Structural and functional characterization of the intrinsically disordered Unique domain of c-Src

Mariano Maffei

Structural and functional characterization of the intrinsically disordered Unique domain of c-Src

Memòria presentada per

Mariano Maffei

Per obtenir el grau de doctor per la Universitat de Barcelona

Director de Tesi:

Prof. Miquel Pons i Vallès

Departament de Química Orgànica, Universitat de Barcelona

Tutor de Tesi:

Prof. Ernest Giralt Lledo

Departament de Química Orgànica, Universitat de Barcelona

Programa de doctorat:

Biotecnologia

Models, Metodologies i aplicacions

Facultat de Química

Universitat de Barcelona

Universitat de Barcelona

IRB BARCELONA
INSTITUTE FOR RESEARCH IN BIOMEDICINE

A mio Padre e mia Madre.

The cover picture has been realized by Laura Caroletti

and is entitled “Structure and Function”.

ACKNOWLEDGEMENTS

Formal acknowledgements

The research described in this thesis has been made possible with the financial support provided by the Spanish Ministry of Science and Innovation - MICIN (project no. BIO2010-15683, fellowship no. BES-2011-047340), the Government of Catalunya and the Bio-NMR Consortium.

I thank the support of the NMR "Instalación Científico Tecnológica Singular" of the Universitat de Barcelona, the "Mass Spectrometry facility" of the IRB Barcelona and all the IRB-Barcelona Science Park for the excellent scientific environment.

I am very grateful to Angel R. Nebreda and Ana Igea for the fruitful collaboration and for the knowledge transfer you gave me during the first years of my PhD.

I would like to thank Dr. Serge Roche for hosting me in the "Tyrosine kinase signalling and oncogenesis group" at the CRBM in Montpellier, for his kindness and for his availability to discuss about science.

I am also thankful to Dr. Andrew Marsh and Prof. Dr. Michael Overduin for hosting me at University of Warwick and Birmingham.

A special thanks goes to the secretariat of the Department of Organic Chemistry at the Universitat de Barcelona for the assistance in dealing with bureaucracy and for the constant help and availability during the years.

Informal acknowledgements

My first and sincere acknowledgment goes to my thesis director Miquel. I thank you for the extraordinary efforts you made for training me as a scientist. You love science and you are able to transmit your immense passion and knowledge to the others. Apart of this, you constantly demonstrated your enormous sensitivity and humanity for people: this makes you not only a good teacher but also a special person ("You'll never walk alone"). My work and my efforts are also dedicated to you.

"Queridos Ponsis", thank you for your infinite support you gave me during my years of PhD. Without your existence it wouldn't be the same. Ildefonso ("el Sevillano") and Irene, thanks for teaching me the "art crafts", for sharing with me adventures and for helping me in time of need. You are really good friends. To Tiago, Orilo and Carles, "the old guard": my friendly and "loving" thanks go to you. Tiago, you have been my first "mentor" and one of my best friends. I hope I have taken care of "Can Pons" appropriately. I also thank you for your help and availability during my stay in Montpellier (*L'enfant prodige*). Orilo, you have been a mate, a friend and a colleague. We have shared wonderful moments together and I will always bring with me our friendship (*Tu italiano preferido*). Carles (Capoccione), you are a sincere and straight arrow friend. I won't forget the time spent together (*Tu capitán*). Dear Isabel, thanks for introducing me to the Pons' world (and to the Catalan!). A special thanks goes also to Miguel for the last year spent together in the lab (and outside): thanks for your support and for the effort made for helping me. Take care of the project my good friend. Anabel-Lise, good luck for your last year of thesis, there is always the light at the end of the tunnel. Don't give up! Borja, I wish all the bests for your future brilliant career. I also acknowledge Yandi,

Pau, Jesús, Xavi R., Lucas, Joao and all the colleagues that have passed through the NMR lab. Once Ponsis, forever Ponsis!

A felt acknowledge goes also to the French colleagues in Montpellier for the utmost help given me in the last steps of my thesis: Audrey, Romain, Nicolas, Valerie, Priscilla and the rest of the group.

Personally, I would like to thank all the IRB friends: Giorgia, Giulio, Milos, Ivan, Pablo, Silvia, Pep, Lorena(s), Benjamí, Elisa, Daria, Giuseppe, Felipe, Michela, Federica, Radek, Marta, Adriana, Janire, Helena and all the rest. Thanks for sharing with me this wonderful period of my life. I am also grateful to my "Aribau" flatmates: Fabio, almost four years of cohabitation, Giulia, Bérénice and Anne-Sophie, Javi, Natasha, Francesca and all the rest that have stopped by Aribhouse! Thanks to all the people that I have met in Barcelona during my life journey, everyone has left something inside me.
Thanks to Barcelona, my second mother. Mentioning the Latin poet Catullo "*Odi et amo. Quare id faciam, fortasse requiris.*"

And now the most important acknowledgements.

To my mother, a strong woman that had allowed me to reach the highest peak of the mountain. My achievements are always dedicated to you. My strength is your strength. To my father, my guiding light despite your agonizing absence ("*Bit by bit I have realized that he was with me*"). To the rest of my family (Nonna Vilma, Nonna Titina, Nonno Benedetto, Nonno Natalino, Zia Cinzia e Zio Massimo, Zio Lello e Zia Grazia, Zia Angela e Zio Piero, Arianna, Alessio, Mario, Ginny etc.) THANK YOU for being always there, in every moment of my life and for giving me all of your love.

To my best friends: Ardo, Lupo, Marco, Marcozzi and Marmotta (you have been sorted in alphabetical order, not by preference). We have grown up together; we have shared thousands of adventures; we were always present in times of joy and in times of pain; nothing and nobody has scraped our friendship during 15 years. My heartfelt acknowledgment goes to you, for being the brothers that I have never had, for being those who were always present without complaining and for being those who I will always count on. “I’ll be there for you”.

To my wonderful native city: Roma. You are a mother, a wife and a lover. I will always be grateful to the emotions that I feel when I am home. (“*Se, infine, troverai che Itaca é povera non pensare che ti abbia ingannato. Perché sei divenuto saggio, hai vissuto una vita intensa, e questo é il significato di Itaca”*).

And last but not least, the most important acknowledge goes to the wonderful woman of my life: Laura. I was lost and you gave me back the joie of vivre, I was disappointed and you gave me back the enthusiasm, I had lost the desire and you taught me what it means to love again (“*Certi amori non finiscono, fanno dei giri immensi e poi ritornano*”). Thanks for offering me your love in every moment of life, without it I wouldn’t be able to reach my objectives. All this work is truthfully dedicated to you; you are my muse, you are my sunshine, I love you with all my might. *...Affaccete alla finestra occhioni belli...*

Always be yourself, always pursue your dreams, and never change.

Yours,

Mariano

INDEX

Part I – Introduction and Objectives

Part II – Results and Discussion

Part IV - Resum en català

Part V - Addenda

Part I

Introduction and Objectives

I.1 – Introduction part I

SFKs: The Src family of non-receptor kinases

From the simplest to the most complex eukaryotic organism, protein kinases, the enzymes that catalyze the transfer of the terminal phosphoryl group of ATP (Adenosine triphosphate) onto specific protein substrates, usually onto a serine, threonine or tyrosine residue, have been shown for more than 50 years to have important functions in regulating various cellular aspects[1]. There are over 500 protein kinases in the human/mouse genomes[2,3] that can be divided into various subfamilies bases on their structure and function (Figure I.1.2.A).

Among all the kinase families, one of the most studied during the past nearly four decades has been the Src family of non-receptor kinases (SFKs). The family consists of nine members (Figure I.1.2.B): Lyn, Hck, Lck, Blk, Src, Fyn, Yes, Yrk and Fgr that show similar modes of regulation and act as important signaling intermediaries regulating a variety of outputs, such as cell proliferation, differentiation, apoptosis, migration, and metabolism[4,5,6]. In addition, these kinases can be grouped into two subfamilies, those that are Lyn related (Lyn, Hck, Lck and Blk) and those that are Src related (Src, Yes, Fyn, Yrk and Fgr).

All the SFKs share a conserved domain arrangement[7] (Figure I.1.1) constituted by: a N-terminal region followed by an intrinsically disordered Unique domain which is divergent among family members; the SH3 and SH2 domains; the SH1 or kinase domain and a C-terminal tail. A short membrane-anchoring SH4 domain with myristoylation and sometimes palmitoylation sites forms the N-terminal region. The SH3 domain can mediate interactions with proline-rich motifs related to the *PXXP* consensus[8]. The SH2 domain can bind to phosphotyrosine motifs, with SFKs showing highest affinity for the consensus sequence *pYEEI*[8]. The kinase domain (SH1) is responsible

for the catalytic activity but also harbours an activation loop (A-loop) with a positive-regulatory tyrosine (Y418) that when phosphorylated allows maximum kinase activity. The C-terminal tail contains a negative-regulatory tyrosine (Y527) that, when phosphorylated, by an intramolecular interaction with the SH2 domain[9,10]. The intramolecular interaction between the SH3 domain with a polyproline region[11] in the connector between the SH2 and SH1 domains[9] further contributes to lock the protein in an enzymatically inactive closed form. Also grouped in the larger SFKs family there are the two kinases Brk and Frk, however, none of these are myristoylated and even they conserved the potential regulatory C-terminal tyrosines, no significant sequence homology with the corresponding Src/Lyn subfamilies is present. Active and inactive configurations of the SFKs have been biochemically and structurally detailed[12], which has delineated differential phosphorylation, domain interaction and competitive domain displacement as the common modes of their regulation.

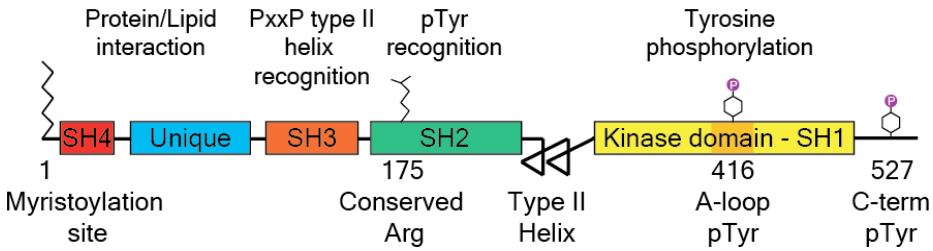


Figure I.1.1. The domain structure of Src family kinases. The Src kinase architecture consists of four domains: the Unique region, which varies among family members, followed by the SH3, SH2, and tyrosine kinase domains. The activation loop (A-loop) of the kinase domain, the activating (Tyr 416) and autoinhibitory (Tyr 527) phosphorylation sites, the conserved residue Arg 175 in the SH2 domain and the polyproline type II helix are indicated. By convention, amino-acid residues are numbered as in chicken Src.

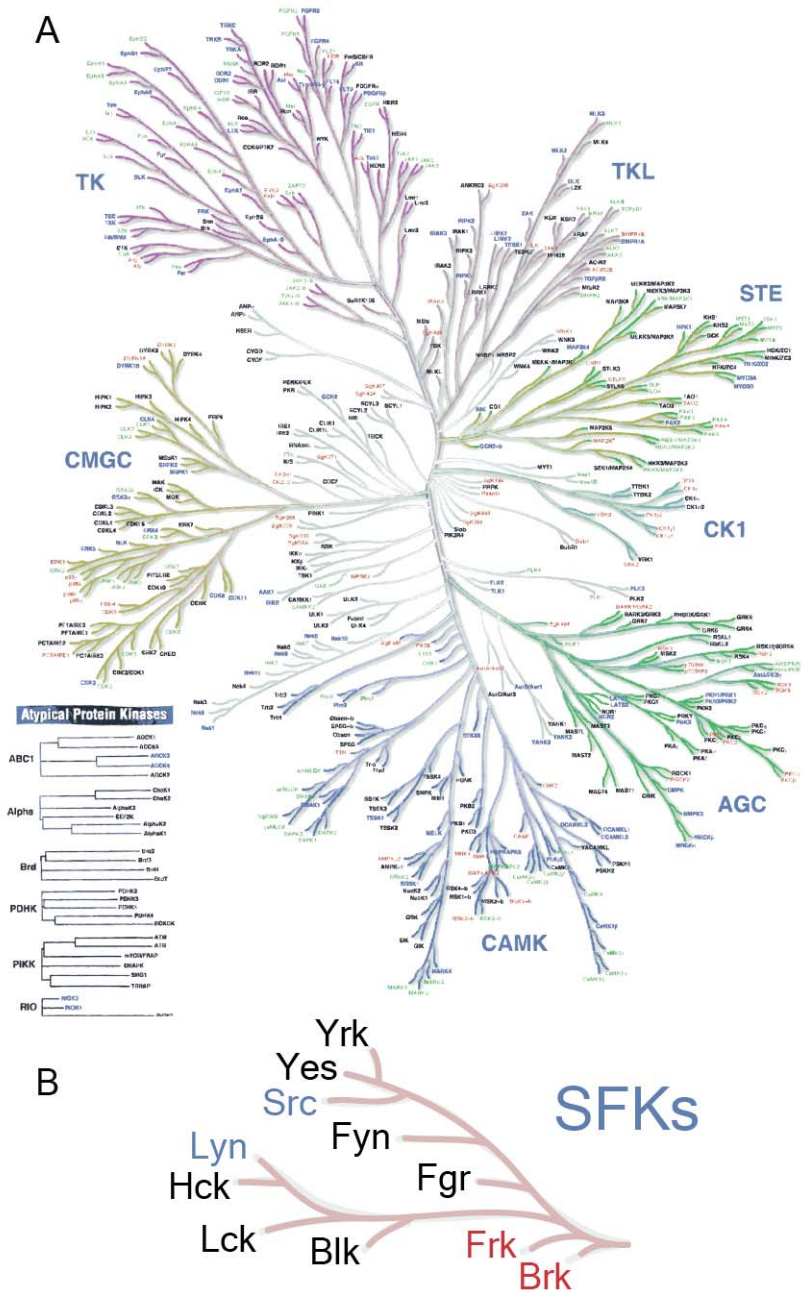


Figure I.1.2. (A) The "human kinome" *(modified from Chartier et al., PeerJ. 2013)* **(B) Schematic representation of the SFKs.**

c-Src tyrosine kinase

c-Src is the leading member of the Src family of non-receptor kinases and has been the subject of intense investigation for decades (Figure I.1.3). These studies spring from the work on the Rous sarcoma virus, a chicken tumor virus discovered in 1911 by Peyton Rous[13]. The viral src (*v-src*) gene was the first retroviral oncogene to be identified, and its cellular counterpart was the first proto-oncogene to be discovered in the vertebrate genome[13].

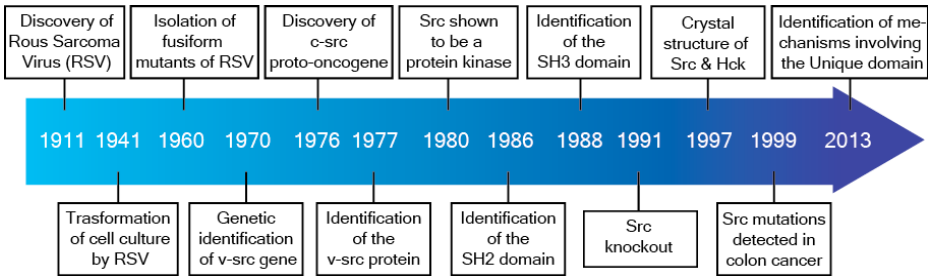


Figure I.1.3. Key events in the discovery of Src.

Src is composed by 533 aminocids (chicken Src; 536 for human c-Src) and is expressed ubiquitously; however, brain, osteoclasts, and platelets express 5 – 200-fold higher levels of this protein than most other cells[1]. Alternatively spliced forms of Src, which contain 6 or 11 amino acid insertions in the SH3 domain (18 or 33 nucleotide insertions between exons 3 and 4), are expressed in nerve cells. Nevertheless, there is little evidence for a functional difference among Src splice variants[14]. c-Src can be found inside cells in an activated or in an not-activated form. Inactive c-Src is mainly localized at perinuclear sites, but c-Src activation causes its SH3 domain to become indirectly associated with actin. Activated c-Src is ultimately translocated to the cell periphery to sites of cell adhesion, where it attaches to the plasma membrane inner surface through its myristoylated SH4 domain[15]. This location allows for interactions with membrane-bound receptor tyrosine kinases and integrins associated

with adhesion functions. Src kinase activity is very low during most of the cell cycle, but is transiently activated by cell-cycle-dependent C-terminal tail dephosphorylation and activation loop phosphorylation[16]. The association of c-Src with the plasma membrane is also considered essential for cellular transformation[17], and the autophosphorylation of Tyr 419 (active form) that occurs with membrane targeting, which is enabled by interactions with activated receptor tyrosine kinases, is associated with the highest level of c-Src transforming activity[18] . Src shares the classical domain architecture of the SFKs consisting in a N-terminal Unique region; the SH3, SH2 and SH1 folded domains followed by a C-terminal regulatory tail (see above, Figure I.1.2). The first X-ray structure of c-Src was determined in 1997[12,19] and is shown in Figure I.1.4.

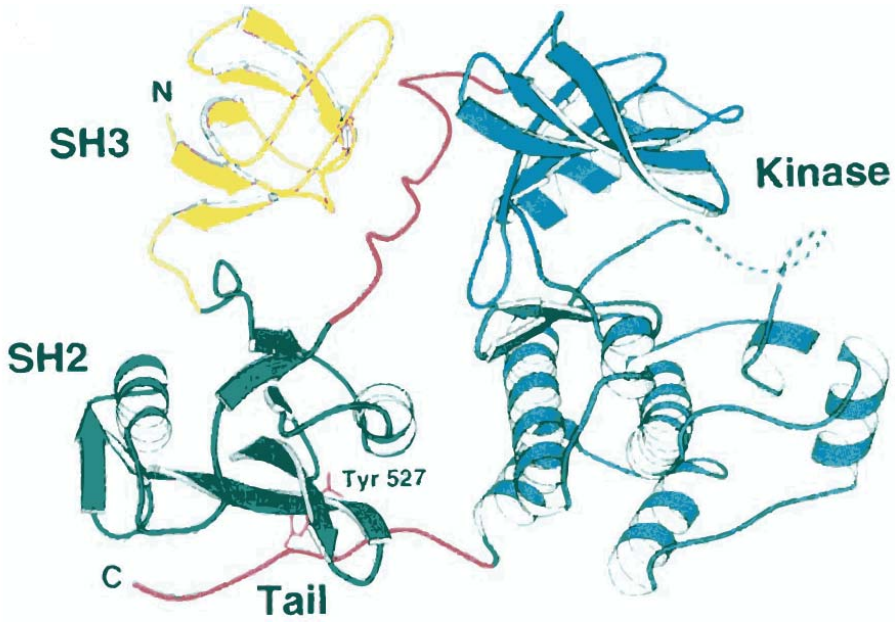


Figure I.1.4. The first determined crystal structure of c-Src. *(modified from Xu et al., Nature 1997)*[12] The inactive conformation of c-Src is shown. The SH3 and SH2 domain are shown in yellow and green, respectively. The N- and C- lobe are colored in cyan (top) and blue (bottom), respectively. pTyr 527 is bound to the SH2 domain.

Regulation mechanisms involving the human tyrosine kinase c-Src have been extensively studied. Mutations in the SH3 and SH2 domains usually activate Src kinase; the Tyr 527 Phe (Y527F) mutation also promotes Src enzyme activity[1]. The SH2 domain binds to phosphotyrosine 527 in the C-terminal tail and this interaction stabilizes the attachment of the SH2 domain to the large lobe (C-lobe) of the kinase domain. The linker between the SH2 and kinase domains contains prolines at positions 246, 250, and 263 that function as a motif that binds the SH3 domain and attaches the SH3 domain to the small lobe (N-lobe) of the SH1 domain. The linker does not resemble the classical PXXP motif[20], but this stretch of residues readily forms a left-handed polyproline type II helix. The activation loop (A-loop) can switch between active and inactive conformations. In the inactive state, Tyr 416, located in the activation loop, is sequestered and is not a substrate for phosphorylation by another kinase. The SH2 domain binds to pTyr 527 and the SH3 domain interacts with the polyprolyne type II helix stabilizing the enzyme. When pTyr 527 is dephosphorylated it dissociates or is displaced from the SH2 binding pocket trigging an “unclamping” process in the active form. Thus, Tyr 416 can then undergo autophosphorylation by another Src kinase molecule and the enzyme is stabilized in its active state. Exogenous substrates decrease autophosphorylation *in vitro* during activity measurements[21]. This finding is consistent with the fact that activation loop phosphorylation occurs in trans, and this process is inhibited by competition with exogenous substrates. Inactivation of c-Src by phosphorylation of the terminal tyrosine residue is known to be performed by Csk[22] and its homologue Chk. Csk is structurally related to c-Src, but lacks the negative-regulatory domain of the c-Src C-terminus, and there is evidence that reduced expression of Csk might have a role in c-Src activation in human cancer[23]. Conversely, the C-terminal phosphate of c-Src can be removed by several protein

phosphatases that function as activators of c-Src. Protein tyrosine phosphatase-α (PTPα) has been shown to dephosphorylate the terminal tyrosine residue[24] *in vitro* and *in vivo*, and PTP1 (Protein tyrosine phosphatase-1), SH2-containing phosphatase 1 (SHP1) and SH2-containing phosphatase 2 (SHP2) might also regulate c-Src[25].

Activation of c-Src can also occur by interactions with its partners through the SH2 or SH3 domain, as the intramolecular interactions that maintain the closed configuration are displaced[26]. Until very recently[27], the role of the intrinsically disordered Unique domain was poorly understood. Recent results have demonstrated that the Unique domain could play an active role in c-Src function suggesting a new layer of regulation.

As a cytoplasmic protein, c-Src has a critical role in mediating signal transduction via interactions with multiple proteins and protein complexes. Depending on its cellular localization, c-Src phosphorylates different substrates in the cytosol or at the inner face of the plasma membrane, at cell–matrix or cell–cell adhesions. Tyrosine phosphorylation may affect the functions of these proteins directly. Alternatively, the phosphotyrosyl residues serve as docking sites for the binding of signaling proteins containing SH2 domains. These signaling complexes initiate pathways that regulate protein synthesis, gene expression, cytoskeletal assembly and many other aspects of cell function such as motility or survival[1,14]. Some of the substrates regulated by c-Src are shown in Figure I.1.5. These include the Ras–mitogen-activated-protein-kinase (Ras–MAPK) pathway, which regulates *fos* expression through phosphorylation of members of the Ets family of transcription factors; the phosphatidylinositol-3- kinase–Akt (PI3K–Akt) pathway, which regulates translational initiation and cell survival; and the Stat3 pathway, which regulates the expression of myc and cyclin D1[13,28].

Of all of the SFKs, c-Src is the one that is most often implicated in cancer. In 1999, Tim Yeatman's group reported that a subset of advanced human colon cancers contained a mutation in c-Src that results in a carboxy-terminal truncation of the Src protein and activation of its catalytic and transforming activities[29]. Beyond colorectal cancer, increased c-Src activity has been demonstrated in several other gastrointestinal malignancies, including hepatocellular, pancreatic, gastric and esophageal malignancies[30], as well as in breast[31], ovarian[32] and lung cancers[32]. Hepatocellular cancers[33,34] and colon carcinomas[35] are of interest because they can both overexpress c-Src and underexpress the negative-regulatory c-Src tyrosine kinase (CSK) protein concurrently, leading to higher levels of c-Src activation. Increased specific activity of c-Src can also occur in the presence of relatively normal levels of c-Src protein expression[28]. In many cancers increased c-Src kinase activity is associated with advanced-stage tumors that readily metastasize and is thought to have an important role in the increased metastatic potential of these tumors. c-Src is proposed to affect the processes of tumor growth and cancer-cell adhesion, motility and invasion, which directly influence metastatic potential. Overexpression of wild type c-Src does not seem to be sufficient to enhance metastatic potential[36], whereas more active forms of Src such as c-Src-531, a truncated form that lacks C-terminal inhibitory amino acids, and v-Src[29,30] might be capable of permitting cells to achieve distant organ colonization[37]. However, in contrast to viral oncoproteins, activating mutations of cellular c-Src are rarely observed in human cancers and overexpression of Src by itself does not induce cell transformation. Therefore, the oncogenic potential of c-Src is probably related to failures in the network of interactions that regulate their activity in normal cells.

