.* 36(31):9388-94.
- Stein A, Céol A, Aloy P. (2011) 3did: identification and classification of domain-based interactions of known three-dimensional structure. *Nucleic Acids Res.* 39(Database issue):D718-23.

REFERENCES

- Stein R. (1998) SH2 and SH3 domains. Unraveling signaling networks with peptide antagonists. *Methods Mol Biol.* 88:187-95.
- Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzclaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksöz E, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE. (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell.* 122(6):957-68.
- Sturtevant JM. (1977) Heat capacity and entropy changes in processes involving proteins. *Proc Natl Acad Sci U S A.* 74(6):2236-40.
- Sudol M. (1996) The WW module competes with the SH3 domain? *Trends Biochem Sci.* 21(5):161-3.
- Sudol M, Chen HI, Bougeret C, Einbond A, Bork P. (1995) Characterization of a novel protein-binding module--the WW domain. *FEBS Lett.* 369(1):67-71. Review.
- Suh JY, Tang C, Clore GM. (2007) Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J Am Chem Soc.* 129(43):12954-5.
- Suhy DA, Simon KD, Linzer DI, O'Halloran TV. (1999) Metallothionein is part of a zinc-scavenging mechanism for cell survival under conditions of extreme zinc deprivation. *J Biol Chem.* 274(14):9183-92.
- Sun L, Chen ZJ. (2004) The novel functions of ubiquitination in signaling. *Curr Opin Cell Biol.* 2004 16(2):119-26. Review.
- Sundquist WI, Schubert HL, Kelly BN, Hill GC, Holton JM, Hill CP. (2004) Ubiquitin recognition by the human TSG101 protein. *Mol Cell.* 13(6):783-9.
- Suzuki N, Wakatsuki S, Kawasaki M. (2010) Crystal structure of GFP-Wrnip1 UBZ domain fusion protein in complex with ubiquitin. PDB code 3VHT.
- Svergun DI. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 76, 2879-86
- Svergun DI, Koch MH, Pedersen JS, Serdyuk IN. (1996) Structural model of the 50S subunit of *E. coli* ribosomes from solution scattering. *Basic Life Sci.* 64:149-74.
- Svergun DI, Barberato C, Koch MHJ. (1995) CRYSOLO - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* 28, 768-773
- Swanson KA, Kang RS, Stamenova SD, Hicke L, Radhakrishnan I. (2003) Solution structure of Vps27 UIM-ubiquitin complex important for endosomal sorting and receptor downregulation. *EMBO J.* 22(18):4597-606.
- Szymkiewicz I, Kowanetz K, Soubeyran P, Dinarina A, Lipkowitz S, Dikic I. (2002) CIN85 participates in Cbl-b-mediated down-regulation of receptor tyrosine kinases. *J Biol Chem.* 277(42):39666-72.
- Tang C, Ghirlando R, Clore GM. (2008) Visualization of transient ultra-weak protein self-association in solution using paramagnetic relaxation enhancement. *J Am Chem Soc.* 130(12):4048-56.
- Tang C, Iwahara J, Clore GM. (2006) Visualization of transient encounter complexes in protein-protein association. *Nature.* 444(7117):383-6.
- Teo H, Gill DJ, Sun J, Perisic O, Veprintsev DB, Vallis Y, Emr SD, Williams RL. (2006) ESCRT-I core and ESCRT-II GLUE domain structures reveal role for GLUE in linking to ESCRT-I and membranes. *Cell.* 125(1):99-111.
- Teo H, Veprintsev DB, Williams RL. (2004) Structural insights into endosomal sorting complex required for transport (ESCRT-I) recognition of ubiquitinated proteins. *J Biol Chem.* 279(27):28689-96.
- Thomson KJ, Peggs KS, Blundell E, Goldstone AH, Linch DC. (2007) A second autologous transplant may be efficacious in selected patients with Hodgkin's lymphoma relapsing after a previous autograft. *Leuk Lymphoma.* 48(5):881-4.
- Tibaldi EV, Reinherz EL. (2003) CD2BP3, CIN85 and the structurally related adaptor protein CMS bind to the same CD2 cytoplasmic segment, but elicit divergent functional activities. *Int Immunol.* 15(3):313-29.