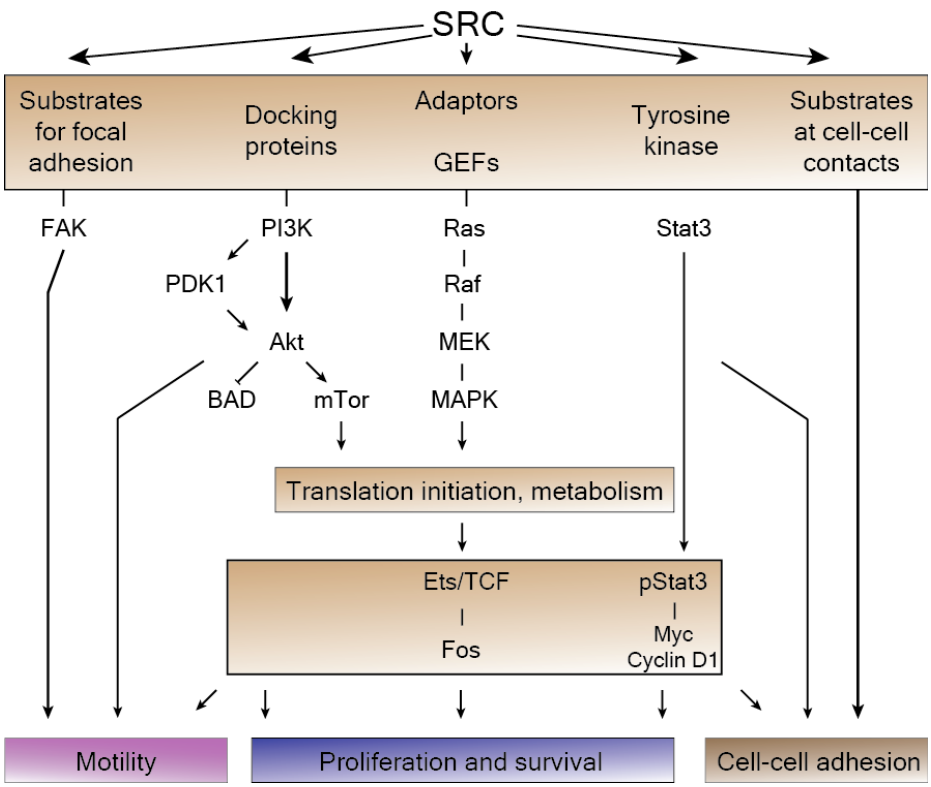


Figure I.1.5. Signaling by c-Src. The Ras-mitogen activated protein kinase (Ras-MAPK) pathway; the phosphatidylinositol-3-kinase-Akt (PI3K-Akt) pathway; and the Stat3 pathway, are shown. (GEFs, guanine-nucleotide-exchange factors; mTOR, mammalian target of rapamycin; PDKs, 3′-phosphoinositide-dependent kinases.).

The Unique domain of c-Src: an intrinsically disordered domain

Although c-Src was the first discovered oncogene, the role of its intrinsically disordered N-terminal region, called the Unique domain (USrc), has remained obscure for a long time. The reported X-ray structures of c-Src[12,19] do not contain the N-terminal region that hinders crystallization and is easily degraded by proteases[11]. This phenomenon is presumably related to the fact that USrc is intrinsically unfolded[38]. The sequence homology within the Src family of non-receptor kinases is rather high with the exception of this N-terminal

region where the only conserved amino acids are Gly 2 and Cys 3 that are sites for myristoylation and palmitoylation, respectively (Figure I.1.6.A). However, the Unique domain of each individual SFK member is well conserved between different organisms suggesting a more specific role than that of a simple spacer[39]. The SH4 and the Unique domains of c-Src have been related to subcellular localization and trafficking[40]. A cluster of 6 basic residues in the first 15 residues of c-Src, as well as the 38-111 stretch, has been implicated in membrane anchoring[41]. The N-terminal region of c-Src was also shown to be implicated in protein-protein interactions[42,43,44]. Recent studies have started to shed light on possible functional roles of the Unique domain of human c-Src. In 2009 Perez et al.[38] characterized the spectroscopic properties and the conformational space sampled by USrc using state-of-the-art NMR techniques. A partially structured region comprising residues 60-75 was identified within the intrinsically disordered N-terminal fragment (Figure I.1.6.B). Moreover, in the same work phosphorylation events involving residues Ser 17, Thr 37 and Ser 75 were identified *in vitro*. Further insights about USrc were provided later always from Pons' group[27]. It was shown, for the first time, that human c-Src contains two additional lipid binding regions in addition to the previously described SH4 domain: the Unique Lipid Binding Region (ULBR) and the SH3 domain. Interestingly, the ULBR was shown to be located within the partially structured region previously determined by NMR. The USrc capability to interact with lipids was demonstrated *in vitro* to be modulated by phosphorylation events involving residues previously identified (Ser 17, Thr 37 and Ser 75). The Unique domain of human c-Src was also shown to be involved in inter- and intra-molecular interaction. Indeed, USrc interacts with Calmodulin in presence of calcium and with its neighbor SH3 domain. This intra-molecular interaction is allosterically inhibited by the binding of a high-affinity polyproline peptide to the SH3 domain of c-Src.

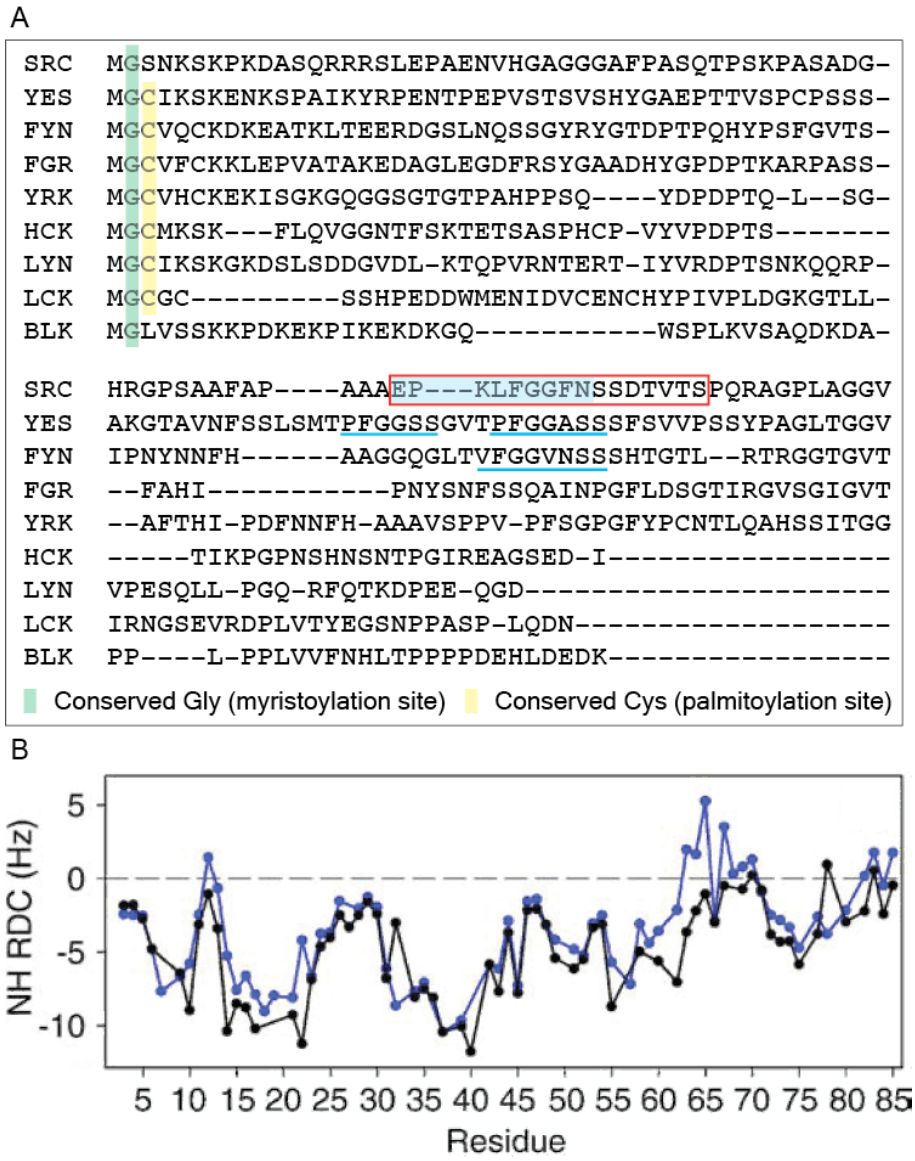


Figure I.1.6. (A) Sequences alignment of the Unique domains of SFKs. Residues highlighted in green are identified as conserved sites of myristoylation (Gly 2) while conserved cysteines (palmitoylation sites) are shown in yellow. The ULBR is represented in blue. Blue lines highlight the ULBR sequence conserved in Yes and Fyn kinases. Red box indicates the identified partially structured region. **(B) N-H^N RDCs measured for the Unique domain of the c-Src.** *(modified from Perez et al., JMB 2009)*[38] Comparison of the RDCs measured at pH 4.5 (blue) and those measured in 4 M urea (black).

REFERENCES

1. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287,** 121-149 (1996).

2. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298,** 1912-34 (2002).

3. Caenepeel, S., Charydczak, G., Sudarsanam, S., Hunter, T. & Manning, G. The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 11707-12 (2004).

4. Frame, M. C. Src in cancer: deregulation and consequences for cell behaviour. *Biochim. Biophys. Acta* **1602,** 114-30 (2002).

5. Parsons, S. J. & Parsons, J. T. Src family kinases, key regulators of signal transduction. *Oncogene* **23,** 7906-9 (2004).

6. Chong, Y.-P., Ia, K. K., Mulhern, T. D. & Cheng, H.-C. Endogenous and synthetic inhibitors of the Src-family protein tyrosine kinases. *Biochim. Biophys. Acta* **1754,** 210-20 (2005).

7. Boggon, T. J. & Eck, M. J. Structure and regulation of Src family kinases. *Oncogene* **23,** 7918-27 (2004).

8. Koch, C., Anderson, D., Moran, D., Ellis, V. & Pawson, T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins, *Science* **252**, 668-674 (1991).

9. Xu, W., Doshi, a, Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3,** 629-38 (1999).

10. Cowan-Jacob, S. W. *et al.* The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13,** 861-71 (2005).

11. Bernadó, P., Pérez, Y., Svergun, D. I. & Pons, M. Structural characterization of the active and inactive states of Src kinase in solution by small-angle X-ray scattering. *J. Mol. Biol.* **376,** 492-505 (2008).

12. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385,** 595-602 (1997).

13. Martin, G. S. The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2,** 467–75 (2001).

14. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13,** 513–609 (1997).

15. Sefton, B. M., Trowbridge, I. S., Cooper, J. A. & Scolnick, E. M. The transforming proteins of Rous sarcoma virus, Harvey sarcoma virus and Abelson virus contain tightly bound lipid. *Cell* **31**, 465–474 (1982).

16. Roskoski, R. Src protein-tyrosine kinase structure and regulation. *Biochem. Biophys. Res. Commun.* **324,** 1155–64 (2004).

17. Nigg, E. a, Sefton, B. M., Hunter, T., Walter, G. & Singer, S. J. Immunofluorescent localization of the transforming protein of Rous sarcoma virus with antibodies against a synthetic src peptide. *Proc. Natl. Acad. Sci. U. S. A.* **79,** 5322–6 (1982).

18. Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85,** 149–58 (1996).

19. Sicheri, F. & Kuriyan, J. Structures of Src-family tyrosine kinases. *Curr. Opin. Struct. Biol.* **7,** 777–85 (1997).

20. Cohen, G. B., Ren, R. & Baltimore, D. Modular binding domains in signal transduction proteins. *Cell* **80,** 237–48 (1995).

21. Sun, G. *et al.* Effect of autophosphorylation on the catalytic and regulatory properties of protein tyrosine kinase Src. *Arch. Biochem. Biophys.* **397,** 11–7 (2002).

22 Cooper, J. A., Gould, K. L., Cartwright, C. A. & Hunter, T. Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science* **231**, 1431–1434 (1986).

23 Masaki, T. et al. Reduced C-terminal Src kinase (Csk) activities in hepatocellular carcinoma. *Hepatology* **29**, 379–384 (1999). 32.

24 Zheng, X. M., Wang, Y. & Pallen, C. J. Cell transformation and activation of pp60c-src by overexpression of a protein tyrosine phosphatase. *Nature* **359**, 336–339 (1992).

25 Jung, E. J. & Kim, C. W. Interaction between chicken protein tyrosine phosphatase 1 (CPTP1)-like rat protein phosphatase 1 (PTP1) and p60(v-src) in v-src- transformed Rat-1 fibroblasts. *Exp. Mol. Med.* **34**, 476–480 (2002). 40.

26. Thomas, J. W. *et al.* SH2- and SH3-mediated Interactions between Focal Adhesion Kinase and Src. *J. Biol. Chem.* **273,** 577–583 (1998).

27. Pérez, Y. *et al.* Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* **3,** 1295 (2013).

28. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4,** 470–80 (2004).

29. Irby, R. B. *et al.* Activating SRC mutation in a subset of advanced human colon cancers. *Nat. Genet.* **21,** 187–90 (1999).

30. Irby, R. B. & Yeatman, T. J. Role of Src expression and activation in human cancer. *Oncogene* **19**, 5636–5642 (2000).

31. Hynes, N. E. Tyrosine kinase signalling in breast cancer. *Breast Cancer Res*. **2**, 154–157 (2000).

32. Wiener, J. R. et al. Activated SRC protein tyrosine kinase is overexpressed in late-stage human ovarian cancers. *Gynecol. Oncol.* **88**, 73–79 (2003).

33. Masaki, T. et al. pp60c-src activation in hepatocellular carcinoma of humans and LEC rats. *Hepatology* **27**, 1257–1264 (1998).

34. Masaki, T. et al. Reduced C-terminal Src kinase (Csk) activities in hepatocellular carcinoma. *Hepatology* **29**, 379–384 (1999).

35. Cam, W. R. et al. Reduced C-terminal Src kinase activity is correlated inversely with pp60(c-src) activity in colorectal carcinoma. *Cancer* **92**, 61–70 (2001).

36. Irby, R. et al. Overexpression of normal c-Src in poorly metastatic human colon cancer cells enhances primary tumor growth but not metastatic potential. *Cell Growth Differ.* **8**, 1287–1295 (1997).

37. Boyer, B., Bourgeois, Y. & Poupon, M. F. Src kinase contributes to the metastatic spread of carcinoma cells. *Oncogene* **21**, 2347–2356 (2002).

38. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391,** 136–48 (2009).

39. Amata, I., Maffei, M. & Pons, M. Phosphorylation of unique domains of Src family kinases. *Front. Genet.* **5,** 1–6 (2014).

40. Kasahara, K. *et al.* Rapid trafficking of c-Src, a non-palmitoylated Src-family kinase, between the plasma membrane and late endosomes/lysosomes. *Exp. Cell Res.* **313,** 2651-66 (2007).

41. Kaplan, J. M., Varmus, H. E. & Bishop, J. M. The src protein contains multiple domains for specific attachment to membranes. *Mol. Cell. Biol.* **10,** 1000-1009 (1990).

42. Resh, M. D. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta* **1451,** 1-16 (1999).

43. Lee, H. *et al.* Palmitoylation of caveolin-1 at a single site (Cys-156) controls its coupling to the c-Src tyrosine kinase: targeting of dually acylated molecules (GPI-linked, transmembrane, or cytoplasmic) to caveolae effectively uncouples c-Src and caveolin-1 (TYR-14). *J. Biol. Chem.* **276,** 35150-8 (2001).

44. Gingrich, J. R. *et al.* Unique domain anchoring of Src to synaptic NMDA receptors via the mitochondrial protein NADH dehydrogenase subunit 2. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6237-6242 (2004).

I.2 – Introduction part II

NMR in structural biology: Paramagnetic Relaxation Enhancement (PREs)

Since its discovery[1,2], Nuclear Magnetic Resonance (NMR) experiments have been extensively applied in the biomedical field, chemistry and structural biology. NMR spectroscopy is used for investigating the structure, dynamics, reaction state, and chemical environment of molecules by exploiting the interaction of certain atomic nuclei with applied magnetic fields and the capacity to probe this interaction by absorption and emission of electromagnetic radiation. In contrast to X-ray crystallography, Nuclear Magnetic Resonance is a suitable tool to study intrinsically disordered proteins (IDPs). Conventional NMR observables like chemical shifts or scalar couplings are intrinsically of short-range nature. However, long-range information can be obtained from paramagnetic induced effects, such as Paramagnetic Relaxation Enhancement (PRE). PRE is a high sensitive technique that allows detecting weak and transient intra- or inter-molecular interactions. The technique is based on the use of paramagnetic tags with unpaired electrons, which enhance the relaxation of the surrounding nuclei due to dipolar magnetic interactions. Because of the large magnetic moment of the unpaired electron (657 times that of a proton), the PRE effects can be observed at distances up to 35 Å, depending on the paramagnetic tag used. Thus, the longer range of the PRE allows exploring a wider region of the conformational space sampled by the protein with a single experiment. The common approach consists on binding covalently a stable radical (paramagnetic probe) to a cysteine residue and measuring the ^{1}H-^{15}N-HSQC (Heteronuclear Single Quantum Coherence) spectra on the paramagnetic species. Enhanced nuclear

relaxation of nuclei that are close to the paramagnetic centers results in broadening of the signals and hence a reduction on the intensity of the peaks. Subsequently, the spin label is reduced using a reducing agent such as ascorbic acid yielding a diamagnetic species and the spectra are acquired again (Figure I.2.1). The observed signals in the paramagnetic and diamagnetic species have identical chemical shifts and do not require re-assignment. Given the distance dependence and sensitivity of the PRE, even low populated conformers presenting transient interactions contribute to signal broadening and can be observed. Another advantage is that using site directed mutagenesis the tag position can be easily and precisely selected in absence of native cysteine residues. Using PRE-data the structures have been characterized for folded and unfolded proteins, protein-nucleic acids[3,4], and membrane proteins[5].

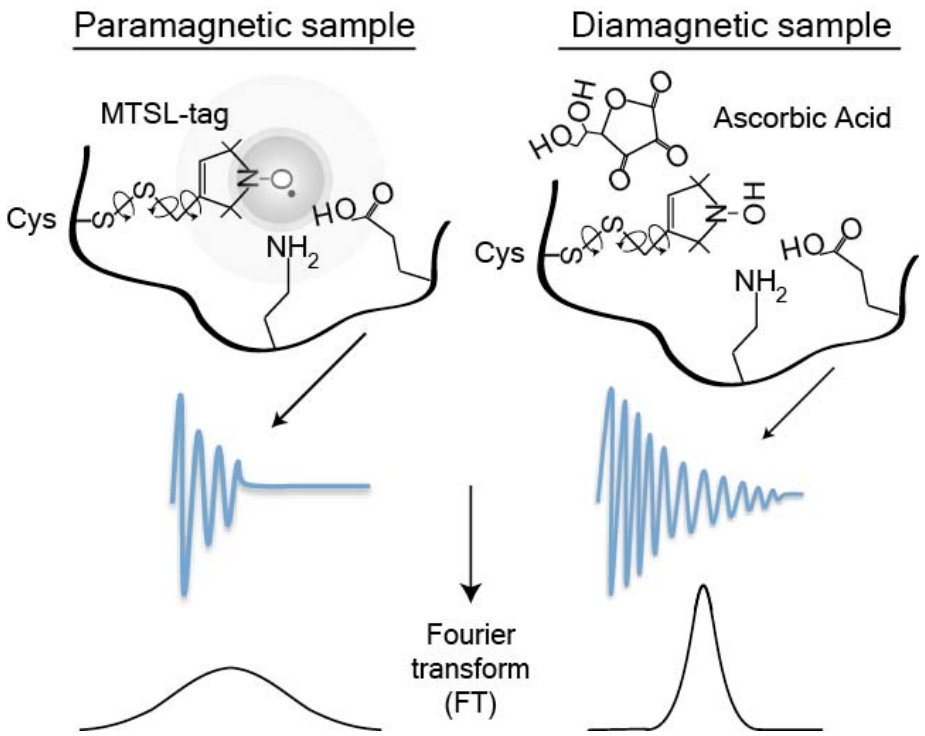


Figure I.2.1. Schematic representation of the PRE technique. The paramagnetic tag (e.g. MTSL) is attached through disulfide bond to a Cysteine present on the protein. The presence of the unpaired electron on the tag enhances the nuclear relaxation of the surrounding nuclei and hence the broadening of the signals (reduction of intensity). Addition of Ascorbic Acid causes probe reduction and the formation of the diamagnetic specie (right panels). The reduced form of the MTSL and the oxidation product of the Ascorbic Acid are represented in the right panel.

In-cell NMR

Remarkable advances have been done in the last 10 years in high-resolution image techniques[6]. Super-resolution light microscopy as well as cryo-electron tomography have allowed to provide high detailed information about the complex internal environment of the cell[7,8]. However, both methods cannot furnish details about protein structure under *in-vivo* conditions. In this direction, *in-cell* NMR spectroscopy has opened the way to the investigation of protein conformation and activation under the intracellular crowded conditions of live cells, virtually at an atomic resolution[9].

NMR exhibits low invasiveness against living specimens, such as cells, tissues and whole organisms, thus it has become a powerful and pioneering tool of investigation in the structural biology field[10].

In-cell NMR spectroscopy represents an extension of a basic magnetic resonance (MR) principle: most atomic nuclei in natural substances, with the exception of protons (^{1}H), are NMR-insensitive and hence not detected by NMR methods. Therefore, these nuclei are substituted with stable isotopes in order to make them NMR 'visible'[11]. This approach allows focusing in the molecules of interest in spite of the bewildering complexity of the cellular milieu. Strictly speaking, while the spectroscopic observation is non-invasive, in-cell NMR spectroscopy does perturb the *in vivo* system as it introduces labelled biomolecules

at concentrations usually higher than the ones found in the unperturbed system and may create non-natural metabolic situations. The first report of an NMR observation of isotopically labelled biomolecules was reported by Guy Lippens[12] in 1998. The molecules were complex carbohydrates present in the periplasm of Gram negative bacteria. Since then, it has expanded to other prokaryotes and eukaryotic cells, including *Pichia Pastoris* yeast[13], *Xenopus Laevis* oocytes[14,15,16] and such mammalian cultured cells as HeLa, COS-7[17] and 293F cells[18] (Figure I.2.2).

Cell lines / transduction	Target molecules	Studies
HeLa (human) / COS-7 (monkey) / CPP (HIV-TAT)	Ub, FKBP12, GB1	Protein-drug interaction[17] enzymatic cleavage H/D exchange
293F (human) / SLO	Thymosin β4	N-terminal acetylation[18]
Xenopus Laevis oocytes (frog) / micro-injection	DNA / RNA hairpins Tau protein GB1	Structure and folding Protein phosphorylation Macromolecular crowding[14]
Pichia Pastoris (budding yeast) / over-expression	Ubiquitin	Influence of metabolic[13] change on protein
E. Coli (prokaryotic) / over-expression	TTHA1718	3D structure determination[20]

Ub = Ubiquitin; GB1 = protein G B1 domain; FKBP = FK506 binding protein;

Figure I.2.2. Table representing selected works related to in-cell NMR spectroscopy.

In the case of *E. coli*, the cells are firstly grown in isotopically unenriched medium, and at the time of IPTG induction, the culture medium is exchanged with an isotope-enriched medium[9] (Figure I.2.3.A). If the expression level of the induced proteins is sufficiently high, the NMR signals from the protein of interest can be distinguished from the other molecules and in-cell NMR spectra can be recorded. Using this system, protein-metal interaction or protein-protein interactions[19] have been studied in *E. coli* cells. Recently, 3D structure determination has been completed only on the basis of NMR data

recorded for proteins in living *E. coli* cells[20,21]. The first eukaryotic cellular system employed for in-cell NMR studies was oocytes of the African clawed frog, *Xenopus Laevis*[14,15]. In this case, isotope-enriched proteins, generally overexpressed in *E. coli*, are directly injected into the oocytes (Figure I.2.3.B), which is a procedure that is easily achieved due to their large size (approximately 1 mm). The system has been employed to study protein phosphorylation[16,22,23,24] and nucleic acid structures[25]. Mammalian cultured cells are widely used in cell biology, including medical and pharmaceutical science. Two protocols have been reported that enable in-cell NMR studies in these cells: one that utilizes a CPP (cell-penetrating peptide)[17], also called PTD (protein transduction domain), and another that uses SLO (streptolysin O)[18] (Figure I.2.3.C). The CPP sequence promotes cellular entry when fused to a cargo protein; however, the CPP-fused protein tends to be targeted to endosomes, preventing translocation to the cytosol. In the other case, the cells are first treated with SLO that induces the formation of pores of 35 nm in size in the plasma membrane. Through those pores the isotope-enriched proteins enter the cytosol. As this translocation process depends purely on the passive diffusion of the proteins, their concentration in the extracellular medium needs to be sufficiently high (approximately 1 mM) to achieve efficient delivery. In 2013, Aricescu and Banci groups[26,27] generated mammalian in-cell NMR samples by transient protein overexpression and isotopic labelling directly in human embryonic kidney HEK-293T cells driven by the cytomegalovirus (CMV) promoter. Despite the strong promoter, protein expression and labelling times were generally longer than in bacteria and led to substantial degrees of background labelling. Because these methods rely on plasmid- and strong promoter-driven protein overexpression coupled to induction time-matched isotopic labelling, they exclusively enable the production of protein in-cell NMR samples[28].

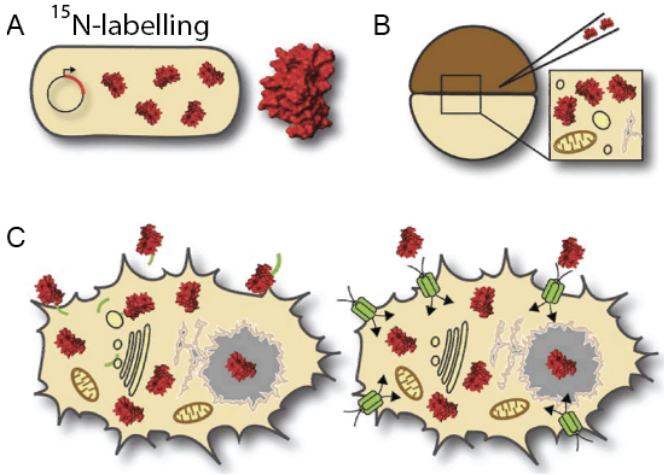


Figure I.2.3. Schematic overview of the protocols employed for in-cell NMR sample preparations. (A) Prokaryotic in-cell NMR. Recombinant protein production from strong plasmid promoters under isotope-labeled growth conditions. (B) Recombinant ^{15}N-labeled proteins are produced and purified in bacterial cells and directly microinjected into the cell cytoplasm of *Xenopus laevis* oocytes. (C) In-cell NMR applications in mammalian cultured cells necessitate intracellular protein transduction procedures with cell-penetrating peptides (left) or via cell-permeabilizing toxins (right).

In-cell NMR has been extensively used for the study of PTMs (post-translational modifications). Post-translationally modified proteins represent the majority of the proteome and establish, to a large part, the impressive level of functional diversity in higher, multi-cellular organisms[29]. In eukaryotes, PTMs constitute reversible, covalent additions of small chemical entities such as phosphate-, acyl-, alkyl- and glycosyl-groups onto selected subsets of modifiable amino acids. The addition of those groups induces highly specific changes in the chemical environments of individual protein residues, which can be readily detected by high-resolution NMR spectroscopy. Intrinsically disordered regions, including also protein loop regions[30,31] are often enriched in PTM sites. This, is probably related to the fact that fast, cellular signalling responses usually require modifying enzymes to rapidly access individual protein PTM sites, which is easier achieved

when modifiable amino acids are solvent exposed. Among all PTMs, modification of serine and threonine protein residues by reversible phosphorylation constitutes the most abundant PTM in eukaryotes[32]. Indeed, phosphorylatable serine/threonine residues are mostly solvent exposed and located in intrinsically disordered regions[30]. 2D ^{1}H-^{15}N correlation experiments are particularly suitable to identify phosphorylated serines and threonines, as these residues experience prominent backbone amide chemical shift perturbations[29,33] ($\Delta\delta$ = 0.5 - 1.5 p.p.m.) (Figure I.2.4). These changes are primarily caused by intra-residue hydrogen bonds between amide protons and the phosphate moieties[34,35]. Serine/threonine phosphorylation at multiple protein sites can lead to considerable increases in spectral complexity, especially when the individual modification sites are closely spaced. In fact, the large average number of protein serine/threonine residues that are phosphorylated by endogenous kinases in physiological environments such as cell extracts, often result in NMR spectra of that sort[29]. In those cases, other techniques such as mass spectrometry or immunoblotting are generally used in parallel with NMR to unequivocally identify which are the residues that are phosphorylated. As well as phosphorylation, dephosphorylation events have been also mapped through high-resolution NMR[16,36].

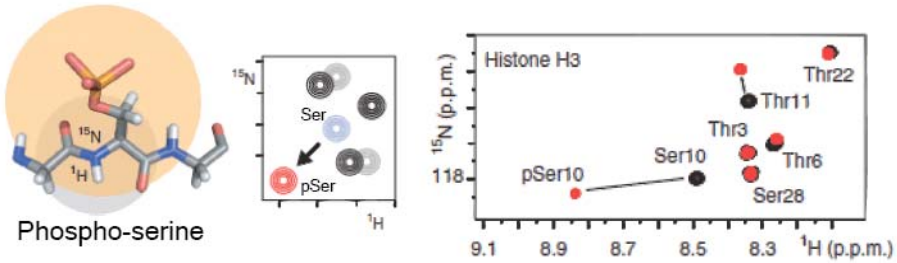


Figure I.2.4. Example of NMR properties of serine/threonine phosphorylation. *(modified from Theillet, F.-X. et al., Nat. Prot. 2013).* Structure and chemical shift changes of phosphorylated Serine are shown (left and middle panels). Overlay correlation spectra of histone H3 unmodified (black) and phosphorylated at Ser10 (red)[37].

Motivations and aims

Although c-Src has been extensively studied worldwide for decades, little is known about the function of its intrinsically disordered Unique domain (USrc). In a recent work a rich repertoire of interactions involving specific residues in the Unique domain has been uncovered, suggesting a more specific role than that of a simple spacer.

This thesis aims to uncover the functional role of the intrinsically disordered domain of human c-Src starting from the atomistic comprehension of the isolated domain (structural characterization) to the understanding of the biological functions in which it could be involved inside the cell (functional characterization).

c-Src is also implicated in several types of cancer. In contrast to viral oncogenes, the oncogenic potential of c-Src is not associated to mutations but to deregulation of signalling networks. Thus, the discovery of a new regulation mechanism involving the Unique domain could led to the development of new specific drugs that can inhibit the oncogenic activity of c-Src. Targeting the Unique domain has the potential of a much higher selectivity and, potentially, cleaner effects with respect to most existing drugs that, acting on the ATP binding site of SFKs, suffer from poor specificity. These perspectives further strengthen the biological relevance of this research.

REFERENCES

1. Bloch, F. Nuclear Induction. *Physical Review*, **70**, 460-474 (1946).

2. Purcell, E. M., Torrey, H. C., & Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review*, **69**, 37-38 (1946).

3. Varani L, Gunderson SI, Mattaj IW, Kay LE, Neuhaus D, Varani G. The NMR structure of the 38 kDa U1A protein - PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nature Struct. Biol* **7**, 329-335 (2000).

4. Iwahara, J., Schwieters, C. D., & Clore, G. M. Characterization of nonspecific protein-DNA interactions by 1H paramagnetic relaxation enhancement. *Journal of the American Chemical Society*, **126**, 12800-12808 (2004).

5. Roosild, T. P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., & Choe, S. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science*, **307**, 1317-1321 (2005).

6. Walter, T. *et al.* Visualization of image data from cells to organisms. *Nat. Methods* **7,** S26-S41 (2010).

7. Leis, A., Rockel, B., Andrees, L., Baumeister, W. Visualizing cells at the nanoscale. *Trends Biochem Sci* **34**, 60-70 (2009).

8. Robinson, C. V, Sali, A. & Baumeister, W. The molecular sociology of the cell. *Nature* **450,** 973-82 (2007).

9. Serber, Z. *et al.* Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. *Nat. Protoc.* **1,** 2701-9 (2006).

10. Tochio, H. Watching protein structure at work in living cells using NMR spectroscopy. *Curr. Opin. Chem. Biol.* **16,** 609-613 (2012).

11. Ito, Y. & Selenko, P. Cellular structural biology. *Curr. Opin. Struct. Biol.* **20,** 640-8 (2010).

12. Pons, M. (Ed.) NMR in Supramolecular Chemistry, Kluwer, Dordrecht, (1999).

13. Bertrand, K., Reverdatto, S., Burz, D. S., Zitomer, R. & Shekhtman, A. Structure of proteins in eukaryotic compartments. *J. Am. Chem. Soc.* **134,** 12798-806 (2012).

14. Selenko, P., Serber, Z., Gadea, B., Ruderman, J. & Wagner, G. Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 11904–9 (2006).

15. Sakai, T. *et al.* In-cell NMR spectroscopy of proteins inside Xenopus laevis oocytes. *J. Biomol. NMR* **36,** 179–88 (2006).

16. Amata, I. *et al.* Multi-phosphorylation of the intrinsically disordered unique domain of c-Src studied by in-cell and real-time NMR spectroscopy. *Chembiochem* **14,** 1820–7 (2013).

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

18. Ogino, S. *et al.* Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O. *J. Am. Chem. Soc.* **131,** 10834–5 (2009).

19. Burz, D.S., Dutta, K., Cowburn, D., Shekhtman, A. In-cell NMR for protein-protein interactions (STINT-NMR). *Nat Protoc* **1**, 146-152 (2006).

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

21. Ikeya, T. *et al.* NMR protein structure determination in living E. coli cells using nonlinear sampling. *Nat. Protoc.* **5,** 1051–60 (2010).

22. Lippens, G., Landrieu, I. & Hanoulle, X. Studying posttranslational modifications by in-cell NMR. *Chem. Biol.* **15,** 311–2 (2008).

23. Selenko, P. *et al.* In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **15,** 321–9 (2008).

24. Bodart, J.-F. *et al.* NMR observation of Tau in Xenopus oocytes. *J. Magn. Reson.* **192,** 252–7 (2008).

25. Hänsel, R. *et al.* Evaluation of parameters critical for observing nucleic acids inside living Xenopus laevis oocytes by in-cell NMR spectroscopy. *J. Am. Chem. Soc.* **131,** 15761–8 (2009).

26. Banci, L., Barbieri, L., Luchinat, E. & Secci, E. Visualization of redox-controlled protein fold in living cells. *Chem. Biol.* **20,** 747–52 (2013).

27. Banci, L. *et al.* Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nat. Chem. Biol.* **9,** 297–9 (2013).

28. Freedberg, D. I. & Selenko, P. Live cell NMR. *Annu. Rev. Biophys.* **43,** 171–92 (2014).

29. Theillet, F.-X. *et al.* Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* 217–236 (2012).

30. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).

31. Radivojac, P. *et al.* Intrinsic disorder and functional proteomics. *Biophys J.* **92,** 1439-1456 (2007).

32. Cohen, P. The origins of protein phosphorylation. *Nat. Cell. Biol.* **4**, 127–130 (2002).

33. Theillet, F.-X. *et al.* Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat. Protoc.* **8,** 1416–32 (2013).

34. Du, J.T. *et al.* Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J. Am. Chem. Soc.* **127**, 16350-16351 (2005).

35. Ramelot, T. A. & Nicholson, L. K. Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR. *J. Mol. Biol.* **307**, 871–884 (2001).

36. Landrieu, I. *et al.* Molecular implication of PP2A and Pin1 in the Alzheimer's disease specific hyperphosphorylation of Tau. *PLoS ONE* **6**,e21521 (2011).

37. Liokatis, S. *et al.* Phosphorylation of histone H3 Ser10 establishes a hierarchy for subsequent intramolecular modification events. *Nat. Struct. Mol. Biol.* **19,** 819–23 (2012).

I.3 – Objectives

The objectives of this thesis are:

Objective 1 – Effect of mutations perturbing the ULBR of the Unique domain of c-Src.

The first objective of this thesis was the *in vitro* characterization of the effect of mutations in the ULBR on the previously determined main interactions of the Unique domain: binding to lipids and to the SH3 domain.

Objective 2 - Phosphorylation of c-Src Unique domain in Xenopus laevis oocytes and mammalian cell extracts.

Phosphorylation of the Unique domain of c-Src is one of the modulation parameters affecting lipid binding and had previously been reported to affect c-Src activity.
The second objective was to study the phosphorylation of the isolated Unique domain in *Xenopus laevis* oocytes, that enabled the study of phosphorylation in a living cell and in cell extracts, that allowed further manipulation of the phosphatase/kinase network.

Objective 3 - Functional studies of the Unique domain in the context of the full-length protein.

The last objective was the evaluation of the functional significance of the modifications investigated in the isolated Unique domain as objectives 1 and 2 in a human colorectal cell line and in the context of the full length protein.

Part II

Results and Discussion

II.1 – Effect of mutations perturbing the ULBR of the Unique domain of c-Src

INTRODUCTION

c-Src is the leading member of the Src family of non-receptor tyrosine kinases (SFKs) and is involved in many signaling pathways[1,2,3]. The c-Src protein is constituted by: a N-terminal region formed by a short membrane-anchoring SH4 domain with a Glycine in position 2 within a site of myristoylation, and the intrinsically disordered Unique domain; the two folded SH3 and SH2 domains; the SH1 or kinase domain and a C-terminal tail.

Although c-Src was the first discovered oncogene[2], and since then extensively studied worldwide, the role of its intrinsically disordered N-terminal region, called the Unique domain, has remained obscure.

A recent publication has shown that the Unique domain of c-Src presents a previously unrecognized regulatory region (Unique Lipid Binding Region – ULBR) involved in lipid binding and intramolecular allosterically modulated interactions with the SH3 domain[4] (Figure II.1.1). For the first time it has been demonstrated that the intrinsically disordered Unique domain of c-Src might play an important role in the protein function. This discovery leads to the need of a full characterization of the Unique domain of c-Src, and in particular to the atomistic comprehension of the mechanisms of actions of the lipid / SH3 binding residues of the Unique domain, in the regulation of the whole protein. Further structural and functional studies are needed to elucidate the details of possible mechanisms in which the Unique domain may be involved.

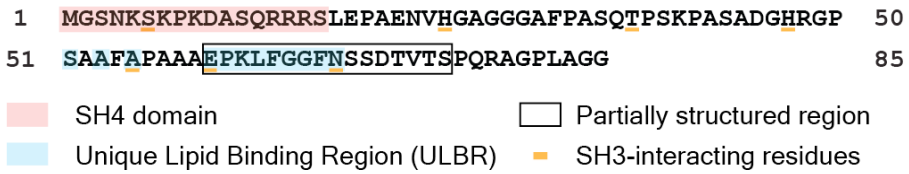


Figure II.1.1. Aminoacids Sequence of the Unique domain of human c-Src. In the red box residues belonging to the SH4 domain are shown. Residues perturbed by lipid binding are shown in the blue box (ULBR). Orange dashes refer to SH3-interacting residues. Black box indicates the previously identified partially structured region[19]. Myristoylation occurs on Glycine 2 after Methionine processing.

RESULTS

ULBR mutations modulate lipid binding by the Unique domain

Previous studies demonstrated that residues Ser 51, Ala 53, Ala 55 and ^{60}EPKLFGGFN^{68} belonging to the ULBR were highly perturbed in presence of negatively-charged lipids[4]. In order to provide further insights into the biological role of this novel region, mutations affecting residues located in the core of ULBR were introduced by site-directed mutagenesis. A charge-conserving mutant was generated by replacing residues 63-65 (LFG) of the Unique domain (USrc) of human c-Src by ^{63}AAA^{65} (mutant AAA). A second mutant had residues 63-68 of the c-Src ULBR (LFGGFN) replaced by ^{63}AAAEAE^{68} (mutant EAE), where two additional negative charges (E66 and E68) were inserted in order to electrostatically abolish binding to negatively charged membranes.

The effect of these mutations on USrc lipid binding was studied by NMR. ^{1}H-^{15}N HSQC spectra of ^{15}N-labelled constructs (residues 1-85 of c-Src) of mutants AAA and EAE (Figure S1) were acquired in the presence of bicelles containing variable proportions of dimyristoyl phosphatidyl choline (DMPC) and the acidic lipid dimyristoyl phosphatidyl glycerol (DMPG). Bicelles are formed when bilayer-forming long-chain lipids are mixed with short chain lipids (such as dihexanoyl phosphatidylcholine, DHPC) or detergent molecules[5,6,7]. They present a disk-like structure, with a central planar bilayer formed by the long-chain phospholipids, surrounded by a rim of short-chain phospholipids or detergent protecting the long-chain lipid tails from water (Figure II.1.2.A). These model membranes have been previously used to study protein-lipid interactions[8,9,10].

By comparing the lipid induced chemical shift perturbations of the wild type form of USrc (previously measured by Y. Perez)[4] with those of the mutants (AAA, EAE), it becomes clear that, while the binding by the

SH4 domain (residues 5-22) is not perturbed, lipid binding by the ULBR is, indeed, abolished (Figure II.1.2.B-C).

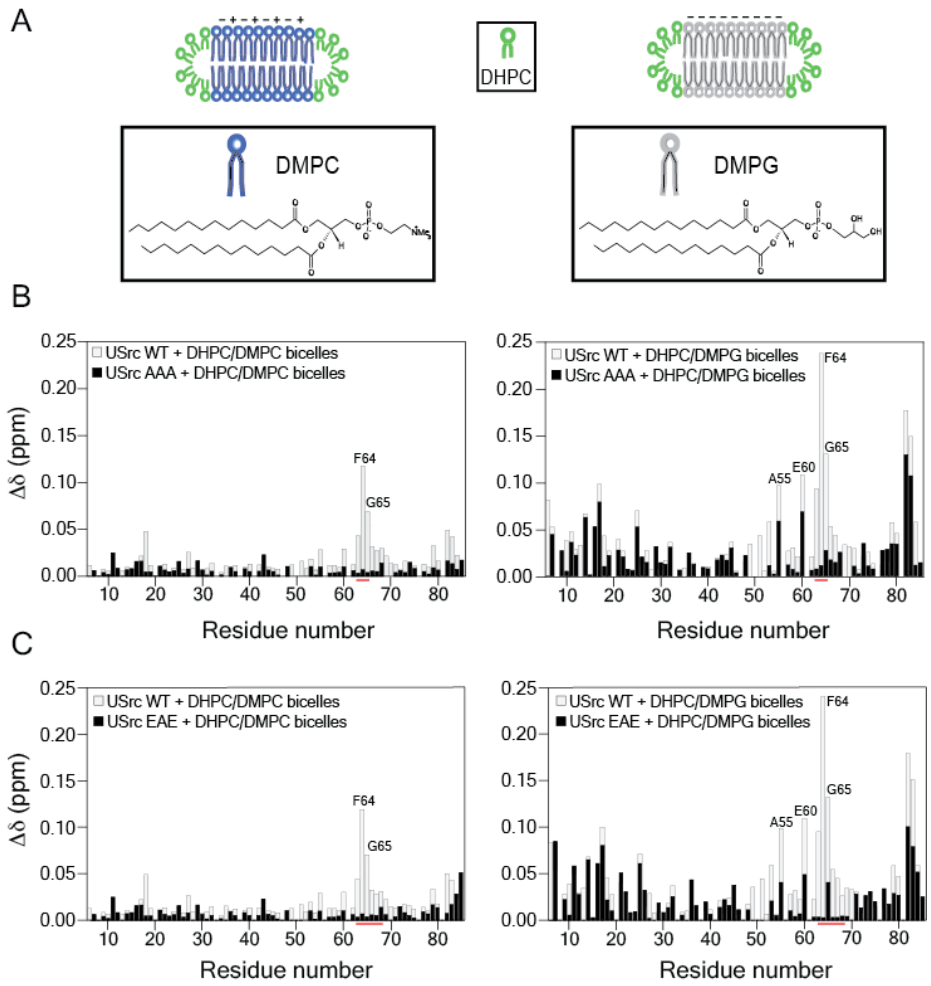


Figure II.1.2. Lipid binding by the Unique domain of human c-Src. (A) Schematic representation of the structure and charge of the lipid bicelles used (DMPC, left panel; DMPG, right panel). (B-C) Combined absolute values of ^{1}H-^{15}N chemical shift changes induced by lipid bicelles (DHPC/DMPC, left panels; DHPC/DMPG, right panels) in mutant USrc AAA (B) and mutant USrc EAE (C) are shown. Total lipid concentration was 8% w/v and the ratio of long chain lipids to DHPC (q) was 0.8. Combined NH chemical shift differences ($\Delta\delta$) were computed as in equation 1 (see Material and Methods).

Lipid binding by the SH3 domain of human c-Src had been observed[4]. Thus, the ULBR is located between two other lipid binding regions: the SH4 and SH3 domains.

Therefore, we next investigated whether mutations in the ULBR could affect the lipid binding capability of a construct with linked Unique and SH3 domains (USH3). ^{1}H-^{15}N HSQC spectra of the USH3-AAA construct were acquired in the presence of bicelles containing a mixture of short chain lipids (dihexanoyl phosphatidylcholine, DHPC) and long chain negatively-charged lipids (dimyristoyl phosphatidyl glycerol, DMPG). Chemical shifts perturbations are shown in Figure II.1.3.B. Significant shifts were observed for residues 98-103 and 115-118 belonging to the SH3 domain and for residues 7-13 and 21-23 belonging to the SH4 domain. Residues located in the ULBR that are perturbed by DHPC/DMPG bicelles lipids in USH3 wt (Figure II.1.3.A), were not perturbed in the USH3-AAA construct. A small-effect is still observed for residues Ala 55 and Glu 60 probably due to electrostatic repulsion of Glu60 when lipid binding takes place through other sites.

These results demonstrate that mutations in the ULBR affect the lipid binding capability of the Unique domain also in presence of its first neighbor domain (SH3). On the other hand, the SH4 and the SH3 domains are still capable to bind lipids.

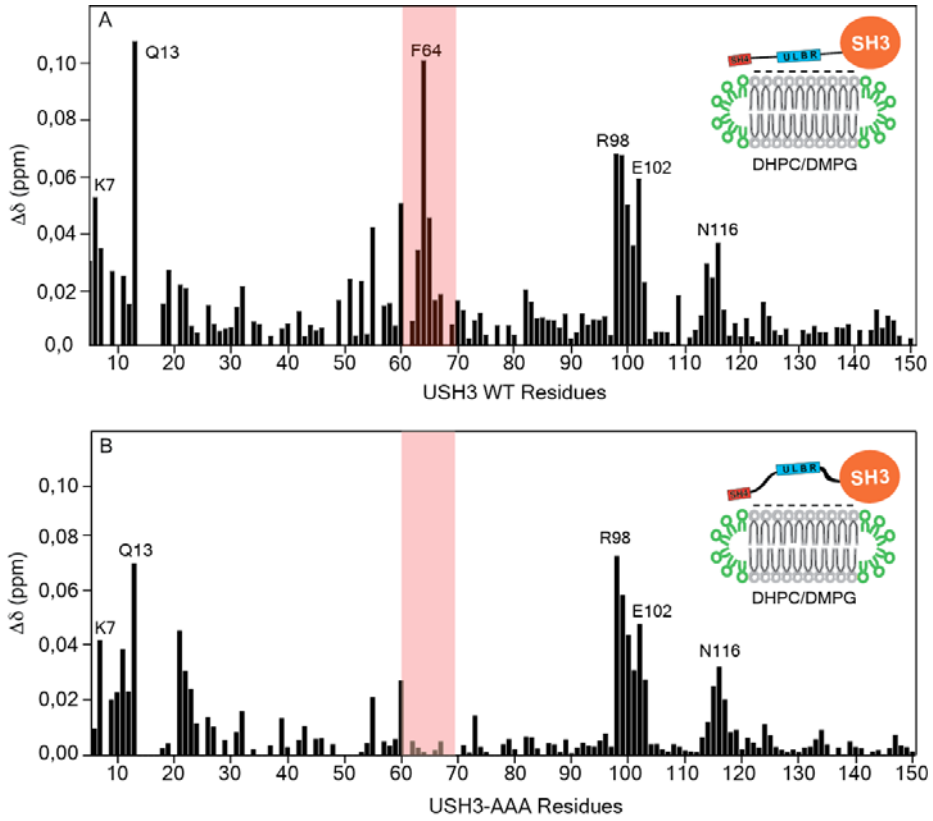


Figure II.1.3. Effect of AAA mutation on lipid binding by Unique-SH3 construct. (A) Combined absolute values ^{1}H-^{15}N NMR changes induced by DHPC/DMPG bicelles (8% w/v total lipid concentration, q = 0.8) in USH3 WT residues (previously measured by Y. Perez)[4]. (B) Combined absolute values ^{1}H-^{15}N NMR changes induced by DHPC/DMPG bicelles (8% w/v total lipid concentration, q = 0.8) in USH3-AAA construct. All spectra were recorded at 298 K and 600 MHz spectrometer. Combined NH chemical shift differences ($\Delta\delta$) were computed as in equation 1 (see Material and Methods). Red areas refer to the ULBR.