- Tjandra N, Szabo A, Bax A. (1996) Protein backbone dynamics and ^{15}N chemical shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.* 118, 6986-6991.
- Tjandra N, Feller SE, Pastor RW, Bax A. (1995) Rotational diffusion anisotropy of human ubiquitin from ^{15}N relaxation. *J. Am. Chem. Soc.* 117, 12562-12566.
- Tokunaga F, Nakagawa T, Nakahara M, Saeki Y, Taniguchi M, Sakata S, Tanaka K, Nakano H, Iwai K. (2011) SHARPIN is a component of the NF- κ B-activating linear ubiquitin chain assembly complex. *Nature.* 471(7340):633-6.
- Tokunaga F, Sakata S, Saeki Y, Satomi Y, Kirisako T, Kamei K, Nakagawa T, Kato M, Murata S, Yamaoka S, Yamamoto M, Akira S, Takao T, Tanaka K, Iwai K. (2009) Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat Cell Biol.* 11(2):123-32.
- Tong AH, Drees B, Nardelli G, Bader GD, Brannetti B, Castagnoli L, Evangelista M, Ferracuti S, Nelson B, Paoluzi S, Quondam M, Zucconi A, Hogue CW, Fields S, Boone C, Cesareni G. (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science.* 295(5553):321-4.
- Tossidou I, Teng B, Drobot L, Meyer-Schwesinger C, Worthmann K, Haller H, Schiffer M. (2010) CIN85/RukL is a novel binding partner of nephrin and podocin and mediates slit diaphragm turnover in podocytes. *J Biol Chem.* 285(33):25285-95.
- Trempe JF, Chen CX, Grenier K, Camacho EM, Kozlov G, McPherson PS, Gehring K, Fon EA. (2009) SH3 domains from a subset of BAR proteins define a Ubl-binding domain and implicate parkin in synaptic ubiquitination. *Mol Cell.* 36(6):1034-47.
- Turnbull WB, Daranas AH. (2003) On the value of c : can low affinity systems be studied by isothermal titration calorimetry? *J Am Chem Soc.* 125(48):14859-66.
- Ubbink M, Ejdebäck M, Karlsson BG, Bendall DS. (1998) The structure of the complex of plastocyanin and cytochrome f , determined by paramagnetic NMR and restrained rigid-body molecular dynamics. *Structure.* 6(3):323-35.
- Ulrich M, Seeber S, Becker CM, Enz R. (2007) Tax1-binding protein 1 is expressed in the retina and interacts with the GABA(C) receptor rho1 subunit. *Biochem J.* 401(2):429-36.
- Usami Y, Popov S, Göttinger HG. (2007) Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. *J Virol.* 81(12):6614-22.
- Uversky VN, Longhi S (eds) (2010) Instrumental Analysis of Intrinsically Disordered Proteins. John Wiley & Sons, Hoboken, NJ.
- Vallabhapurapu S, Karin M. (2009) Regulation and function of NF-kappaB transcription factors in the immune system. *Annu Rev Immunol.* 27:693-733.
- van Duijn TJ, Anthony EC, Hensbergen PJ, Deelder AM, Hordijk PL. (2010) Rac1 recruits the adapter protein CMS/CD2AP to cell-cell contacts. *J Biol Chem.* 285(26):20137-46.
- Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C, Fushman D. (2004) Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *J Biol Chem.* 279(8):7055-63.
- Vega MC, Martínez JC, Serrano L. (2000) Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot. *Protein Sci.* 9(12):2322-8.
- Velazquez-Campoy A, Freire E. (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat Protoc* 1(1): 186-191.
- Velázquez Campoy A, Freire E. (2005) ITC in the post-genomic era...? Priceless. *Biophys Chem.* 115(2-3):115-24.
- Velazquez-Campoy A, Leavitt SA, Freire E. (2004) Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods Mol Biol.* 261:35-54.
- Velazquez-Campoy A, Kiso Y, Freire E. (2001) The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. *Arch Biochem Biophys.* 390(2):169-75.