Lipid-Strips™ assays were carried out to test whether mutations in the ULBR could affect USrc lipid specificity (Figure II.1.4). Lipid Strips™ consist in hydrophobic membranes presenting on their surface phosphoinositides and/or other biologically relevant lipids[11]. They allow detecting protein binding to one or more of these lipids[12,13]. Bound proteins are detected using antibodies against the protein or a suitable affinity tag.

USrc mutants (AAA, EAE) displayed the same lipid specificity of the wild type form suggesting that, in these conditions, the observed signals came mainly from the SH4 binding to the acidic lipids present on the membrane surface (Figure II.1.4.A). In particular, USrc was shown to bind preferentially phosphatidic acid (PA), cardiolipin (CL), phosphatidylserine (PS), phosphatidylinositol-4-phosphate (PtdIns(4)P),phosphatidylinositol-4,5-diphosphate (PtdIns(4,5)P) and phosphatidylinositol-3,4,5-diphosphate (PtdIns(3,4,5)P). Relative intensities are shown in Figure II.1.4.B. Mutant EAE binds with less affinity PS and CL with respect to AAA mutant and USrc wt. No changes between mutants and wt are observed for phosphoinositide phosphate binding. Mutant EAE displays also a small increase for PA affinity with respect to the other USrc forms. These data suggest that mutations in the ULBR have a small overall effect on lipid binding by the Unique domain of c-Src. ULBR is not essential for the lipid specificity.

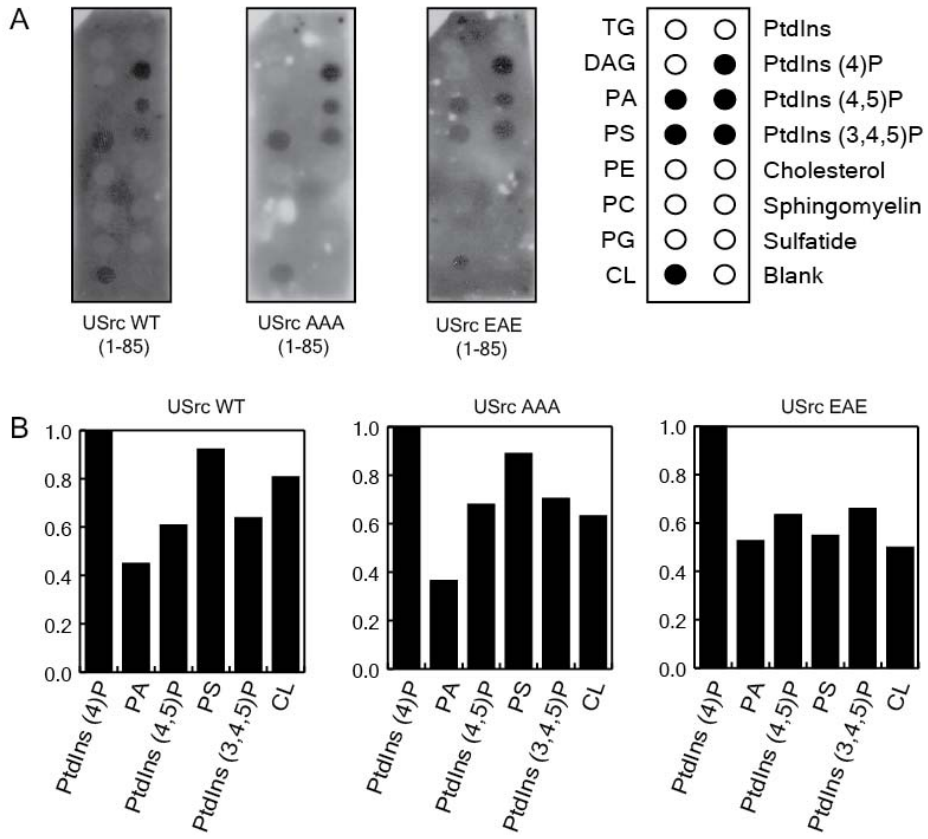


Figure II.1.4. Lipid Strips™ assays. (A) Immunoblots showing the lipid specificity of the wild type (left panel) and mutants form (middle, right panels) of the Unique domain of human c-Src. Bound protein to the immobilized lipids were detected by using an anti-Strep-Tag HRP-conjugated antibody (TG triglyceride, DAC diacylglyceride, PA phosphatidic acid, PS phosphatidylserine, PE phosphatidylethanolamine, PC phosphatidylcholine, PG phosphatidylglycerol, CL cardiolipin, PtdIns phosphatidylinositol). (B) Relative intensities of lipid binding of the wt and mutant forms of USrc. Values were calculated separately for each sample (WT, AAA, EAE) and normalized with respect to the maximum intensity coming from PtdIns(4)P.

Interaction between the Unique and SH3 domains is not affected by ULBR mutations

The regions defined by residues perturbed by lipid binding and by the interaction with the SH3 domain overlap (Figure II.1.1). Therefore, we investigated whether ULBR mutations affecting lipid binding also had an effect on the inter-domain interaction with the SH3.

We introduced the ^{63}AAA^{65} mutation in the linked Unique-SH3 construct (USH3-AAA) and we compared chemical shifts with those of the isolated domain (USrc-AAA). Results are shown in Figure II.1.5.A. Significant changes (larger than 2-fold with respect to the average shifts) were observed for residues Ser 6, Thr 37, Ala 55, Glu 60 and Asn 68 within the Unique domain. In Figure II.1.5.B, differences in the chemical shift perturbations for the isolated SH3 domain and USH3-AAA are shown. Residues of the SH3 domain in positions 98-103, 114-115, His 125, Gly 130, and Tyr 134 were strongly perturbed. Thus, residues involved in the inter-domain interaction are the same observed for the wild type form of the USH3 construct (Figure II.5.C-D). AAA mutation has no effect on the SH3 binding.

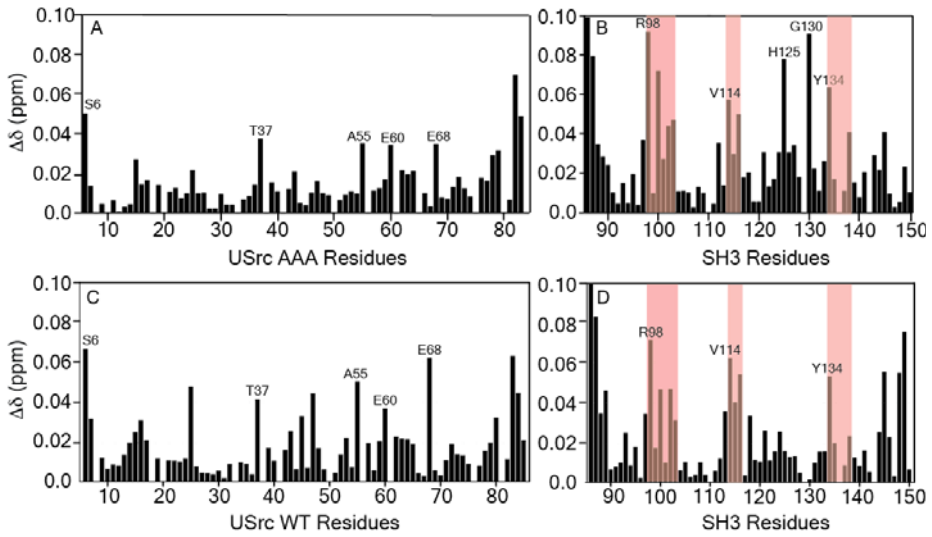


Figure II.1.5. AAA mutation does not affect the Unique-SH3 interaction. Combined absolute values ^{1}H-^{15}N NMR chemical shift changes between linked and isolated domains. (A) USH3-AAA versus USrc-AAA. (B) USH3-AAA versus SH3. (C) USH3 WT versus USrc WT. (D) USH3 WT versus SH3. Combined NH chemical shift differences ($\Delta\delta$) were computed as in equation 1 (see Material and Methods).

We further investigated the inter-domain interaction by Paramagnetic Relaxation Enhancement (PRE) experiments[14]. PRE experiments provide information on the proximity of different nuclei to a free-radical probe attached to a cysteine. PREs are extremely sensitive to transient close encounters between nuclei and paramagnetic labels. The closeness of the unpaired electron enhances the transversal relaxation rate of nuclear spins; as consequence, a paramagnetic 'broadening' is observed for residues in proximity of the spin label. Thus, even if the paramagnetic label is in close proximity only a small fraction of the time (transient contact), the PREs still significantly contribute to experimental relaxation rates. We used PREs as a tool of investigation in order to detect alternative contacts induced by the AAA mutation in

the Unique domain of c-Src. The free-radical paramagnetic probe MTSL (1-Oxyl - 2,2,5,5 - tetramethyl - Δ3 - pyrroline – 3 - methyl , methanethiosulfonate)[15,16] was attached to a cysteine[17] introduced by site-specific mutagenesis at position 59, close to the ULBR, in the Unique-SH3 linked construct carrying the AAA mutation (USH3-AAA-A59C). The ratios between the intensities measured in paramagnetic (I_o) and diamagnetic (reduced) (I_R) samples are shown in Figure II.1.6.A. The mayor effects on the SH3 domain induced by the presence of the MTSL-probe are observed in the highlighted regions, corresponding to residues 98-103, 114-116, and 133-137. The affected amino acids matched those that suffered the largest chemical shifts perturbation between SH3 and USH3 (Figure II.1.5.B). The same effect was observed for the wt form of the Unique domain (USH3-A59C) as shown in Figure II.1.6.B. Thus, the results confirm that the ^{63}AAA^{65} mutation does not affect the inter-domain interaction between the Unique and SH3 domain.

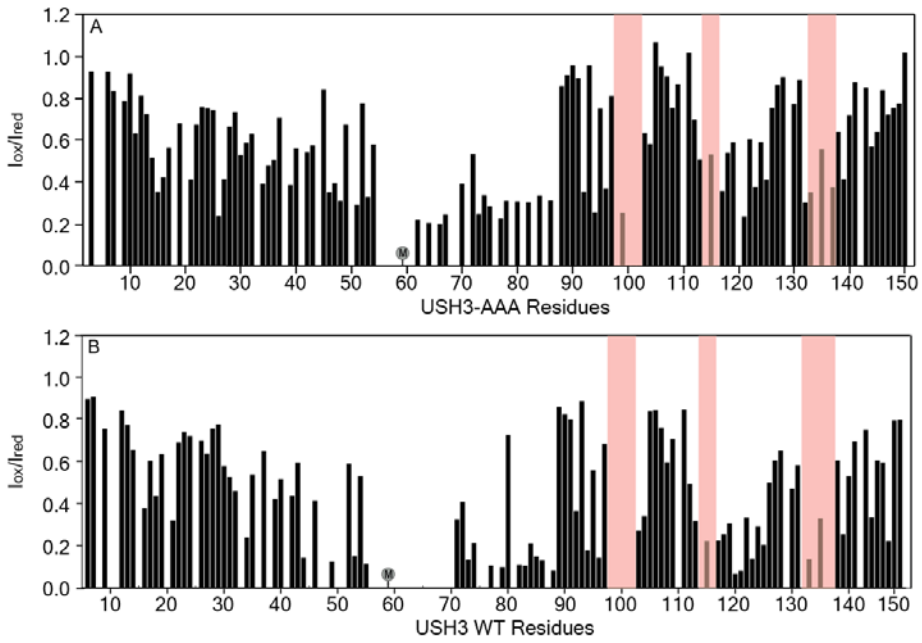


Figure II.1.6. PRE experiments of amide protons in spin-labeled USH3-AAA. (A) Intensity ratios of NH cross-peaks from MTSL-labeled (USH3-AAA-A59C at pH 7.0) between paramagnetic (I_{ox}) and diamagnetic (I_{red} – Ascorbic Acid reduced) forms. (B) Intensity ratios of NH cross-peaks from MTSL-labeled (USH3 WT -A59C at pH 7.0) between paramagnetic (I_{ox}) and diamagnetic (I_{red}) forms. Spectra were recorded at 278 K and 298 K in a 600 MHz spectrometer. M = MTSL-tag. Red highlighted regions represent affected amino acids matching those that suffered the largest chemical shifts perturbation between SH3 and USH3 (Figure II.1.5.B-D).

The binding of a Polyproline peptide to the SH3 domain still prevents the Unique-SH3 interaction

The SH3 domain of c-Src is known to bind to two classes of ligands containing the consensus motif P*xx*P flanked by the presence of charged residues[18]. Previous results[4] have demonstrated that the interaction between a polyproline peptide and the SH3 domain of human c-Src allosterically prevents the Unique-SH3 domain. Moreover, the loss of the intra-molecular interaction between the SH3 and Unique domains of c-Src is not a result of a direct competition with the peptide but an allosteric effect, as the binding sites for the two different ligands are located on two distinct regions within the SH3 domain. In order to test whether the ^{63}AAA^{65} mutation could affect this mechanism, we have monitored the effect of the binding of a high-affinity polyproline peptide containing the P*xx*P motif in a polyproline II helix on the Unique-SH3 interaction by NMR spectroscopy. Increasing amounts (0 mM; 0.04 mM; 0.08 mM; 0.16 mM; 0.2 mM) of unlabeled VSL12 peptide (Ac-VSLARRPLPPLP-OH, GenScript) were added to an initial solution of the ^{15}N-labeled USH3-AAA protein (0.2 mM). Figure II.1.7 shows expansions (A-C) of NMR spectra including signals from the NH group of several residues (A55, E60, T37, T74) within the Unique domain that are perturbed in the presence of the peptide. At

the end of titration with the VSL12 peptide, residues in the Unique domain of USH3-AAA that are perturbed by the binding to the peptide, present chemical shift values corresponding to the isolated Unique domain (Figure II.1.7.D-F). This observation indicates that peptide binding to the SH3 domain prevents its interaction with the Unique domain. Thus, we can conclude that ULBR mutations do not affect the previous described mechanism involving the Unique and SH3 domains of human c-Src.

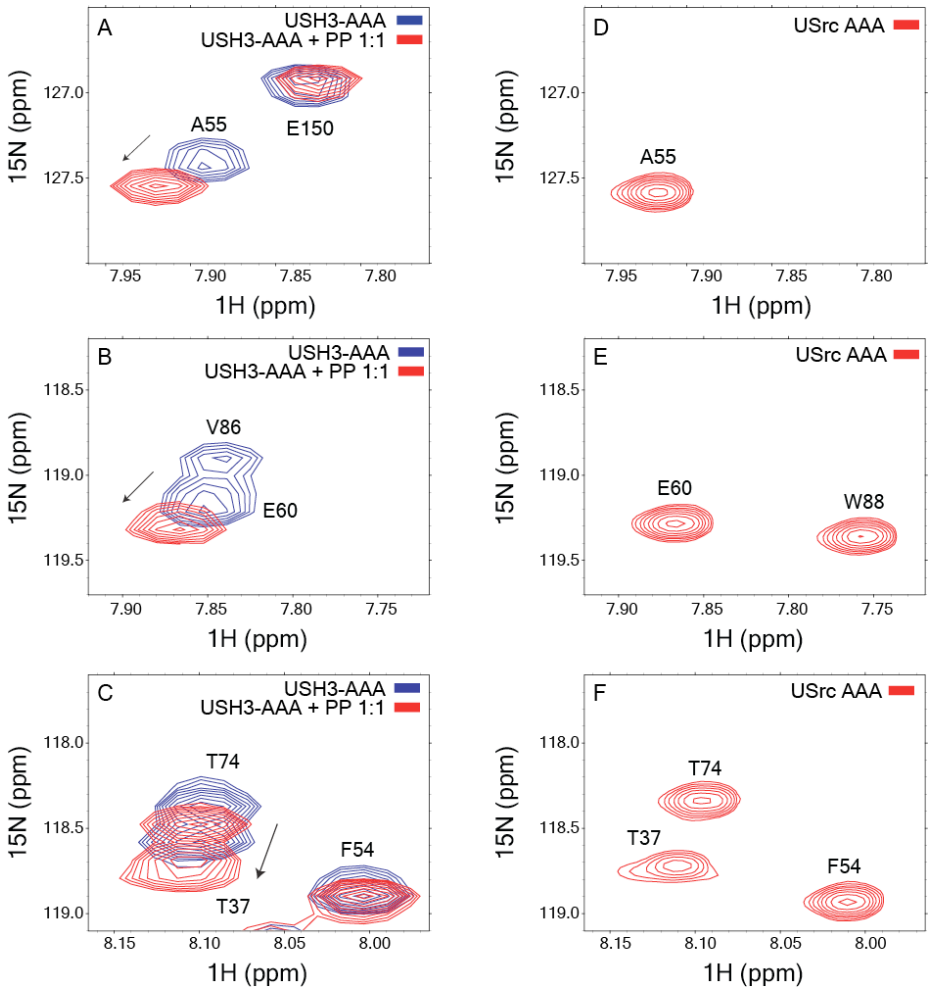


Figure II.1.7. Polyproline peptide binding to the SH3 domain impedes the Unique-SH3 interaction. Expansion of ^{1}H-^{15}N HSQC spectra showing residues A55 (A), E60 (A), T37 and T74 (C) of the Unique-SH3-AAA (USH3-AAA) construct in absence (black) and in presence of Ac-VSLARRPLPPLP-OH peptide (light grey) at molar ratio 1:1. Residues E150 (A) and F54 (C) belonging to the SH3 and Unique domains, respectively, act as a control. Residue V86 (SH3 domain) is not visible at the end of peptide titration because strongly affected by the binding to VSL12 peptide. Residue W88 belongs to the Strep-Tag. (D-F) Expansions of HSQC spectrum of the isolated Unique domain AAA (USrc-AAA) showing residues A55, E60, T37 and T74 (grey). All the spectra were recorded at 278 K and 600 MHz spectrometer.

The SH3 domain capability of binding lipids is modulated by the interaction with a polyproline peptide

The SH3 domain of human c-Src was shown for the first time to be able to interact with lipids through residues located in the RT and nSrc loops[4]. The SH3 - lipids interaction takes place for the isolated SH3 as well as for the USH3 construct, suggesting that the capability of binding lipids is an intrinsic feature of the SH3 domain. Here, we have investigated the effect of the binding of a high-affinity polyproline peptide on lipid interaction capability by the SH3 domain of human c-Src. ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled USH3-AAA were recorded in the presence of unlabeled VSL12 peptide (1:1 molar ratio) and DHPC/DMPG bicelles. Representative ^{1}H-^{15}N HSQC spectra are given in Figure S2.A. Chemical shifts perturbations differences between USH3 in presence of peptide and in presence of peptide plus bicelles, are shown in Figure II.1.8.B. Residues located in the SH4 domain were perturbed by the addition of bicelles while no significant changes were observed for the mutated ULBR and SH3 residues. Therefore, the SH3 domain is not able to interact with lipids when is interacting with a polyproline peptide.

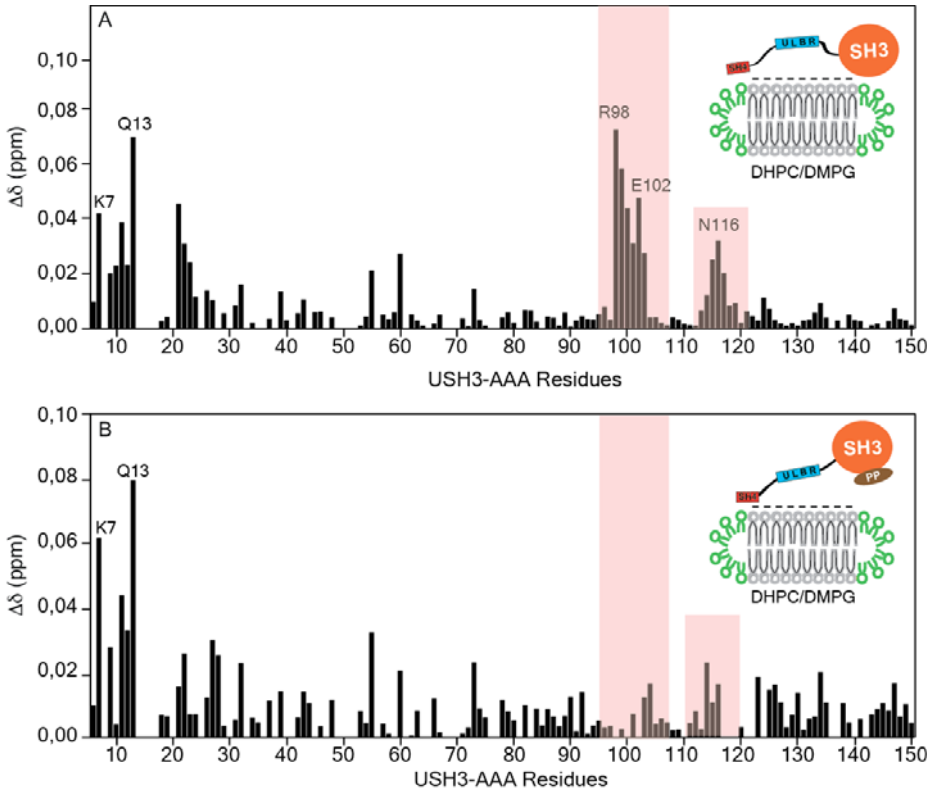


Figure II.1.8. SH3 - lipid interaction is prevented by the binding of a polyproline peptide to the SH3 domain. (A) Combined absolute values ^{1}H-^{15}N NMR changes induced by DHPC/DMPG bicelles (8% w/v total lipid concentration, q = 0.8) in USH3-AAA construct. (B) Combined absolute values ^{1}H-^{15}N NMR chemical perturbations between USH3-AAA in presence of VSL12 peptide (1:1 molar ratio) and in presence of VSL12 peptide (1:1 molar ratio) plus DHPC/DMPG bicelles (8% w/v total lipid concentration, q = 0.8). All spectra were recorded at 298 K and 600 MHz spectrometer. Combined NH chemical shift differences ($\Delta\delta$) were computed as in equation 1 (see Material and Methods). Red highlighted regions represent the perturbed SH3 residues in presence of bicelles.

Monitoring long-range interactions by PREs in the mutated Unique domain

The Unique domain of human c-Src presents a partially structured region comprising residues 60-75[19] (Figure II.1.1). This structural observation had allowed us to identify and characterize the Unique Lipid Binding Region. At the same time it was shown that the conformational space sampled by the Unique domain of c-Src is not random and favors conformations in which intra-molecular interactions take place. Paramagnetic Relaxation Enhancement (PRE) experiments were carried out to uncover transient interactions involving distant residues in USrc. In particular it was observed that residues belonging to the SH4 domain were able to transiently interact with residues located in the ULBR.

To test whether mutations in the ULBR could affect the local structure in the Unique domain of c-Src, we have compared the spectra of the wt form of USrc and of the AAA mutant.

A significant chemical shift variation (more than 0.5 ppm) is observed for Gly 66 (Figure II.1.9.A-B). This residue is the first neighbor aminoacid to the AAA mutation and it is located in the core of the ULBR. This huge perturbation suffered by Gly 66 suggested a major change in the chemical environment of this residue, which may be a consequence of the disruption of the *FGGF* motif present in the ULBR. Therefore, the AAA mutation modifies locally that structural feature.

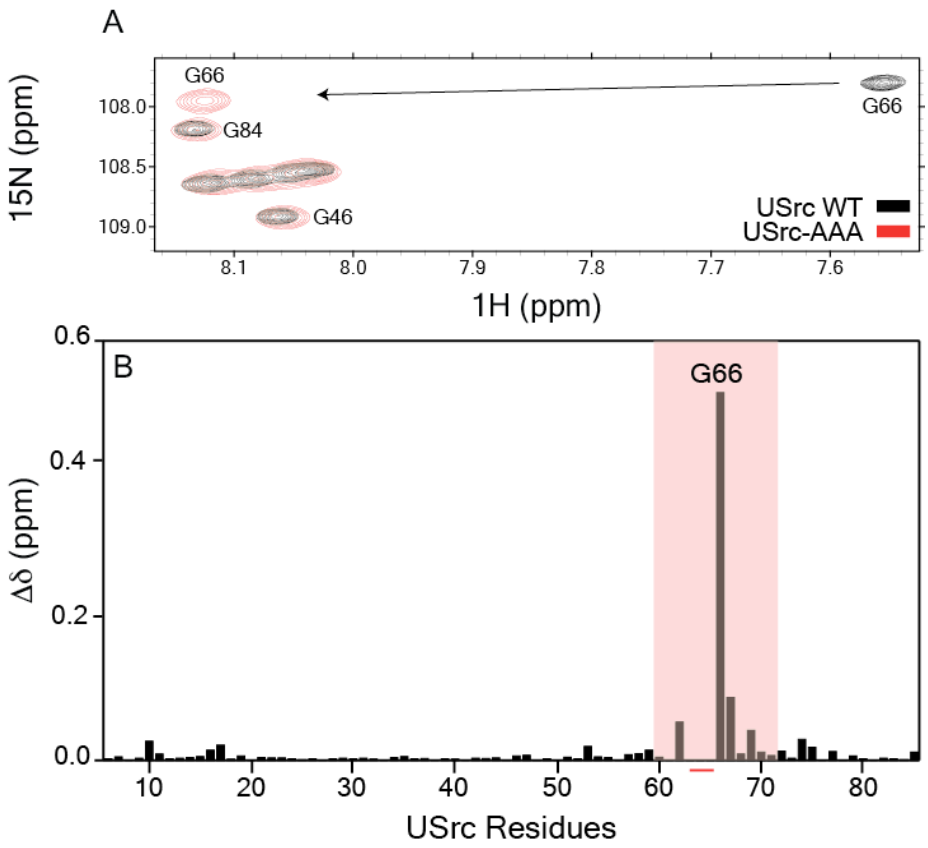


Figure II.1.9. AAA mutation disrupts the partially structured region within the ULBR. (A) Expansion of ^{1}H-^{15}N HSQC spectra showing USrc WT (black) and USrc-AAA (red). Glycine 66 is the most affected residue when the ULBR is mutated. (B) Combined absolute values ^{1}H-^{15}N NMR chemical shift changes between USrc WT and USrc-AAA. Combined NH chemical shift differences ($\Delta\delta$) were computed as in equation 1 (see Material and Methods).

Next, we have monitored long-range interactions within the Unique domain of c-Src by Paramagnetic Relaxation Enhancement (PRE) experiments. A Unique mutated construct (AAA mutant) with a Cysteine in position 59 (A59C) was designed to carry out PREs experiments in order to study the effect of ULBR mutations on the transient interactions taking place in the Unique domain.

From the N-terminal end to the beginning of the ULBR (Glu 60) the observed perturbations are conserved as for the wild type form in

normal conditions (Figure II.1.10.A-B). Regions between residues Gln 13 – Asn 23 and Gly 30 – Gln 36 show interactions taking place, only partially reduced with respect to the wild type. However, the small decrease in the paramagnetic effects is more pronounced for residues Gly 30 – Gln 36 than for those located in the SH4 domain (Gln 13 – Asn 23). On the other hand, towards the C-terminal end (from residues Glu 60), the PRE profile resembles much more the one of USrc wt in denaturing conditions (4M Urea) (Figure II.1.10.C-D). Nevertheless, PRE effects are still observed for residues located in the region 73-84.

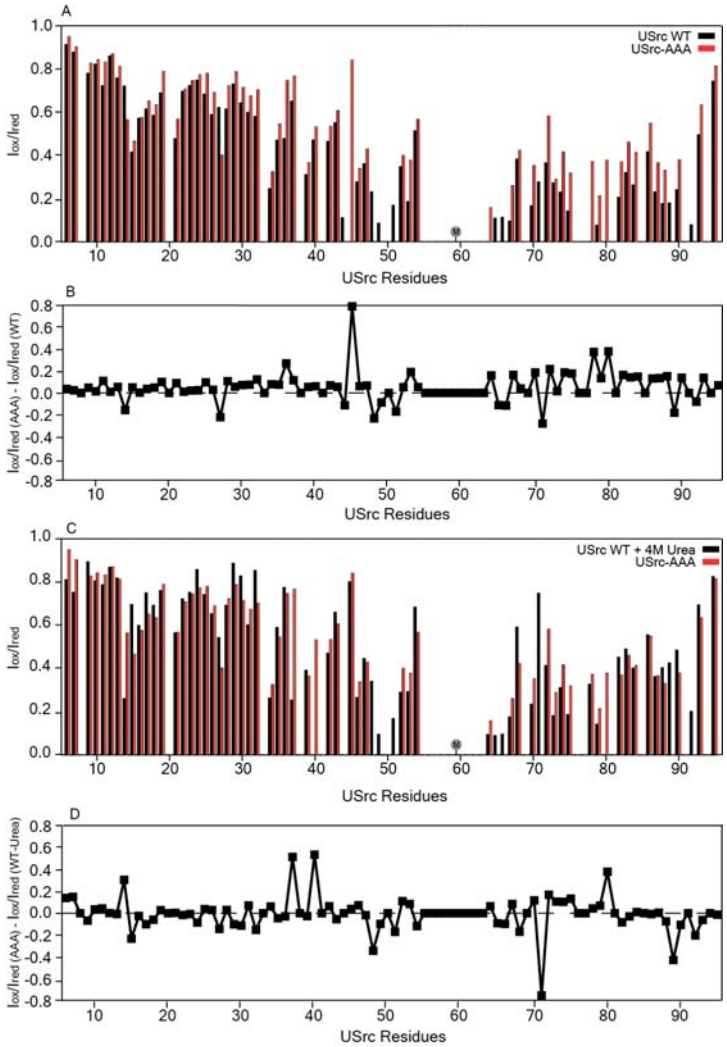


Figure II.1.10. PRE experiments of amide protons in spin-labeled USrc. PREs are plotted as the ratio of the intensity in the oxidized and the reduced states (I_{ox}/I_{red}). (A) Comparison between PREs profiles of USrc wt (black columns) in standard conditions and the USrc-AAA mutant (red columns). (B) Differences between the ratio of the intensity (I_{ox}/I_{red}) of the AAA mutant and the ratio of the intensity (I_{ox}/I_{red}) of the WT in standard conditions. (C) Comparison between PREs profiles of USrc wt (black columns) in denaturing conditions (4M Urea) and mutant USrc-AAA (red columns). (D) Differences between the ratio of the intensity (I_{ox}/I_{red}) of the AAA mutant and the ratio of the intensity (I_{ox}/I_{red}) of the WT in denaturing (4M Urea) conditions. M = MTSL-tag.

DISCUSSION

Functional studies have shown that mutations involving the Unique Lipid Binding Region (AAA, EAE mutants) cause an aberrant phenotype in the maturation process of *Xenopus Laevis* oocytes. This observation demonstrates that the ULBR plays an essential role in the regulation of c-Src. The ULBR and residues surrounding it are involved at least in two important mechanisms: lipid binding and inter-domain interaction with the SH3 domain. The capability to bind lipids by the ULBR has been abolished by replacing residues ^{63}LFG^{65} (AAA mutant) or ^{63}LFGGFN^{68} (mutant EAE) of the Unique domain of human c-Src. This effect has been observed for the isolated (USrc) as well for the linked construct (USH3). However, the loss of interaction does not affect lipid specificity because ULBR mutants display the same recognition pattern of the wild type form.

On the other hand, residues previously described to be involved in the Unique-SH3 interaction were still perturbed in the USH3-AAA construct demonstrating that ULBR mutations do not affect the inter-domain interaction with the SH3 domain. The Unique-SH3 binding was further confirmed by the fact that the previously described allosteric

modulation of a polyproline peptide (VSL12) on the inter-domain interaction is still conserved.

The AAA mutation was designed to partially remove the *FGGF* motif present in the ULBR. This motif formed by the sequence FGG(F/V) followed by a stretch of small polar residues (N/D/S/T) is highly conserved in c-Src of a range of species and in the Unique domain of Fyn and Yes, the two kinases most similar to c-Src. The replacement of those aminoacids causes important local effects that affect the structure of the ULBR. Above all, the major change is observed for Glycine 66. This residue was the only one presenting negative RDCs[19] in a USrc region (63-70) presenting positive values, probably indicating the enhanced local flexibility of this aminoacid and, therefore, its structural importance in the *FGGF* motif (Figure I.1.6.B). However, the modification of the local structure has no effect on long-range interactions taking place in the Unique domain of c-Src. PREs data show that the ULBR - SH4 interaction is still retained in spite of the mutation.

Thus, we have been able to discriminate between alternative interactions involving the Unique domain. Indeed, the AAA mutation mainly affect the ULBR lipid capability probably modifying its structural features that are required for this interaction while no effects have been observed for the inter-domain as well for the intra-domain interactions with the SH3 and SH4 domains in the absence of lipids. The ULBR represents a second lipid binding region within the Unique domain that could cooperate with the SH4 domain in order to correctly locate the protein on the membrane surface (“positional regulation”). The relative position of c-Src on the membrane could be essential for the regulation of the various signaling pathways that are controlled by c-Src. By abolishing the ULBR ability of binding lipids we are modulating the mechanism with which c-Src effects on the different upstream and downstream partners at membrane level.

A third important lipid binding region is present in c-Src: the SH3 domain. Our results have shown that the SH3 domain of c-Src loses its ability to interact with lipids in presence of a polyproline peptide. This observation suggests that a regulation mechanism involving lipid binding may also modulate the SH3 domain. Residues involved in the lipid binding are located on the opposite face of the SH3 domain where the cleft that recognizes polyproline motifs is. In the inactive basal state of c-Src the SH3 domain interacts with the polyproline linker region connecting the SH2 domain and kinase domain. This interaction is released when c-Src is activated[20,21]. The unbound SH3 is then able to interact with lipids or with the Unique domain. The SH3 domain, as well the SH2, is crucial for targeting c-Src to the different cellular membranes[22]. In effect, mutations in the SH2 and SH3 domains have been shown to differentially affect localization to various cellular membranes[22]. Presumably these domains interact with proteins, or lipids[23], that are localized to specific membrane systems.

MATERIAL AND METHODS

Cloning and mutagenesis

The cDNA encoding for human c-Src region (SH4, Unique domains, residues 1-85) with a Strep-tag (*SAWSHPQFEK*) in C-terminal position for purification purposes was cloned into a pET-14b vector (Novagen, UK) (Figure A.3.1). Human USH3 protein (SH4, Unique and SH3 domains of c-Src protein, residues 1–150, Figure S2.B) or human SH3 c-Src domain (residues 86–150) were expressed as His-GST fusion proteins using the pETM-30 vector (EMBL) (Figure A.3.2). A TEV cleavage site was inserted between the His-GST fusion protein and the proteins of interest (SH3, USH3). The glycine residue (Gly 2) following the initial methionine was mutated to alanine. Mutations were introduced using the QuikChange site-directed mutagenesis kit (Stratagene).

Protein expression and purification

Plasmids carrying the constructs of interest were transformed in *Escherichia coli* Rosetta™(DE3)pLysS cells (Novagen, UK). Cells were grown in Luria Broth (LB) with overnight incubation at 25°C. Expression was induced with 0.5 - 1 mM IPTG when an $O.D._{600}$ of 0.7 was reached. For ^{15}N isotopic labeling, cells were grown in M9 minimal medium supplemented with $[^{15}N]H_4Cl$ (Cambridge Isotope Laboratories, UK). Cells overexpressing USrc protein were harvested by centrifugation and resuspended in Tris-HCl 100 mM (pH 7.5), NaCl 150 mM, EDTA 1 mM, NaN_3 0.01%, treated for 30 min with DNAse and sonicated on ice (6 short burst of 30 sec separated by 30 sec). After centrifugation, USrc protein was isolated using Strep-tactin Sepharose (IBA, Gottingen) and further purified by size exclusion chromatography on a Superdex 75 26/60 (GE Healthcare, Spain) in sodium phosphate

(pH 7.0), EDTA 0.2 mM, NaN_3 0.01%. The expression and purification procedure for USrc is described in Ref. 24.

USH3 and SH3 constructs were expressed as His-GST fusion proteins. Cells overexpressing USH3 and SH3 proteins were harvested by centrifugation and resuspended in Tris-HCl 20mM (pH 8.0), NaCl 300 mM, Imidazole 5 mM, 0.01% NaN_3, treated for 30 min with DNAse and sonicated on ice (6 short burst of 30 sec separated by 30 sec). After centrifugation, USH3 and SH3 proteins were purified using a Ni-NTA resin (Qiagen). The purified proteins were cleaved from their fusion partner with TEV protease (S219V mutant with an N-terminal poly-histidine tag), and re-purified with Ni-NTA resin and size exclusion chromatography (Superdex 75 26/60, GE Healthcare, Spain) in sodium phosphate (pH 7.0), NaCl 150 mM, EDTA 0.2 mM, NaN_3 0.01%.

Preparation of bicelles

12.5% (w/w) bicelle dispersions were prepared by mixing long-chain (DMPC or DMPG) and short-chain (DHPC) phospholipids (Avanti Polar Lipids, Alabaster, AL) in chloroform or chloroform/methanol, evaporating the solvent and rehydrating the resulting lipid film in 50 mM phosphate buffer, pH 7.0. The mixture was subjected to five freeze-thaw cycles with pipetting and vortexing until the lipid solution was clear and transparent. The molar ratio of lipids in bicelle samples was 1.0 DHPC : 0.8 DMPG (q = 0.8, 44.4 % DMPG molar ratio). DMPC + DMPG/DHPC (q ratios) of 0.5 – 0.8 provide isotropic fast tumbling bicelles[7,25]. Concentrated protein solutions were added to the bicelles to a final concentration of 0.2 mM protein and 8 % (w/v) of lipids. NMR samples containing 10% D_2O were measured at 298K. The detailed protocol for bicelles preparation is described in Ref. 24.

Protein-phospholipid assays: Lipid Strips™

Lipid binding specificity was assessed with protein-lipid overlay assays using commercially available Lipid Strips™ blotted with 100 pmol of biologically relevant lipids (Echelon Biosciences), following manufacturer's instructions. Echelon Lipids Strips™ were blocked with 3% fatty acid-free BSA (A7030, Sigma) in TBS-Tween for 1 h at room temperature and then incubated with 1 μM USrc WT or mutant forms (10 mg/ml) in TBS-Tween for 1 h at room temperature. The membranes were washed with TBS-Tween and incubated with a 1:3,000 dilution of anti-Streptag HRP antibody (Novagen) in TBS-Tween for 1 h at room temperature. After washing, protein-lipid interactions were detected by enhanced chemiluminiscence (ECLTM detection, GE Healthcare). The detailed protocol is described in Ref. 24.

NMR Spectroscopy

NMR experiments were recorded in Bruker Avance 600 or 800 MHz spectrometers equipped with TCI cryo-probes at 0.2 mM protein concentration in 50 mM phosphate buffer pH 7.0 at either 278 or 298 K. The lower temperature allowed the observation of some NH signals from the Unique domain that exchanged rapidly. Data processing and analysis were carried out with NMRPipe[26] and Sparky[27] software. Reported chemical shift differences are from spectra obtained under identical conditions. USrc and USH3 assignments had been previously reported[4,19]. Combined NH chemical shift differences were computed as in equation (1):

$$\Delta\delta = [\,\Delta\delta_H{}^2 + (\Delta\delta_N / 5)^2\,]^{1/2}$$

where $\Delta\delta_H$ and $\Delta\delta_N$ are the changes in chemical shift for ^{1}H and ^{15}N, respectively.

For PRE experiments the spin label (1-oxy-2,2,5,5-tetramethyl-D-pyrroline-3-methyl)-methanethiosulfonate, MTSL, (Toronto Research Chemicals) was attached as previously described[17] to a cysteine residue introduced by mutation at position 59. Diamagnetic samples were measured after the paramagnetic sample by reducing the paramagnetic tag with the addition of a 5-fold excess of Ascorbic Acid. PRE effects were determined as the ratio of ^{1}H-^{15}N HSQC cross peak intensity, in the paramagnetic (I_{ox}) and diamagnetic (I_{red}) samples.

For the peptide titration the unlabeled VSL12 peptide (Ac-VSLARRPLPPLP-OH, GenScript) was added to an initial solution of 0.2 mM ^{15}N-labeled USH3-AAA. ^{1}H-^{15}N HSQC spectra were acquired at 278K and 298K in the presence of increasing stoichiometric amounts of peptide (0 mM; 0.04 mM; 0.08 mM; 0.16 mM; 0.2 mM).

REFERENCES

1. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287,** 121–149 (1996).

2. Martin, G. S. The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2,** 467–75 (2001).

3. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4,** 470–80 (2004).

4. Pérez, Y. *et al.* Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* **3,** 1295 (2013).

5. Dürr, U. H. N., Gildenberg, M. & Ramamoorthy, A. The magic of bicelles lights up membrane protein structure. *Chem. Rev.* **112,** 6054–74 (2012).

6. Dürr, U. H. N., Soong, R. & Ramamoorthy, A. When detergent meets bilayer: Birth and coming of age of lipid bicelles. *Prog Nucl Magn Reson Spectrosc.* **69**, 1–22 (2013). doi:10.1016/j.pnmrs.2013.01.001.

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

8. Whiles, J. a., Deems, R., Vold, R. R. & Dennis, E. a. Bicelles in structure–function studies of membrane-associated proteins. *Bioorg. Chem.* **30,** 431–442 (2002).

9. Biverståhl, H., Andersson, A., Gräslund, A. & Mäler, L. NMR solution structure and membrane interaction of the N-terminal sequence (1-30) of the bovine prion protein. *Biochemistry* **43,** 14940–7 (2004).

10. Andersson, A., Almqvist, J., Hagn, F. & Mäler, L. Diffusion and dynamics of penetratin in different membrane mimicking media. *Biochim. Biophys. Acta* **1661,** 18–25 (2004).

11. Dowler, S., Kular, G., and Alessi, D. R. Protein lipid overlay assay, *Sci STKE*, 2002, L6. (2002).

12. Sankarshanan, M., Ma, Z., Iype, T. & Lorenz, U. Identification of a Novel Lipid Raft-Targeting Motif in Src Homology 2-Containing Phosphatase 1. *J. Immunol.* **179,** 483–490 (2007).

13. Landeta, O. *et al.* Reconstitution of proapoptotic BAK function in liposomes reveals a dual role for mitochondrial lipids in the BAK-driven

membrane permeabilization process. *J. Biol. Chem.* **286,** 8213–30 (2011).

14. Bloembergen, N., & Morgan, L. O. Proton Relaxation Times in Paramagnetic Solutions. Effects of Electron Spin Relaxation. The Journal of Chemical Physics, **34**(3), 842 (1961).

15. Antoniou, C. & Fung, L. W.-M. Potential artifacts in using a glutathione S-transferase fusion protein system and spin labeling electron paramagnetic resonance methods to study protein-protein interactions. *Anal. Biochem.* **376,** 160–2 (2008).

16. Mchaourab, H. S., Lietzow, M. a, Hideg, K. & Hubbell, W. L. Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* **35,** 7692–704 (1996).

17. Bertoncini, C. W. *et al.* Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 1430–5 (2005).

18. Feng, S., Kasahara, C., Rickles, R. J. & Schreiber, S. L. Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 12408–15 (1995).

19. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391,** 136–48 (2009).

20. Xu, W., Doshi, a, Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3,** 629–38 (1999).

21. Bernadó, P., Pérez, Y., Svergun, D. I. & Pons, M. Structural characterization of the active and inactive states of Src kinase in solution by small-angle X-ray scattering. *J. Mol. Biol.* **376,** 492–505 (2008).

22. Kaplan, J. M., Varmus, H. E. & Bishop, J. M. The src protein contains multiple domains for specific attachment to membranes. *Mol. Cell. Biol.* **10,** 1000–1009 (1990).

23. Rameh, L. E., Chen, C. S. & Cantley, L. C. Phosphatidylinositol (3,4,5)P3 interacts with SH2 domains and modulates PI 3-kinase association with tyrosine-phosphorylated proteins. *Cell* **83,** 821–30 (1995).

24. Perez, Y., Maffei, M., Amata, I., Arbesú, M. & Pons, M. Lipid Binding by Disordered Proteins. *Protoc. Exch.* (2013). at http://dx.doi.org/10.1038/protex.2013.094

25. Vold, R. R., Prosser, R. S. & Deese, A. J. Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. *J. Biomol. NMR* **3**, 329-35 (1997)

26. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **3**, 277-93. (1995)

27. T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco

II.2 - Study of c-Src Unique domain in Xenopus laevis model system

INTRODUCTION

Post-translational modifications (PTMs) are crucial for cells to be able to quickly regulate protein function. Reversible phosphorylation, in particular, is a major regulatory event in controlling signaling pathways. Phosphorylation of serine and threonine residues is the most abundant inside cells. Commonly, the same protein can contain multiple phosphorylation sites that can regulate its function or its ability to interact with different partners[1].

Intrinsically disordered regions (IDRs) are highly abundant in the proteins of eukaryotic organisms, but not in those of prokaryotes, thus suggesting that their existence is linked to the increased regulatory demands of multicellular organisms[2,3]. IDRs are often associated with cell regulation and signaling, and 80% of the proteins involved in human cancers contain long IDRs[4,5]. PTMs sites are often located in exposed, flexible segments and IDRs often display phosphorylation sites[6].

The intrinsically disordered Unique domain of human c-Src contains several sites susceptible to phosphorylation and dephosphorylation. Phosphorylation of Ser 17 by PKA (cAMP-dependent protein kinase) is a well-characterized process but its biological significance remains obscure[7]. Other phosphorylation sites within the Unique domain that have been identified and are believed to modulate the protein activity are Thr 34, Thr 46, and Ser 72 in chicken c-Src, which correspond to Thr 37 and Ser 75 in human c-Src (Thr 46 has no equivalent in humans). Phosphorylation of c-Src at Ser 75 by Cdk5 is a crucial mechanism for the regulation of intracellular c-Src activity, as it is

directly implicated in modulating the ubiquitin-dependent degradation of active c-Src[8]. Moreover, *in vitro* studies have shown that phosphorylation of Ser 17, Thr 37, and Ser 75 can modulate lipid binding by the Unique domain[9,10] (Figure II.2.2). These phosphorylation sites are highly conserved in c-Src of phylogenetically distant species, in particular, human and *Xenopus Laevis* (Figure II.2.1).

```
                10        20        30        40        50
Hum  MGSNKSKPKDASQRRRSLEPAENVHGAGGGAFPASQTPSKPASADGHRGP
Mou  MGSNKSKPKDASQRRRSLEPSENVHGAGG-AFPASQTPSKPASADGHRGP
Chi  MGSSKSKPKDPSQRRRSLEPPDSTH-HGG--FPASQTPNKTAAPDTHRTP
XLa  MGATKSKPREGGPRSRSLDIAEGSHQPFT-SLSASQTPSK--SLDSHRPS

                60        70        80
Hum  SAAFAPAAAEPKLFGGFNSSDTVTSPQRAGPLAGG
Mou  SAAFVPPAAEPKLFGGFNSSDTVTSPQRAGPLAGG
Chi  SRSFGTVATEPKLFGGFNTSDTVTSPQRAGALAGG
XLa  GQPFG-GNCDLTPFGGVNFSDTITSPQRTGPLAGG
```

RxxS site Ser/Thr conserved phosphorylation sites

Figure II.2.1. Sequence alignment of the Unique domain of c-Src. Sequence alignment of the Unique domains of human (Hum), mouse (Mou), chicken (Chi) and Xenopus laevis (XLa) c-Src is shown. Yellow and green boxes represent, respectively, conserved RxxS site and residues Thr 37 and Ser 75 that are conserved among different species.

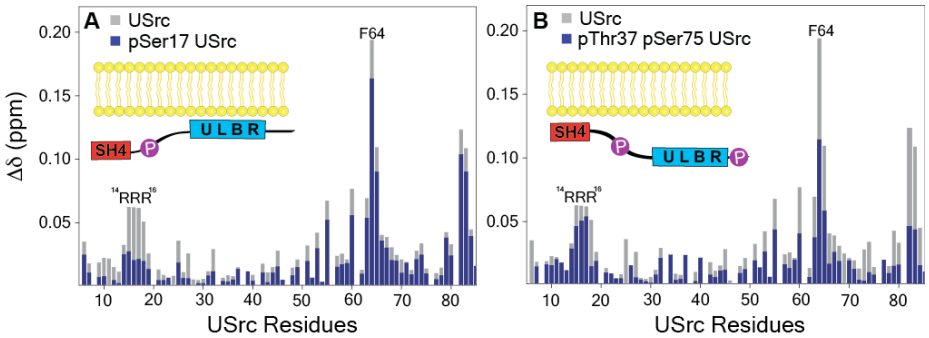


Figure II.2.2. Effect of phosphorylation on lipid binding by USrc WT. Combined ^{1}H-^{15}N NMR shifts induced by DHPC/DMPG bicelles (8% w/v total lipid concentration, q = 0.8) in USrc unphosphorylated (A, B gray), mono-phosphorylated at Ser 17 (A, blue), and di-phosphorylated at Thr 37 and Ser 75 (B, blue).