REFERENCES

- Velazquez-Campoy A, Todd MJ, Freire E. (2000) HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry*. 39(9):2201-7.
- Verstrepen L, Verhelst K, Carpentier I, Beyaert R. (2011) TAX1BP1, a ubiquitin-binding adaptor protein in innate immunity and beyond. *Trends Biochem Sci*. 36(7):347-54. Review.
- Verstrepen L, Verhelst K, van Loo G, Carpentier I, Ley SC, Beyaert R. (2010) Expression, biological activities and mechanisms of action of A20 (TNFAIP3). *Biochem Pharmacol*. 80(12):2009-20. Review.
- Vidal M, Gigoux V, Garbay C. (2001) SH2 and SH3 domains as targets for anti-proliferative agents. *Crit Rev Oncol Hematol*. 40(2):175-86. Review.
- Viguera AR, Martínez JC, Filimonov VV, Mateo PL, Serrano L. (1994) Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry*. 33(8):2142-50.
- Vistica J, Dam J, Balbo A, Yikilmaz E, Mariuzza RA, Rouault TA, Schuck P. (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem*. 326(2):234-56.
- Volkov AN, Ubbink M, van Nuland NA. (2010) Mapping the encounter state of a transient protein complex by PRE NMR spectroscopy. *J Biomol NMR*. 48(4):225-36.
- Volkov AN. (2007) Phd thesis: Transient Complexes of Haem proteins. Leiden University. The Netherlands.
- Volkov AN, Worrall JA, Holtzmann E, Ubbink M. (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc Natl Acad Sci U S A*. 103(50):18945-50.
- Volkov AN, Ferrari D, Worrall JA, Bonvin AM, Ubbink M. (2005) The orientations of cytochrome c in the highly dynamic complex with cytochrome b5 visualized by NMR and docking using HADDOCK. *Protein Sci*. 14(3):799-811.
- Volkov VV, Svergun DI. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Cryst*. 36, 860-864.
- Vonrhein C, Blanc E, Roversi P, Bricogne G. (2007) Automated structure solution with autoSHARP. *Methods Mol Biol*. 364:215-30.
- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687-696
- Vriend G. (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 8(1):52-6, 29.
- Wakasaki T, Masuda M, Niuro H, Jabbarzadeh-Tabrizi S, Noda K, Taniyama T, Komune S, Akashi K. (2010) A critical role of c-Cbl-interacting protein of 85 kDa in the development and progression of head and neck squamous cell carcinomas through the ras-ERK pathway. *Neoplasia*. 12(10):789-96.
- Walczak H. (2011) TNF and ubiquitin at the crossroads of gene activation, cell death, inflammation, and cancer. *Immunol Rev*. 244(1):9-28. Review.
- Wang B, Elledge SJ. (2007) Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brcal/Brc36 complex in response to DNA damage. *Proc Natl Acad Sci U S A*. 104(52):20759-63.
- Wang C, Pawley NH, Nicholson LK. (2001) The role of backbone motions in ligand binding to the c-Src SH3 domain. *J Mol Biol*. 313(4):873-87.
- Wang Q, Young P, Walters KJ. (2005) Structure of S5a bound to monoubiquitin provides a model for polyubiquitin recognition. *J Mol Biol*. 348(3):727-39.
- Wang X, Naidu SR, Sverdrup F, Androphy EJ. (2009) Tax1BP1 interacts with papillomavirus E2 and regulates E2-dependent transcription and stability. *J Virol*. 83(5):2274-84
- Watanabe O, Kinoshita J, Shimizu T, Imamura H, Hirano A, Okabe T, Aiba M, Ogawa K. (2005) Expression of a CD44 variant and VEGF-C and the implications for lymphatic metastasis and long-term prognosis of human breast cancer. *J Exp Clin Cancer Res*. 24(1):75-82.