RESULTS

USrc phosphorylation by endogenous kinases in Xenopus Laevis oocytes

In-cell NMR spectroscopy allows the study of protein conformation and activation in the physiological environment of living cells[11]. Under the intracellular crowded conditions proteins can undergo changes that may alter their structure or function. We aimed to investigate the intrinsically disordered Unique domain of human c-Src inside cells using *Xenopus Laevis* as a model system. A ^{15}N-labeled construct (USrc WT, residues 1-85) was first expressed in *Escherichia coli*, purified and then injected into living oocytes. Figure II.2.3 shows ^{1}H-^{15}N correlations of the NMR spectra of USrc, either injected into intact stage VI *Xenopus* oocytes (Figure II.2.3.A) or dissolved in buffer (Figure II.2.3.B). Well-resolved NMR spectra were obtained under the two different conditions. Most of the resonances of USrc in-cell appeared at very similar frequencies with respect to standard conditions (buffer), thus making their assignment within the cellular environment straightforward. A major difference was the appearance of a new peak (^{1}H-^{15}N chemical shift positions in ppm: 8.79, 117.20), which, on the basis of its chemical shift, corresponded to a phosphorylated serine or threonine[12]. The appearance of this peak matched the disappearance of the signal from Ser 17, thus indicating that this serine residue is spontaneously phosphorylated by endogenous kinases in *Xenopus* oocytes. Y. Perez[10] already observed a similar chemical shift perturbation in vitro in presence of PKA (Figure II.2.3.C), hence confirming that the phosphorylated residue is Ser 17 and suggesting that the endogenous kinase involved in the reaction is probably PKA. After NMR measurements ^{15}N-USrc was isolated from cells by Strep-tag affinity purification and analyzed by Western Blot to

rule out degradation phenomena within intact oocytes. Results are shown in Figure II.2.3.D. No degradation effect was observed.

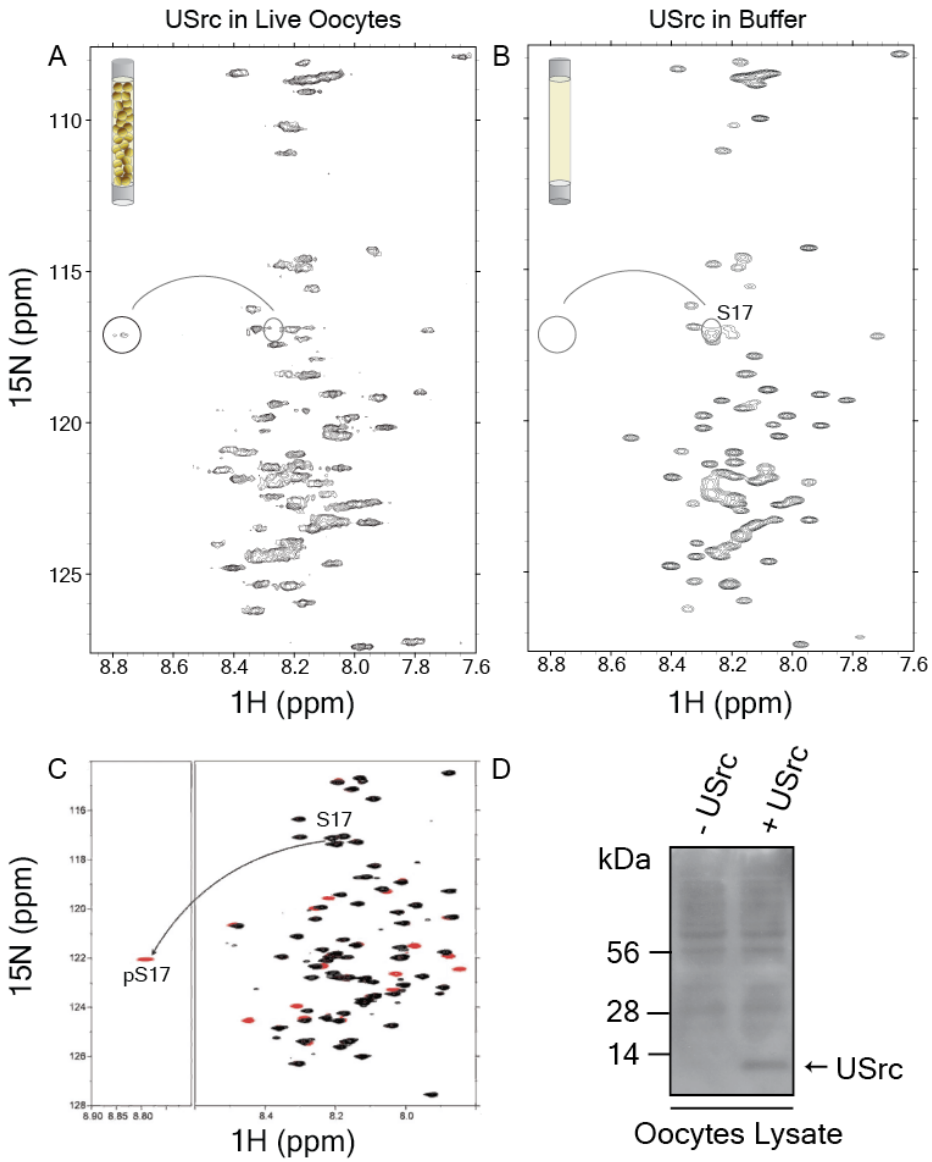


Figure II.2.3. USrc phosphorylation in Xenopus Laevis oocytes. ^{1}H-^{15}N correlation NMR spectra of USrc injected into intact Xenopus oocytes (A), and dissolved in buffer (B). Circles mark the position of phospho-Ser 17. The curved line indicates the shift to the left from unmodified (ovals) to phosphorylated Ser 17. (C) ^{1}H-^{15}N HSQC spectra of USrc of unphosphorylated USrc (black) and phospho-Ser 17 USrc (red) (Y. Perez results) (D) Western blots showing USrc in oocytes lysed after NMR measurements (lane 1, not injected; lane 2 injected with ^{15}N-labeled USrc). Protein was detected by using an antibody against the Strep tag.

We next investigated the intrinsically disordered domain of c-Src into oocyte extracts. ^{15}N-labeled USrc was added to the soluble cytoplasmic fraction obtained from lysed stage VI oocytes. Successively, NMR spectra were recorded at 288 K in a 600 MHz spectrometer. Phosphorylated Serine 17 was observed also in those conditions (Figure II.2.4.A). Chemical shift changes between ^{1}H-^{15}N HSQC spectra in buffer and in oocyte extracts are shown in Figure II.2.4.C. Although most residues displaying small changes were assigned by reference to the spectra obtained in buffer, resonances from residues Gln 13, Arg 14, Arg 15 (or Arg 16), Ser 17, Leu 18, Glu 19, and Ala 21 were significantly shifted or had disappeared in oocyte extracts and could not be assigned. These residues cluster around Ser 17 and included a number of charged residues that are probably affected by the presence of the additional charge of the phosphate group at Ser 17.

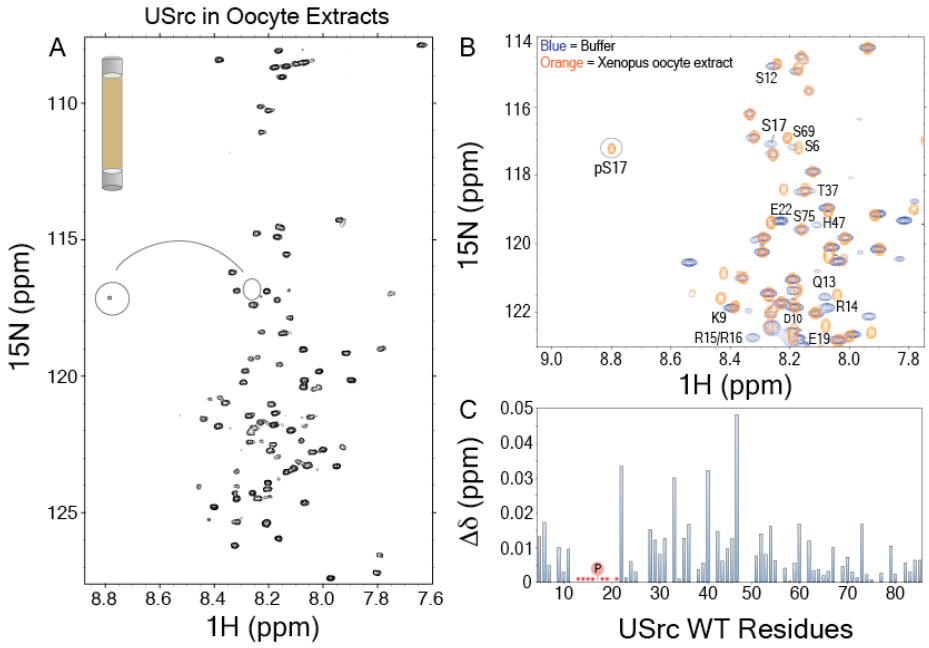


Figure II.2.4. USrc phosphorylation in lysed stage VI oocytes. (A) ^{1}H-^{15}N correlation NMR spectra of USrc in cytoplasmic extract from lysed oocytes. Circles mark the position of phospho-Ser 17. The curved line indicates the shift to the left from unmodified (ovals) to phosphorylated Ser 17. (B) Overlay of expansions of ^{1}H-^{15}N HSQC spectra of USrc in buffer (blue) and in Xenopus oocyte extract (orange). Circle marks the position of pSer 17. (C) Chemical shift changes between the two conditions (buffer vs. extract). Red asterisks indicate residues with large perturbation (preventing their unambiguous assignment in the phosphorylated sample). The phosphorylation site is indicated by a P.

Time-resolved multiple phosphorylations of USrc in Xenopus egg extracts

Extracts obtained from unfertilized *Xenopus Laevis* eggs may be referred to as cytostatic factor (CSF) extracts. Xenopus eggs are arrested in metaphase II of meiosis, and present a larger number of active kinases than do stage VI oocytes, including members of the cyclin dependent kinases (CDKs) family[13]. We aimed to investigate the

Unique domain of c-Src within the environment of Xenopus eggs. ^{15}N-labeled USrc WT was added to CSF extracts and studied by NMR spectroscopy. Figure II.2.5.A shows a comparison between ^{1}H-^{15}N HSQC correlation spectra of USrc in buffer and in Xenopus egg extract. After prolonged incubation of ^{15}N-labeled USrc in the CSF extract, three signals with chemical shifts typical of phosphorylated Ser/Thr were observed at the following ^{1}H-^{15}N chemical shift positions (in ppm): 8.59, 120.48 (peak 1), 8.84, 121.76 (peak 2), and 8.96, 117.72 (peak 3). However, the three new peaks appeared at different times and at different rates. The chemical-shift perturbations of USrc residues in the major species present after reaching the steady state (3 h after the reaction started) in the egg extract are shown in Figure II.2.5.B. Significant changes (higher than 0.02 ppm) were observed for residues: Ser 6, Glu 22, His 25, Ser 35, Gln 36, His 47, Asp 71, Thr 72 and Gly 80. However, almost of USrc residues suffered important perturbations in these conditions probably due to chemical rearrangements caused by phosphorylation events taking place within the domain.

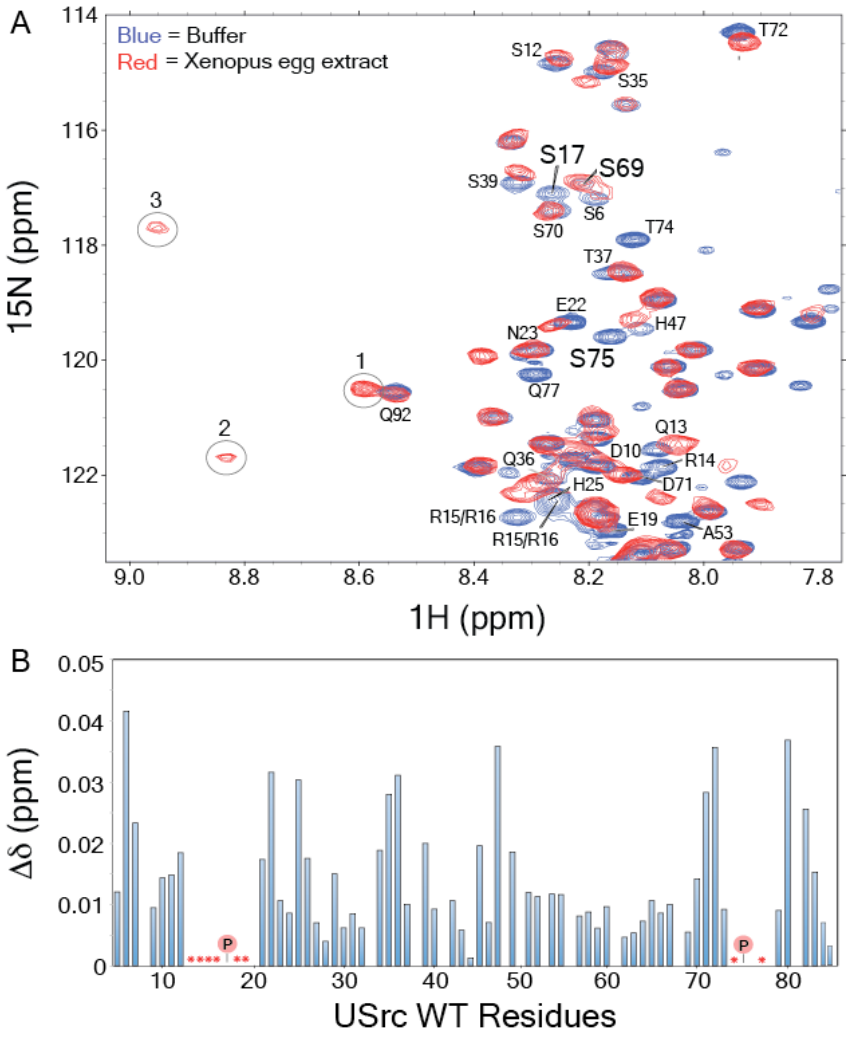


Figure II.2.5. USrc phosphorylation in Xenopus egg extract. (A) Overlay of expansions of ^{1}H-^{15}N HSQC spectra of USrc in buffer (blue) and in Xenopus egg extract (CSF extract; red). Circles mark the signals from phosphorylated Serines/Threonines. (B) Chemical shift changes between the major species in the two conditions (buffer vs. CSF extract). Red asterisks indicate residues with large perturbation (preventing their unambiguous assignment in the phosphorylated sample). P indicates phosphorylation sites unambiguously assigned by NMR.

In order to increase the time resolution in the observation of the multiple phosphorylation events of USrc, we optimized the experimental conditions by using a combination of SOFAST-HMQC experiments (to minimize the recycle time between successive NMR scans) and non-uniform sampling (to achieve good resolution in the indirect dimension in minimal time)[14,15,16]. Under optimized conditions (see Material and Methods), we recorded well-resolved 2D spectra with acceptable signal-to-noise ratio every 4 – 5 mins from 50 mM USrc samples in Xenopus Laevis egg extracts.

Peak 1 resonance appeared in the spectrum 10 – 15 min after addition of USrc to the CSF extract. Peaks 2 and 3 were first observed in the spectra 15 – 20 min after the reaction started. Figure II.3.6.A shows the variation of peak intensities over time, for the unmodified Ser 75 signal and for peak 1. Each intensity point was calculated as a running average of two experiments (5 min each). After adding USrc to the CSF extract, phosphorylation events caused a decrease in the intensity of the Ser 75 NH signal (this matched the increase in peak 1) until a plateau was reached after 2 h of reaction. The time evolution of the concentration of phosphorylated Ser/Thr residues of USrc in Xenopus egg extracts contributing to peak 2 and peak 3 intensities are shown in Figure II.3.6.B and II.3.6.C, respectively. The two peaks increased with different rates and more slowly than peak 1. In untreated extracts, the relative intensities of peaks 1, 2, and 3 were 1:0.3:0.5, after around 2 h of reaction.

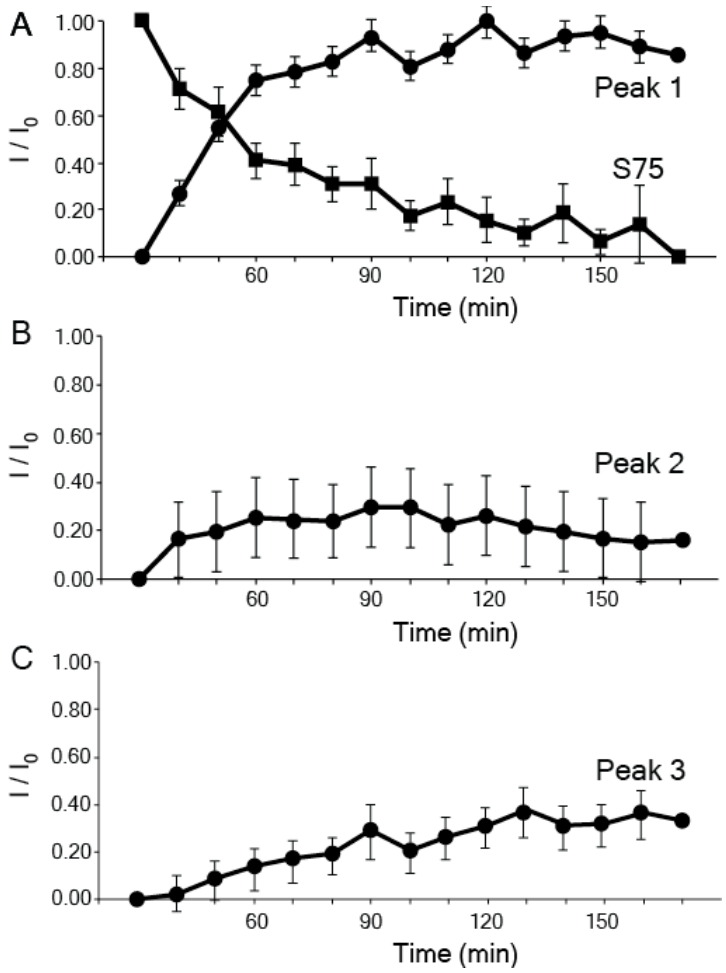


Figure II.2.6. Time evolution of USrc phosphorylation in Xenopus egg extract (CSF) followed by Real-time NMR spectroscopy. (A) Ser 75 and pSer 75 (peak 1) normalized with respect to the maximum intensity of peak 1; (B) peak 2, tentatively assigned to pSer 69; C) pSer 17 (peak 3). Intensities were measured as running averages of two SOFAST-HMQC experiments (5 min each). The intensities of peaks 2 and 3 are shown relative to the maximum intensity of peak 1. Error bars show standard deviations propagated from the variations of peaks unaffected by phosphorylation between experiments. Protein concentration was 50 µM.

Identification of USrc phosphorylation sites and active kinases in Xenopus egg extracts

In order to assign the USrc residues that are phosphorylated in the Xenopus egg extract we combined NMR, mass spectrometry (MS) and biochemical methods. ^{15}N-labeled USrc was recovered from CSF extract after NMR experiments. Phosphorylation of Ser 75 was found by mass-spectrometry upon detection of pSer75-specific tryptic and chymotryptic peptide MS/MS fragment ions, and by Western Blot using a specific anti phosphoS75-Src antibody (Figure II.2.7). Moreover, the previous NMR observation of the appearance of peak 1 parallel to the disappearance of the signal from unphosphorylated Ser 75 suggested that this new peak corresponds to phospho-Ser 75.

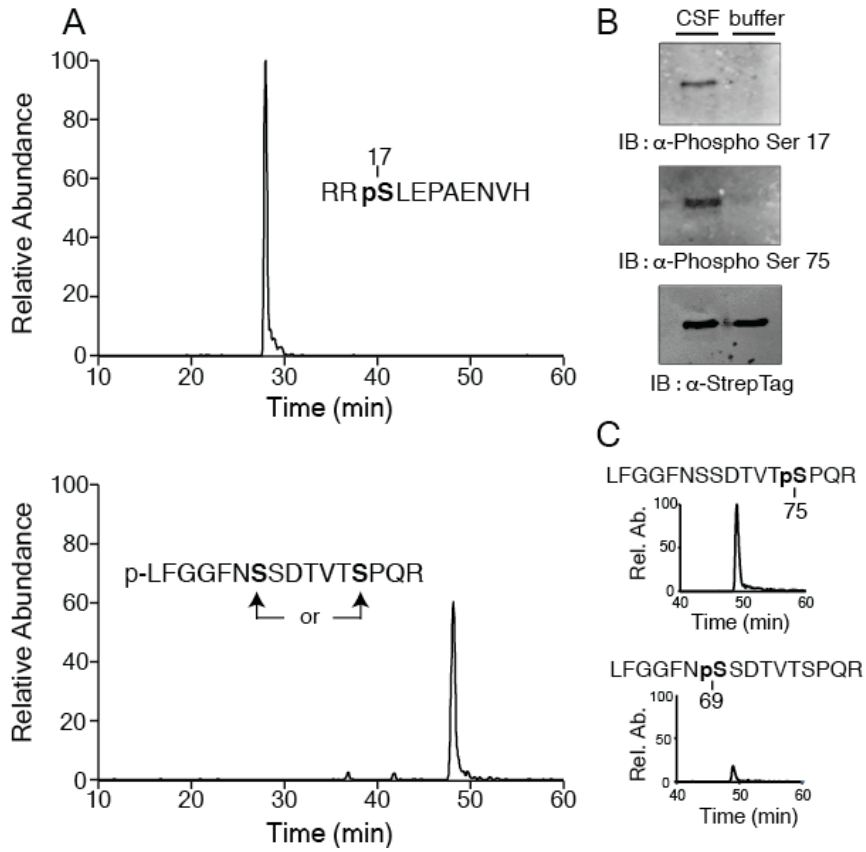


Figure II.2.7. MS and Western blot analysis of USrc in CSF extract. (A) Extracted ion chromatograms of m/z 463.2210 (calculated m/z 463.2208, Δm=0.43 ppm; top) and 896.8997 (calculated m/z 896.8987, Δm=1.11 ppm; bottom) corresponding to RRpSLEPAENVH $[M+3H]^{3+}$ and p-LFGGFNSSDTVTSPQR $[M+2H]^{2+}$ ions, respectively. (B) Western blots showing phosphorylation of Ser 17 and Ser 75 of USrc in CSF extracts. Protein loading was determined by using an antibody against the Strep tag. (C) The mono-phosphorylated peptide **p**-LFGGFNSSDTVTSPQR is a mixture of LFGGFNSSDTVTpSPQR and LFGGFNpSSDTVTSPQR; these can be distinguished by their MS/MS transitions: (896.89 → 567.25, y4) and (896.89 → 972.48, y9-H2O), respectively.

In vitro incubation of USrc with Cdk5/p25 resulted in the phosphorylation of both Ser 75 and Thr 37[10] suggesting that CDKs could be possible candidates for phosphorylating USrc *in vivo.* To test this hypothesis, Xenopus egg extracts were treated with Roscovitine, an inhibitor of several CDKs including Cdk5 and successively mixed with ^{15}N-labeled USrc. Roscovitine treatment prevented the appearance of peaks 1 and 2 confirming that the corresponding phosphorylation events were catalysed by CDKs (Figure II.2.8.A). We then tested a second inhibitor called Ro3306 (selective for Cdk1). CSF extract treatment with Ro3306 had no effect on the phosphorylation of USrc detected by NMR (Figure II.2.8.B). Western Blot analysis confirmed that phosphorylation of USrc on Ser 75 was partially inhibited by adding Roscovitine but not Ro3306 to Xenopus egg extracts (Figure II.2.8.C-D).

Peak 2 was not possible to assign by NMR. To uncover which phosphorylated residue corresponded to this cross peak ^{15}N-labeled USrc was purified from Xenopus egg extract after NMR experiments and analysed by mass-spectrometry. Phosphorylated Ser 69 was unequivocally detected, although at lower abundance than pSer 75 (Figure II.2.7.A and Figure S3).

Peak 3 corresponds to phosphorylated Serine 17 (pSer 17), which was phosphorylated by PKA in oocytes. The chemical shifts of phosphorylated residues are very sensitive to pH and the chemical shifts of phosphorylated residues observed in oocytes or oocyte extracts and in egg-extracts were not identical. However, the assignment was confirmed by the observation of a rapid increase in the intensity of peak 3 after the addition of PKA to the egg extracts (Figure II.2.9). The presence of exogenous PKA within the CSF extract resulted, in addition to complete phosphorylation of Ser 17, in an increase in the level of pSer 75. Interestingly, addition of PKA prevents the appearance of peak 2 (pSer 69). After recovering USrc from the CSF extract, the phosphorylation of residue Ser 17 was confirmed by Western Blot, using a specific phospho-antibody (Figure II.2.7.B). Phosphorylation at Ser 17 was further confirmed by mass spectrometry (Figure II.2.7.A).

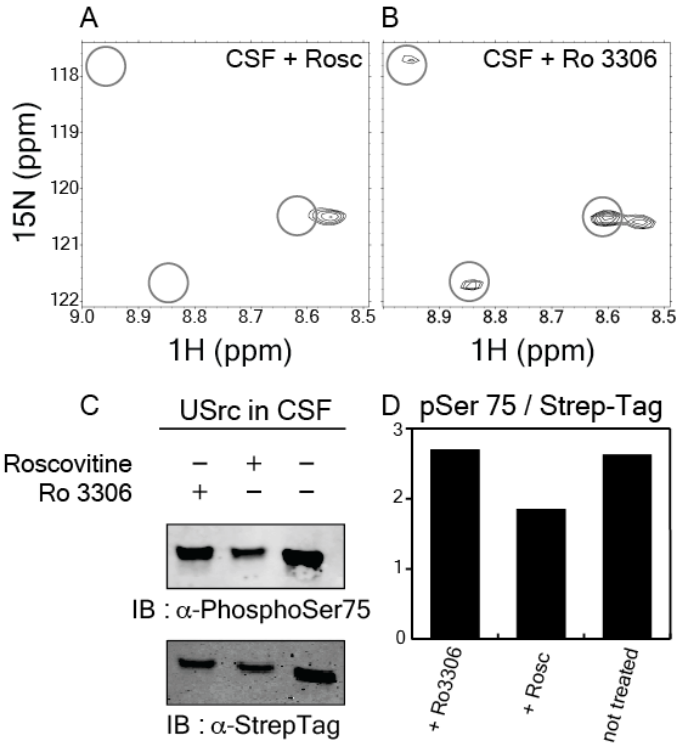


Figure II.2.8. Inhibiting phosphorylation of USrc in CSF extract. Expansions of ^{1}H-^{15}N HSQC spectra of USrc in Xenopus egg extracts (CSF) treated with CDKs inhibitors Roscovitine (A) and Ro 3306 (B). (C) Effects of CDKs inhibitors on Ser 75 phosphorylation. USrc was isolated after being incubated with CSF extracts with added Roscovitine (0.5 mM) and Ro 3306 (25 mM). Membranes were immunoblotted with anti-Phospho Ser 75 (top panels) and anti-StrepTag (bottom panels). (D) Quantification of Western Blots represented in (C).

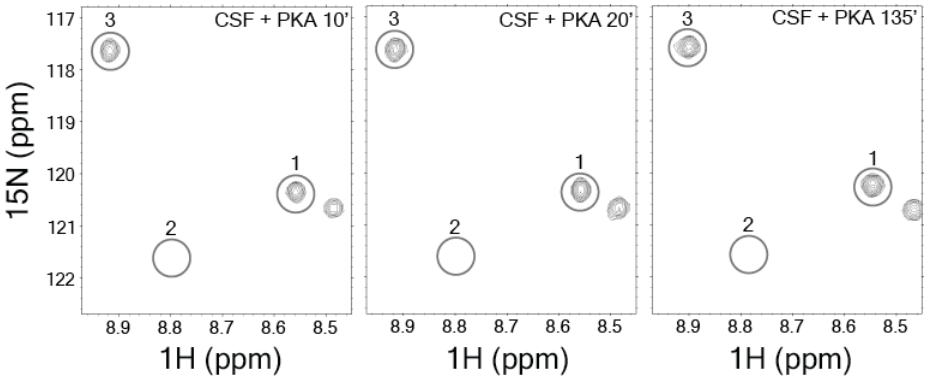


Figure II.2.9. Addition of PKA to CSF extract. Expansions of ^{1}H-^{15}N HSQC spectra of USrc in Xenopus egg extracts (CSF) incubated with an excess of PKA. Peak 1 and peak 3 appear within the first 20 minutes, while peak 2 did not appear even after 135 minutes of reaction.

PKA is not the major kinase phosphorylating Ser 17 of USrc in Xenopus egg extracts

PKA was shown to phosphorylate Ser 17 both *in vitro* and *in vivo*. To test whether PKA was the only kinase able to phosphorylate USrc on residue 17, we treated Xenopus egg extract with a PKA selective inhibitor called PKA-I 6-22 (residues 6-22 of the PKA inhibitory protein PKI). ^{15}N-labeled USrc was then added to CSF treated extract and measured by NMR in a 600 MHz spectrometer. The addition of the PKA inhibitor PKA-I 6-22 to the egg extracts did not impede the

appearance of pSer 17 (Figure II.2.10.A). This result suggested that an additional kinase present in the extract was able to catalyze Ser 17 phosphorylation of USrc. To confirm this observation, CSF extract were successively treated with a broader specificity inhibitor called H89 which inhibits in addition to PKA other kinases such as S6K1, MSK1, ROCKII, PKBα and p90RSK. The addition of H89 to the Xenopus egg extracts prevented phosphorylation of Ser 17 (Figure II.2.10.B), indicating that the kinase responsible for the phosphorylation of this residue is sensitive to H89 in our model system.

Interestingly, in the presence of PKA inhibitors PKA-I 6-22 and H89 the NMR signals corresponding to unmodified Ser 75 was clearly observed and peaks 1 and 2 corresponding to phosphorylated species were not present (Figure II.2.10.A-B). These results suggest that inhibition of PKA might trigger the activity of a phosphatase that reverses or strongly reduces the effect of CDKs keeping Ser 75 and Ser 69 unphosphorylated. In order to test this hypothesis, H89 was added after a 30 minutes preincubation of USrc with the CSF extract to allow the phosphorylation of USrc by CDKs (Figure II.2.10 cartoon). NMR measurements showed a decrease in the intensity of the CDK phosphorylated species following the addition of H89 (Figure II.2.10.C-D). The amount of phosphorylated Ser 17 remained constant showing that the activated phosphatase is not affecting pSer 17. A strong decrease in the phosphorylation of Ser 75 in *Xenopus* egg extracts caused by the addition of PKA-I 6-22 or H89 was also observed using a specific pSer75-Src antibody (Figure II.2.10.E).

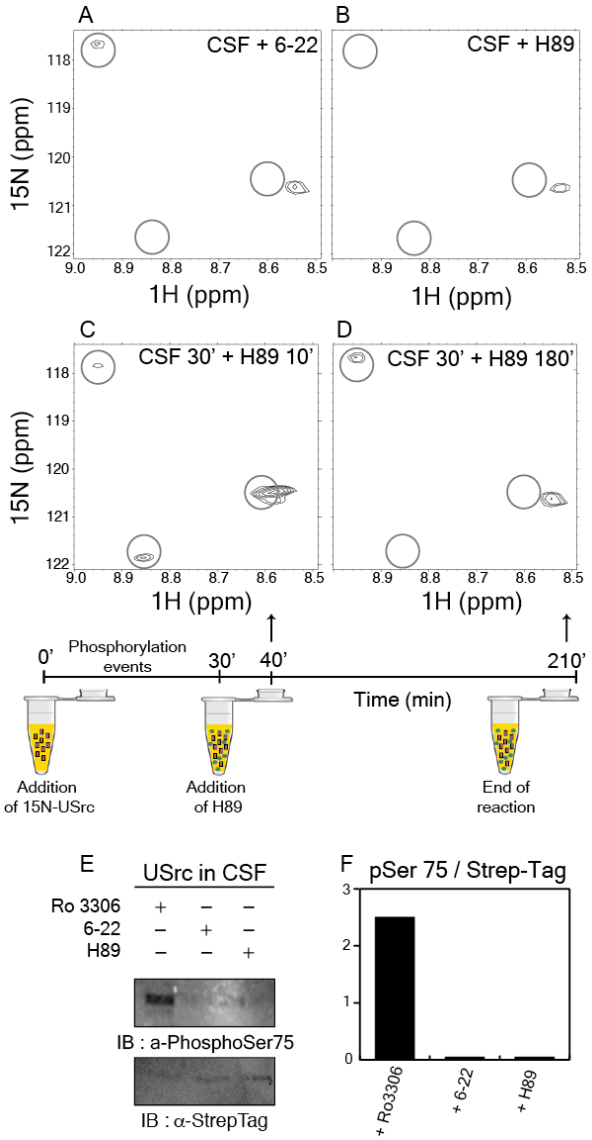


Figure II.2.10. Inhibiting phosphorylation in CSF extract. Expansions of ^{1}H-^{15}N HSQC spectra of USrc in Xenopus egg extracts (CSF) treated with PKA inhibitors 6-22 fragment (A) and H89 (B). Panels C and D compare expansions of ^{1}H-^{15}N HSQC spectra of USrc at 10 and 180 minutes after the addition of H89 following preincubation in CSF for 30 minutes (see cartoon below). (E) Effects of CDKs/PKA inhibitors on Ser 75 phosphorylation. USrc was isolated after being incubated with CSF extracts with added Ro 3306 (25 mM); 6-22 (50 mM) and H-89 (3 mM). Membranes were immunoblotted with anti-Phospho Ser 75 (top panels) and anti-StrepTag (bottom panels). (F) Quantification of Western Blots represented in (E).

Phosphorylation of Ser 69

The NMR observed peak 2 was unequivocally assigned to Ser 69. We had shown that addition of exogenous PKA to CSF extracts prevents the appearance of peak 2 (Figure II.2.9). This effect could indicate that Ser 75 and Ser 69 are targets of a different set of kinases and phosphatases having different dependencies on PKA or may reflect that phosphorylation of Ser 69 cannot occur after Ser 75 is phosphorylated. To test this hypothesis a single point mutant was generated by replacing Ser 69 with the negatively charged Glutamic acid (USrc S69E). This mutation was inserted in order to mimic the phosphate group present on the phosphorylated Ser 69. Xenopus egg extracts were incubated over-night at 25°C with USrc WT and USrc S69E. The day after, proteins were purified from CSF extracts and analyzed by Western blots and mass-spectrometry. Results are shown in Figure II.2.11 and Table S1. Phosphorylation of Ser 75 was detected with both technics even in presence of the mutated Ser 69.

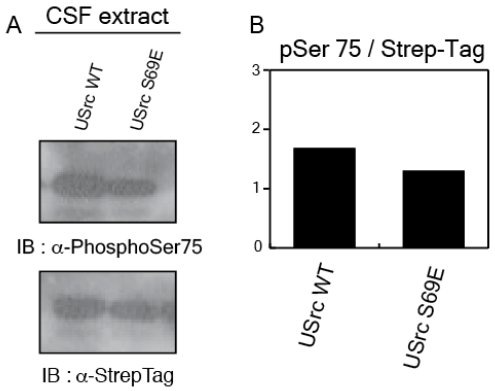


Figure II.2.11. Phosphorylation of USrc in CSF extract. (A) USrc WT and S69E was isolated after being incubated O/N with Xenopus egg extracts. Membranes were immunoblotted with anti-Phospho Ser 75 (top panels) and anti-StrepTag (bottom panels). (B) Quantification of Western Blots represented in (A).

In vivo effects of mutations in the Lipid Binding Region of the Unique domain

The ULBR and phosphorylation sites shown to be relevant for the Unique domain of human c-Src are highly conserved even in phylogenetically distant species (Figure II.2.1). *Xenopus laevis* stage VI oocytes were used as a model system to test the effect of mutations of the ULBR *in vivo*. USrc mutants AAA and EAE, previously shown to lose the capability to bind lipids *in vitro*, were generated by site-direct mutagenesis in the context of the full-length c-Src.

Progesterone induced maturation of Xenopus oocytes has been previously described to be accelerated by the expression of constitutively active viral or *Xenopus* Src[17]. Oocytes were injected with *in vitro* transcribed mRNAs encoding constitutively active human c-Src protein (Y530F), with either wild-type, AAA, or EAE Unique domains. Figure II.2.12.A shows the percentage of oocytes that completed maturation at different times after the addition of progesterone. Maturation was assessed by the appearance of a white spot at the animal pole of the oocytes, which indicates germinal vesicle breakdown (GVBD) and meiosis I entry. Oocytes injected with mRNA encoding wild-type c-Src started maturation around 2 h before control oocytes, either untreated or injected with H_2O. This observation agrees with previous reports[18] and confirms that Xenopus oocytes are an appropriate model to functionally characterize human c-Src. Interestingly, oocytes injected with mRNAs encoding the AAA or EAE Unique domain mutants also showed accelerated maturation in response to progesterone but only around 70%-80% completed the process (Figure II.2.12, inset). In addition, 2 h after the appearance of the white spot, about half of the matured oocytes that expressed AAA or EAE mutants showed progressive depigmentation and started to die (Figure II.2.12.B). These effects were not a consequence of different expression levels of the wild type and mutant forms of c-Src as all

three proteins were expressed at similar levels in the injected oocytes (Figure II.2.12.C).

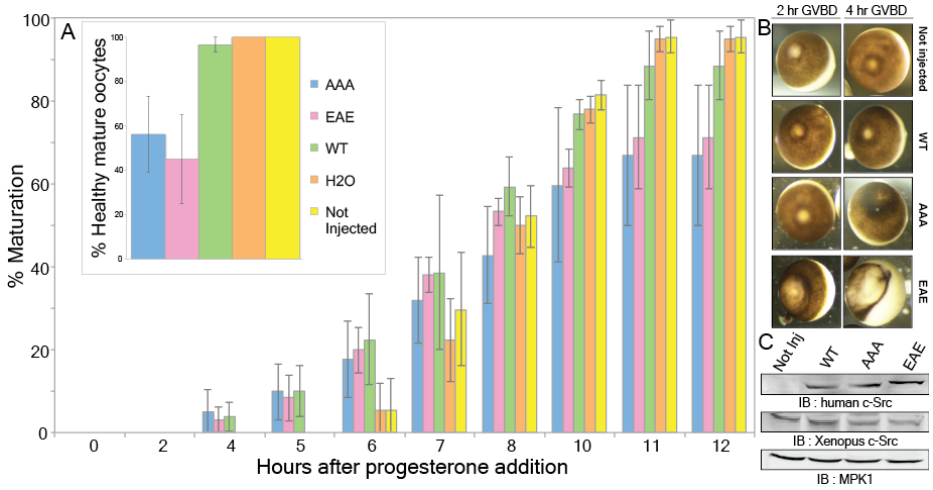


Figure II.2.12. Effect of c-Src mutants on Xenopus oocyte maturation. The effect of the injection of oocytes with mRNAs encoding wild-type c-Src (green), AAA mutant (blue), EAE mutant (magenta), or water (orange) is compared with that of non-injected oocytes (yellow). (A) Percentage of oocytes that underwent maturation, as determined germinal vesicle breakdown (GVBD), at different times after the addition of progesterone. The inset shows the percentages of morphologically normal oocytes at the end of the experiment, when 100% of the control oocytes treated with progesterone reached GVBD. (B) Oocyte appearance 2 and 4 h after GVBD. About half of the matured oocytes that express AAA or EAE mutants show progressive depigmentation starting 2 h after GVBD. (C) Analysis of the expression levels of wild-type and mutants forms of human c-Src and endogenous Xenopus c-Src. Oocyte lysates were separated by SDS-PAGE and analyzed by immunoblotting with anti-human c-Src (top panel), anti-Xenopus c-Src (middle panel) and anti-MPK1 (bottom panel) as a loading control.

DISCUSSION

In-cell NMR and real-time NMR spectroscopy allowed the study of complex phosphorylation/dephosphorylation processes in the intrinsically disordered Unique domain of human c-Src mediated by kinases and phosphatases that are present in *Xenopus Laevis* oocytes and cell extracts. We first observed that USrc was phosphorylated on Ser 17 in Xenopus living oocytes. Similar results were observed when ^{15}N-labeled USrc was added to extracts containing the soluble fraction of lysed stage VI oocytes, indicating that the phosphorylation involved a kinase that is soluble or at least associated to the soluble lipid fraction. Protein kinase A (PKA) was identified as the major candidate to catalyse the reaction. Indeed, Ser 17 can be phosphorylated *in vitro* by purified PKA[10], and this kinase is known to be active in *Xenopus* oocytes[19]. Most of the USrc resonances surrounding Ser 17 were significantly perturbed in the oocyte extracts (Figure II.2.4.C). Residue Glu 22, another charged residue close to Ser 17, also showed chemical shift changes above the standard deviation (0.13 ppm) of the variations observed between buffer and oocyte extracts. Likewise, the charge-sensitive residue His 47 and its neighbor Gly 46, although located far from Ser 17 were also significantly perturbed in the oocyte extracts. Interestingly, hydrophobic residues Ala 27 and Phe 32 in the intermediate region were also affected. USrc phosphorylated at Ser 17 showed only very small perturbations around His 47 in buffer[10], thereby suggesting that the perturbations observed in the oocyte extract are mediated by the interaction with some components of the extract. Residues Arg 48, Asn 68 and Arg 78 could not be distinguished in the cell extract due to spectral overlap.

We successively studied USrc in the unfertilized *Xenopus Laevis* eggs (CSF extract) system in which we were able to identify three different phosphorylation sites that were unequivocally assigned to Ser 17, Ser 69 and Ser 75. The major candidates able to phosphorylate USrc on

residues Ser 69 and Ser 75 were CDKs. USrc was previously shown to be phosphorylated *in vitro* by Cdk5/p25 on Ser 75 and Thr 37[10]. However, phosphorylated Thr 37 could not be detected by mass-spectrometry after incubation of USrc with *Xenopus* egg extracts. In contrast, mass-spectrometry unequivocally detected phosphorylated Ser 69. Phosphorylation of residue Ser 69 of c-Src had previously been detected only by MS in cell extracts from cancer lines HCT116 and MDA-MB-435S[20], although its function is presently unknown. Ser 69 and Ser 75 are present in the same tryptic peptide. Nevertheless, MS detected only mono-phosphorylated forms (Table S2). The appearance of the peak assigned to pSer 69 does not result in an observable decrease of the signal from the unmodified residue. This apparent contradiction could be accounted for by the observation of a general increase in the intensities caused by the phosphorylation of Ser 75, masking the small decrease expected by the (minor) phosphorylation of Ser 69 (Figure S4). In fact, in the major species, the chemical shift of Ser 69 neighbour residues 66, 67 and 70 were not affected (Figure II.2.5.B) suggesting that this form is phosphorylated at Ser 75 but not at Ser 69. We interpreted the intensity increase of signals from the major species as a result of the elimination of residual interactions with the 60-75 region caused by phosphorylation of Ser 75, as previously observed[9].

Addition of CDKs or PKA kinase inhibitors or purified PKA clearly showed the cross talk between kinases and phosphatases in Xenopus egg extracts. Thus, by changing the experimental conditions of the cellular environment, we have been able to uncover a very complex scenario of multi-phosphosphorylation events of USrc driven by endogenous kinases. Roscovitine (CDK inhibitor) affects the phosphorylation of Ser 75, Ser 69 as well of Ser 17. On the other hand, CSF treatment with the PKA inhibitor H89 had a striking effect not only on Ser 17 but also on the phosphorylation state of USrc residues

targeted by CDKs (Ser 69 and Ser 75). Inhibition of PKA (and other kinases) probably induces the activation of a phosphatase that controls CDKs. A candidate to mediate the effect of H89 is inhibitor-1 (I-1) of Protein Phosphatase 1 (PP1). This ubiquitously distributed intrinsically disordered protein is phosphorylated by PKA and become a strong inhibitor of PP1[21]. Inhibition of PKA would allow the dephosphorylation of I-1 and release of active PP1, leading to the observed absence of phosphorylation of Ser 75.

The excess of exogenous PKA prevents the appearance of peak 2 (pSer 69). We have demonstrated that even if MS detected only mono-phosphorylated fragments, Ser 75 can be still phosphorylated in the presence of a negative charge (S69E) mimicking the phosphate group of the pSer 69. Thus, this result suggests that Ser 75 and Ser 69 are probably targets of a different set of kinases and phosphatases having different dependencies on PKA.

We had shown that the Unique domain contains a secondary lipid-binding region, in addition to the SH4 lipid-anchoring moiety. The observation of well-resolved spectra in the intact oocyte as well in the different extracts suggests that interactions involving this secondary site are weak, at least in the absence of myristoylation.

We have finally demonstrated the biological relevance of the interactions involving the Unique Lipid Binding Region in the context of full length c-Src, by comparing the effects of proteins with wild type or mutated ULBR sequences in the maturation of *Xenopus Laevis* oocytes, used as a model system. Injection of mRNA encoding a constitutively active form of human c-Src caused an acceleration of progesterone-induced maturation of oocytes, which is very similar to that induced by constitutively active Xenopus c-Src or v-Src[17,18]. This result validates the heterologous model. However, injection of in vitro transcribed mRNAs encoding for mutants in the ULBR caused that a significant portion (20%–25%) of the oocytes failed to mature after

treatment with progesterone. Furthermore, about half of the oocytes that matured, showed depigmentation and possible apoptotic symptoms. Thus, while wild-type and mutant forms of human c-Src are similarly active in the initiation of oocyte maturation, the failure to complete maturation and the lethal phenotype observed only in matured oocytes expressing mutated c-Src demonstrates that the ULBR plays an essential role in the regulation of c-Src. These results constitute a proof of concept of the functional importance of the ULBR and support the idea that the observed interactions are indeed part of a new level of regulation for c-Src.

The characterized phosphorylation events, that were already shown to control lipid binding *in vitro* (Figure II.2.2), could represent an important mechanism of regulating the ULBR also *in vivo*.

MATERIAL AND METHODS

Cloning, Protein Expression and Purification

The cDNA encoding (1-85) human c-Src region with a Strep-tag in C-terminal position for purification purposes was cloned into a pET-14b vector (Novagen, UK) (Figure A.3.1). The glycine residue (G2) following the initial methionine was mutated to alanine to prevent myristoylation after injection into the oocytes. The plasmid was transformed in *Escherichia coli* Rosetta^{TM} (DE3) pLysS cells (Novagen, UK) and cells were grown in M9 minimal medium supplemented with [^{15}N]H_4Cl (Cambridge Isotope Laboratories, UK). USrc protein was isolated using Strep-tactin Sepharose (IBA, Göttingen). After elution with 2.5 mM of desthiobiotin, it was further purified by size exclusion chromatography (Superdex 75 26/60, GE Healthcare, Spain) in phosphate buffer (50 mM phosphate, 0.2 mM EDTA, pH 7.0). Mutations were introduced using the QuikChange site-directed mutagenesis kit (Stratagene). Further details are in II.1 – Material and Methods.

Xenopus Oocytes and Extracts

Xenopus laevis ovaries were surgically removed from full-grown females and treated with collagenase and dispase[22]. Stage VI oocytes were selected and maintained in Barth's medium. Oocytes were microinjected with 50 nl of purified ^{15}N-labeled USrc protein (1 mM) and were left recovering for 3 hrs in modified Barth's saline (MBS) at 18°C before NMR measurements.

For the preparation of lysates, oocytes were homogenized in 10 µl per oocyte of ice-cold H1K buffer (80 mM β-glycerophosphate, pH 7.5, 20 mM EGTA, 15 mM $MgCl_2$, 1 mM DTT, 1 mM AEBSF, 2.5 mM Benzamidine, and 10 µg/ml each of Aprotinin and Leupeptin). Lysates

were centrifuged at 10,000 g for 10 min at 4 ºC, and the cleared supernatants were used for NMR measurements. Cytostatic factor (CSF) extracts were prepared with unfertilized *Xenopus Laevis* eggs. An energy mix solution (150 mM creatine phosphate, 20 mM ATP, 2 mM EGTA pH 7.7, 20 mM $MgCl_2$) was added before storage. Protocols are detailed in[22,23]. Immediately after preparation, CSF extracts were stored in aliquots of 100 or 200 µl at -80 ºC.

NMR sample preparation

Cell-extract samples for NMR experiments were prepared by mixing 25 µL of 0.5 mM stock solution of ^{15}N-labeled USrc protein (in 50 mM phosphate buffer, 0.2 mM EDTA, pH 7.0) with 25 µL of D_2O and 200 µl of CSF or oocyte extracts. Prior to NMR measurements CSF extracts were thawed in ice and quickly span down. In order to minimize time for temperature equilibration in the spectrometer, D_2O and ^{15}N-labeled protein stock solutions were kept between 16 °C and 18 °C before being added to the extract. Immediately after adding the protein to the extract, the samples (250 µl, final volume) were transferred to Shigemi tubes and inserted in the spectrometer for measurements.

The in-cell NMR samples were prepared as described above and details are in[24]. After recovering (see above), oocytes were first washed twice with MBS and, 30 minutes before NMR measurements, transferred to MBS containing 10% D_2O. 50 nL of a 1 mM stock solution of ^{15}N-labeled USrc protein were microinjected in each oocyte to provide an in-cell concentration of around 50 µM, assuming that the intracellular volume is around 1 mL. Two hundred and fifty oocytes were used for individual in-cell NMR samples.