- Watanabe S, Take H, Takeda K, Yu ZX, Iwata N, Kajigaya S. (2000) Characterization of the CIN85 adaptor protein and identification of components involved in CIN85 complexes. *Biochem Biophys Res Commun.* 278(1):167-74.
- Weber C, Wider G, von Freyberg B, Traber R, Braun W, Widmer H, Wüthrich K. (1991) The NMR structure of cyclosporin A bound to cyclophilin in aqueous solution. *Biochemistry.* 30(26):6563-74.
- Weber G, Anderson SJ. (1965) Multiplicity of binding. Range and practical test of Adair's equation. *Biochemistry.* 10:1942-1947.
- Weiss M, Elsner M, Kartberg F, Nilsson T. (2004) Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophys J.* 87(5):3518-24.
- Welsch T, Endlich N, Kriz W, Endlich K. (2001). CD2AP and p130Cas localize to different F-actin structures in podocytes. *Am. J. Physiol.* 281, F769-F777.
- Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S, Wu P, Wiesmann C, Baker R, Boone DL, Ma A, Koonin EV, Dixit VM. (2004) De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature.* 430(7000):694-9.
- Whisstock JC, Lesk AM. (1999) SH3 domains in prokaryotes. *Trends Biochem Sci.* 24(4):132-3.
- Wienands J, Schweikert J, Wollscheid B, Jumaa H, Nielsen PJ, Reth M. (1998) SLP-65: a new signaling component in B lymphocytes which requires expression of the antigen receptor for phosphorylation. *J Exp Med.* 188(4):791-5.
- Wiseman T, Williston S, Brandts JF, Lin LN. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem.* 179(1):131-7.
- Wittekind M, Mapelli C, Farmer BT 2nd, Suen KL, Goldfarb V, Tsao J, Lavoie T, Barbacid M, Meyers CA, Mueller L. (1994) Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry.* 33(46):13531-9.
- Wolfe SA, Nekludova L, Pabo CO. (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct.* 29:183-212. Review.
- Woodruff RV, Bomar MG, D'Souza S, Zhou P, Walker GC. (2010) The unusual UBZ domain of *Saccharomyces cerevisiae* polymerase η . *DNA Repair (Amst).* 9(11):1130-41.
- Worrall JA, Liu Y, Crowley PB, Nocek JM, Hoffman BM, Ubbink M. (2002) Myoglobin and cytochrome b5: a nuclear magnetic resonance study of a highly dynamic protein complex. *Biochemistry.* 41(39):11721-30.
- Xiao G, Cvijic ME, Fong A, Harhaj EW, Uhlik MT, Waterfield M, Sun SC. (2001) Retroviral oncoprotein Tax induces processing of NF-kappaB2/p100 in T cells: evidence for the involvement of IKKalpha. *EMBO J.* 20(23):6805-15.
- Xu XX, Yang W, Jackowski S, Rock CO. (1995) Cloning of a novel phosphoprotein regulated by colony-stimulating factor 1 shares a domain with the *Drosophila* disabled gene product. *J Biol Chem.* 270(23):14184-91.
- Yaffe MB. (2002) Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol.* 3(3):177-86. Review.
- Yamazaki T, Forman-Kay JD, Kay LE (1993) Two-dimensional NMR experiments for correlating ^{13}C and ^1H chemical shifts of aromatic residues in ^{13}C -labeled proteins via scalar couplings. *J Am Chem Soc.* 115:11054-11055.
- Yan J, Kim YS, Yang XP, Li LP, Liao G, Xia F, Jetten AM. (2007) The ubiquitin-interacting motif containing protein RAP80 interacts with BRCA1 and functions in DNA damage repair response. *Cancer Res.* 67(14):6647-56.
- Ye J, Mayer KL, Stone MJ. (1999) Backbone dynamics of the human CC-chemokine eotaxin. *J Biomol NMR.* 15(2):115-24.
- Yee AA, Savchenko A, Ignachenko A, Lukin J, Xu X, Skarina T, Evdokimova E, Liu CS, Semesi A, Guido V, Edwards AM, Arrowsmith CH. (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *J Am Chem Soc.* 127(47):16512-7.

REFERENCES

- Yoon HY, Kales SC, Luo R, Lipkowitz S, Randazzo PA. (2011) ARAP1 association with CIN85 affects epidermal growth factor receptor endocytic trafficking. *Biol Cell*. 103(4):171-84.
- Yoshida M, Korfhagen TR, Whitsett JA. (2001) Surfactant protein D regulates NF-kappa B and matrix metalloproteinase production in alveolar macrophages via oxidant-sensitive pathways. *J Immunol*. 166(12):7514-9.
- Zarrinpar A, Bhattacharyya RP, Lim WA. (2003) The structure and function of proline recognition domains. *Sci STKE*. 2003(179):RE8. Review.
- Zeng W, Sun L, Jiang X, Chen X, Hou F, Adhikari A, Xu M, Chen ZJ. (2010) Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell*. 141(2):315-30.
- Zhang J, Yang W, Piquemal JP, Ren P. (2012) Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J Chem Theory Comput*. 8(4):1314-1324.
- Zhou HX, Rivas G, Minton AP. (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys*. 37:375-97.
- Zuiderweg ER. (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry*. 41(1):1-7. Review.
- Zwier MC, Chong LT. (2010) Reaching biological timescales with all-atom molecular dynamics simulations. *Curr Opin Pharmacol*. Review.