The following products were used for the different in vivo assays: Protein Kinase A (P2645-400UN, Sigma), H89 dihydrochloride (#2910, Tocris Bioscience), Protein Kinase A Inhibitor Fragment 6-22 amide

(P6062, Sigma), Ro 3306 (#4181, Tocris Bioscience), and Roscovitine (#557360, Calbiochem). Prior to the addition of the ^{15}N labelled USrc, the kinases inhibitors were incubated 30 min with the CSF extract.

Western Blotting

USrc was recovered from CSF extracts by Strep-tag affinity purification after NMR measurements. 0.5 µg of purified protein was separated by SDS-PAGE and transferred onto Hybond ECL nitrocellulose membrane (GE Healthcare). Membranes were blocked with 5% Milk in TBS-Tween and immunoblotted with the following antibodies: 1:1,000 anti-PhosphoSrc-Ser17 (#5473, Cell Signaling Technology), 1:1,000 anti-PhosphoSrc-Ser75 (#8A8186, AAT Bioquest) and 1:20,000 anti-Streptag antibody (IBA, Göttingen).

Mass Spectrometry

20 mg of ^{15}N-labeled USrc recovered from NMR experiments, or unlabeled USrc treated in identical way, was digested by adding trypsin or chymotrypsin (2% w/v) and incubated at 37°C overnight. Digestions were stopped by adding formic acid to a final concentration of 1%. The resulting peptide mixtures were diluted in 1% FA and loaded in a nano-LC-MS/MS system connected to an Advion TriVersa NanoMate (Advion) fitted on an LTQ-FT Ultra mass spectrometer (Thermo Scientific). Further experimental details are given as supplementary information (see Addendum – A.4).

A database search was performed with Bioworks v3.1.1 SP1 and Proteome Discoverer software v1.3 (Thermo Scientific) by using the Sequest search engine and a home-made database in SwissProt format, which included USrc protein and the common repository of adventitious proteins (http://www.thegpm.org/crap/index.html). Search

parameters included no-enzyme specificity, methionine oxidation and phosphorylation in serine and threonine as dynamic modification and, depending on the sample, amino acids labeled with ^{15}N as static modification.

Extracted ion chromatograms of MS or MS/MS ions were obtained using Xcalibur software vs 2.0SR2 (Thermo Scientific).

NMR Spectroscopy

NMR experiments of ^{15}N labeled USrc in 50 mM phosphate buffer pH 7.0 were recorded at 288 K with a 600 MHz Bruker Avance III spectrometer equipped with a TCI cryo-probe at 0.5 mM protein concentration. NMR experiments in live oocytes as well as in the cellular extracts were performed at 288K, which is the optimal temperature for live oocytes, and 50 µM concentration.

Combined NH chemical shift differences were computed as in equation (1):

$$\Delta\delta = [\Delta\delta_H^2 + (\Delta\delta_N / 5)^2]^{1/2}$$

where d_H and d_N are the changes in chemical shift for ^1H and ^{15}N, respectively. In vitro kinase reactions were reconstituted with 50 µM of ^{15}N-labeled protein and recombinant PKA, 400 enzymatic units, in the presence of ATP (0.5 mM) and Mg_2SO_4 (0.1 mM).

In buffer and in-cell NMR experiments were performed in Shigemi tubes. For experiments with cell extracts Shigemi or 3mm NMR tubes were used. Probe tuning and shimming was performed in a dummy sample equilibrated at 288K immediately before the insertion of the real sample to minimize dead time in the real time measurements setting. The time from protein addition to the start of NMR data

acquisition was 5 min. For quantitative time-resolved experiments, we acquired consecutive ^{1}H-^{15}N SOFAST-HMQC[14] correlation spectra while the kinase reaction proceeded. The data were recorded with 4 or 8 transients and 800 complex points for ^{1}H and 256 or 512 complex points for the ^{15}N dimension, using traditional or non-uniform sampling schemes. Spectral width in the ^{15}N dimension was minimized in order to improve resolution in the indirect dimension. NMR data were processed with NMRPipe[25] and spectra were analyzed with Sparky[26].

In vivo assays with Xenopus oocytes

Xenopus Laevis ovaries were surgically removed from full-grown females and treated with collagenase and dispase. Stage VI oocytes were then selected and cultured in Barth's medium (see above).

Wild type human c-Src cDNA or mutated variants in the Unique domain were cloned in the vector FTX6 (Figure A.3.3). Constitutively active forms were obtained replacing tyrosine 530 by phenylalanine. The different constructs were then used to prepare mRNAs by in vitro transcription with the mMessage mMachine kit (Ambion). Oocytes were microinjected with 50 nl of mRNAs (250 ng) and maintained overnight in modified Barth's medium at 18°C before progesterone stimulation (5 mg/ml, Sigma) to induce maturation. For the preparation of lysates, oocytes were homogenized in 10 ml per oocyte of ice-cold H1K buffer (80 mM b-glycerophosphate, pH 7.5, 20 mM EGTA, 15 mM MgCl2, 1 mM DTT, 1 mM AEBSF, 2.5 mM Benzamidine, and 10 mg/ml each of Aprotinin and Leupeptin). Lysates were centrifuged at 10,000 g for 10 min, and the cleared supernatants were used for western blotting. Protocols are detailed in[22]. The following antibodies were used for western blotting: anti-Src [clone 327] (ab16885, Abcam), anti c-Src [SRC 2] (sc-18, Santa Cruz biotechnology), anti-MPK1 (rabbit serum).

REFERENCES

1. Selenko, P. *et al.* In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **15,** 321–9 (2008).

2. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337,** 635–45 (2004).

3. Oldfield, C. J. *et al.* Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **44,** 1989–2000 (2005).

4. Dunker, a K., Silman, I., Uversky, V. N. & Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18,** 756–64 (2008).

5. Uversky, V. N. *et al.* Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* **10 Suppl 1,** S7 (2009).

6. Stein, A., Pache, R. a, Bernadó, P., Pons, M. & Aloy, P. Dynamic interactions of proteins in complex networks: a more structured view. *FEBS J.* **276,** 5390–405 (2009).

7. Obara, Y., Labudda, K., Dillon, T. J. & Stork, P. J. S. PKA phosphorylation of Src mediates Rap1 activation in NGF and cAMP signaling in PC12 cells. *J. Cell Sci.* **117,** 6085–94 (2004).

8. Pan, Q. *et al.* Cdk5 targets active Src for ubiquitin-dependent degradation by phosphorylating Src(S75). *Cell. Mol. Life Sci.* **68,** 3425–36 (2011).

9. Pérez, Y. *et al.* Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* **3,** 1295 (2013).

10. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391,** 136–48 (2009).

11. Selenko, P. & Wagner, G. Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* **158,** 244–53 (2007).

12. Theillet, F.-X. *et al.* Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* 217–236 (2012). doi:10.1007/s10858-012-9674-x

13. J. E. Ferrell Jr. Xenopus oocyte maturation: new lessons from a good egg. *BioEssays*, 21, 833- 842 (1999).

14. Schanda, P., Kupce, E. & Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **33,** 199-211 (2005).

15. Kazimierczuk, K. & Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chemie* **123,** 5670-5673 (2011).

16. Orekhov, V. Y. & Jaravine, V. a. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **59,** 271-92 (2011).

17. Unger, T. F. & Steele, R. E. Biochemical and Cytological changes associated with Expression of deregulated pp60src In Xenopus Oocytes. *Mol. Cell. Biol.* 12, 5485-5498 (1992).

18. Tokmakov, A. et al. Regulation of Src kinase activity during Xenopus oocyte maturation. *Dev. Biol.* 278, 289-300 (2005).

19. Schmitt, A. & Nebreda, A. R. Signalling pathways in oocyte meiotic maturation. *J. Cell Sci.* **115,** 2457-9 (2002).

20. Oppermann, F. S. *et al.* Large-scale proteomics analysis of the human kinome. *Mol. Cell. Proteomics* **8,** 1751-64 (2009).

21. Nimmo, G.A. & Cohen, P. The regulation of glycogen metabolism. Purification and characterisation of protein phosphatase inhibitor-1 from rabbit skeletal muscle. *Eur. J. Biochem.* 87, 341−351 (1978)

22. Perdiguero, E. & Nebreda, A. R. Use of Xenopus oocytes and early embryos to study MAPK signaling. *Methods Mol Biol* 250, 299-314 (2004).

23. Serber, Z. *et al.* Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. *Nat. Protoc.* **1,** 2701-9 (2006).

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

25. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **3**, 277-93. (1995)

26. T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco

II.3 - Functional studies of the Unique domain in the context of the full-length protein

INTRODUCTION

c-Src is the leading member of the Src family of non-receptor tyrosine kinase (SFKs). It plays important roles in cell proliferation, survival, migration and cell adhesion[1]. It has been shown that c-Src displays oncogenic activity when deregulated, a situation originally found with v-Src, the transforming product of the avian retrovirus Rous sarcoma virus[2]. c-Src was found to be over-expressed and up-regulated in several tumor types, including colorectal cancer (CRC) and prostate cancer[2] and has been associated to survival of bone metastasis breast cancer cells[3].

Remarkably, elevated kinase activity has been found in >80% of colon carcinoma (CRC) compared with the normal counterpart and this has been associated with a poor clinical outcome[4]. It has not be clarified yet how c-Src is deregulated in this type of cancer. An activating mutation has been reported in a minority of advanced tumors, suggesting the existence of alternative mechanisms for kinase deregulation[5]. Indeed, in contrast to viral oncoproteins, activating mutations of c-Src are rarely observed in human cancers and its over-expression is not always sufficient to induce cell transformation. Therefore, the oncogenic potential of c-Src is probably related to failures in the network of interactions that regulates their activity in normal cells.

It has been postulated that c-Src activation additionally implicates alteration of important Src regulators by genetic or epigenetic mechanisms, including the inhibitory COOH-terminal Src tyrosine kinase Csk[2].

c-Src deregulation is an important event for colon tumorigenesis and metastasis[4]. For instance, c-Src regulates growth, survival, and invasion of some CRC cell-lines *in vitro*[4,6]. Moreover, it contributes to tumor growth, angiogenesis, and metastasis in xenografts nude mouse models[4]. Therefore, c-Src has become an attractive therapeutic target in this type of cancer and several small inhibitors are under clinical trials. c-Src substrates important for cellular growth and invasiveness in CRC that are phosphorylated in normal and transformed cells have been characterized by phosphoproteomic analysis[7]. Specifically, c-Src activity induces a cluster of tyrosine kinases composed of receptor and cellular tyrosine kinases (MET, EPHA2, FAK, SGK223, SYK) that is crucial for oncogenic signalling[7].

These observations highlight the importance of context in c-Src signalling and the need to characterize mechanisms by which it senses the complex environment. The described interactions detected for the Unique domain of c-Src and the multi-phosphorylation events may play a key role in the integration of c-Src in the signalling networks and still required a full characterization.

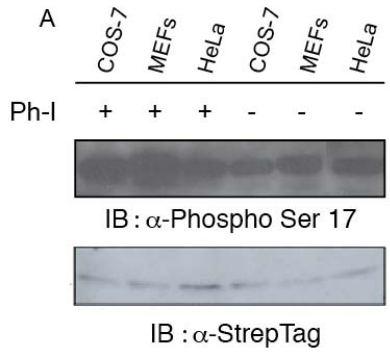
RESULTS

USrc phosphorylation by endogenous kinases in mammalian cells

We have shown that the intrinsically Unique domain of human c-Src is phosphorylated on Ser 17, Ser 69 and Ser 75 by endogenous kinases in *Xenopus Laevis*. In order to study our protein of interest in a more suitable biological system, we decided to investigate USrc in the context of different types of mammalian cell extracts. A ^{15}N-labeled USrc WT construct (residues 1-85) was first expressed in *Escherichia coli*, purified and then incubated into COS-7, MEFs or HeLa cellular extracts. The COS-7 cell line was derived from the kidney of the

African Green Monkey[8], *Cercopithecus aethiops*. MEF cellular line was initiated from murine embryonic fibroblasts transformed with the Simian Virus 40 (SV40). HeLa immortal cellular line was derived from human cervical cancer and is the oldest and most commonly used human cell line[9]. Simultaneously, USrc was also incubated in cellular extracts treated with a cocktail of phosphatases inhibitors (Ph-I) in order to test whether enzymatic reactions involving the Unique domain could be affected by the inhibition of phosphatase activity.

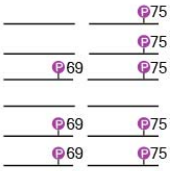
After the incubation, ^{15}N-USrc was purified from cellular extracts by Strep-tag affinity purification and analyzed by Western Blot and mass spectrometry (MS). Biochemical analysis revealed the phosphorylation of Ser 17 (Figure II.3.1.A) while no signals relative to phosphorylated Ser 75 were observed. However, MS detected pSer 75 and pSer 69 (Figure II.3.1.B). Interestingly, different patterns of USrc phosphorylation on serines 69/75 were found by MS depending on the cellular extracts (Figure II.3.B). Ser 75 was found to be phosphorylated in COS-7 cell extracts either in presence or in absence of phosphatases inhibitors while no pSer 69 was detected in this cellular system. pSer 75 and pSer 69 were phosphorylated in MEFs cellular extracts only in presence of phosphatases inhibitors. Moreover, Ser 69 was found phosphorylated almost 4-fold more than Ser 75. In contrast with other conditions, no Ser 69 or 75 phosphorylation was observed in absence of phosphatases inhibitors in MEFs. Phosphorylation of both serines was always detected in HeLa cellular extracts. The pSer75/pSer69 ratio was around 1 in absence of Ph-I while an increase of pSer75 (about 3-fold higher than pSer69) was observed in presence of Ph-I.



B

Area

Sample	LFGGFNSSDTVTSPQR	S75-p-LFGGFNSSDTVTSPQR	S69-p-LFGGFNSSDTVTSPQR
COS-7 + PhI	10867388,53	1427,38	NF
COS-7 - PhI	4885068,98	274,55	NF
MEFs + PhI	8794071,05	282,74	1052,82
MEFs - PhI	3093002,62	NF	NF
HeLa + PhI	17918360,9	11611,43	4017,66
HeLa - PhI	1659341,67	328,12	357,18



Normalized area

Sample	LFGGFNSSDTVTSPQR	S75-p-LFGGFNSSDTVTSPQR	S69-p-LFGGFNSSDTVTSPQR		RATIO 69/75
COS-7 + PhI	2,62325	0,00034	0,00000		NM
COS-7 - PhI	1,38760	0,00008	0,00000		NM
MEFs + PhI	0,57461	0,00002	0,00007		3,72
MEFs - PhI	0,54564	0,00000	0,00000		NM
HeLa + PhI	2,61717	0,00170	0,00059		0,35
HeLa - PhI	3,25044	0,00064	0,00070		1,08

Figure II.3.1. MS and Western blot analysis of USrc in mammalian cell extracts. (A) Western blots showing phosphorylation of Ser 17 of USrc in COS-7 (1st - 4th lanes), MEFs (2nd - 5th lanes) and HeLa (3rd - 6th lanes) cellular extracts through a pSer 17-specific antibody. Protein loading was determined by using an antibody against the Strep tag. (B) Mass-Spec semi-quantitative analysis of the phosphorylated fragment ^{63}LFGGFNSSDTVTSPQR^{78} of USrc showing pSer 69 and pSer 75 in the different cellular extracts. PhI = Phosphatases inhibitors. NF = not found. NM = not measurable.

USrc mutations lead to a reduction of the activity of the tyrosine kinase c-Src in HEK-293T cells

In order to elucidate the biological functions of the discovered Unique Lipid Binding Region (ULBR) and, as well of the observed

phosphorylation sites in USrc, cDNAs encoding for full-length human c-Src WT and Unique mutant forms (AAA or A3, S69E, S75E) were transiently transfected into HEK-293T cells. The serine to glutamic acid mutation can mimic the effect of phosphorylation if the negative charge of the phosphate group plays an important role. We had previously shown that phosphorylation of Ser 75 does not significantly modify the conformational space of USrc but decrease its capacity to bind to negatively charged lipids[10]. HEK cells are a specific cell line originally derived from human embryonic kidney cells. The 293T cells are created from 293 cells but stably express the SV40 large T antigen, which can bind to SV40 enhancers of expression vectors to increase protein production. After 48 hrs of protein expression (checked by Green Fluorescent Protein, GFP-expression), cells were lysed in ice with RIPA buffer (see Material and Methods). Proteins were separated by SDS-PAGE and analysed by Western blots (Figure II.3.2). A strong decrease in the overall tyrosyl-phosphorylation was observed for USrc AAA and S75E mutants (Figure II.3.2.A-B) with respect to the WT form. The S69E mutant promoted no significant effect. The AAA and S75E mutants were also observed to be less active (Figure II.3.2.C-D) and less stable (Figure II.3.2.E-P) with respect to the WT or S69E counterparts. Moreover, both mutants (AAA, S75E) failed to activate the Ras-mitogen-activated-protein-kinase (Ras-MAPK) signalling pathway normally regulated by c-Src (Figure II.3.2.F-G-H). The observed effects were not a consequence of different transfection yields as the GFP-expression levels, used as a report, were roughly the same in all the conditions (Figure II.3.2.N-Q). No changes were observed in the levels of phosphorylated AKT (Figure II.3.2.I-L-M), most probably due to the fact that in HEK293T cells mTOR is strongly deregulated and, as a consequence, so it is its downstream effector AKT.

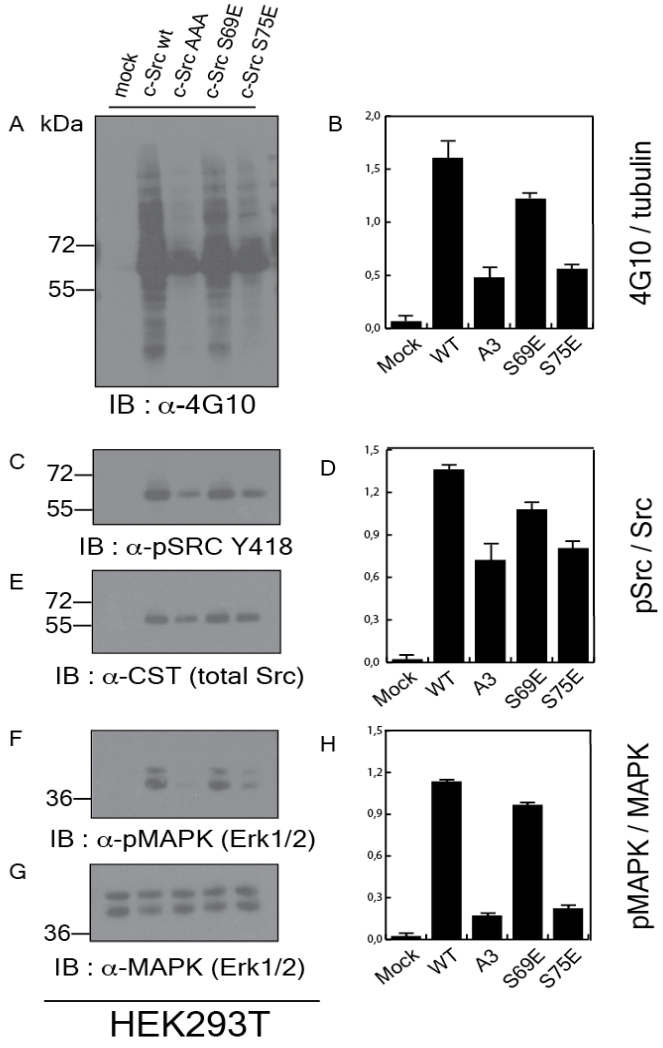


Figure II.3.2.Part I - Unique mutations affect c-Src activity. (A) Western blot showing the tyrosine phosphorylation content of HEK-293T cells transfected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (B) Quantification of (A) as the ratio between A/P (4G10/tubulin). (C) Western blot showing c-Src activation levels. (D) Quantification of (C) as the ratio between C/E (pSrc/total Src). (E) Western blot showing Src level in transfected HEK-293T cells. (F-G) Western blotting analysis of phosphorylated and total MAPK (Erk1, Erk2). (E) Quantification of (F) as the ratio between F/G (pMAPK/total MAPK). Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

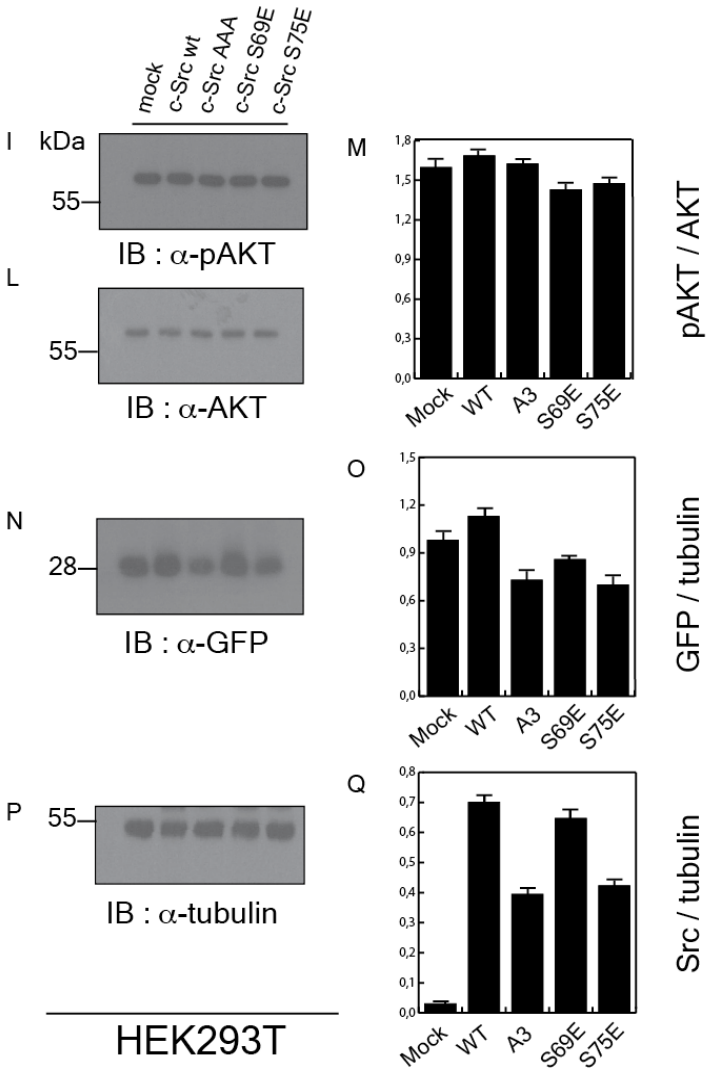


Figure II.3.2.Part II - Unique mutations affect c-Src activity. (I-L) Western blot showing phosphorylated and total AKT in HEK-293T cells transfected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (M) Quantification of (I) as the ratio between I/L (pAKT/total AKT). (N) Western blot showing GFP expression levels. GFP is used as a report of the transfection yields (O) Quantification of (N) as the ratio between N/P (GFP/tubulin). (P) Western blot showing tubulin content in HEK-293T cells. Tubulin is used as a loading control. (Q) Quantification of (E) as the ratio between E/P (total Src/tubulin). Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

Treatment of HEK-293T cells with the MG132 proteasome inhibitor

The observation that USrc mutants AAA and S75E were less stable with respect to the WT form could be attributable to a degradation of c-Src[11]. To test this hypothesis, we inhibited the proteasomal degradation in HEK-293T with MG132 inhibitor. cDNAs encoding for full-length human c-Src WT and Unique mutant forms (AAA or A3, S69E, S75E) were transiently transfected into HEK-293T cells. After 40 hrs of protein production, cells were treated for 8 hrs with MG132 to inhibit the proteasomal machinery. Cells were then lysed in ice with RIPA buffer. Proteins were separated by SDS-PAGE and analysed by Western blots. Results are shown in Figure II.3.3. The inhibition of the proteasomal complex restored the expression levels of mutant S75E (Figure II.3.3.E-Q). A complete recover of its activity and the capacity of regulating its substrates were also observed (Figure II.3.3.A-B-C-F-G-H). Treatment with MG132 also led to the stabilization of AAA mutant (Figure II.3.3.E-Q) but its activity was still 2-3 fold less than the WT form of c-Src (Figure II.3.3.C-D). Interestingly, this mutant is capable to activate itself (by autophosphorylation) but is not able to completely phosphorylate all its subset of substrates (Figure II.3.3.A-B-F-G-H). These effects were not a consequence of different transfection yields as the GFP-expression levels, used as a report, were roughly the same in all the conditions (Figure II.3.3.N-Q). No changes were observed for pAKT (Figure II.3.3.I-L-M).

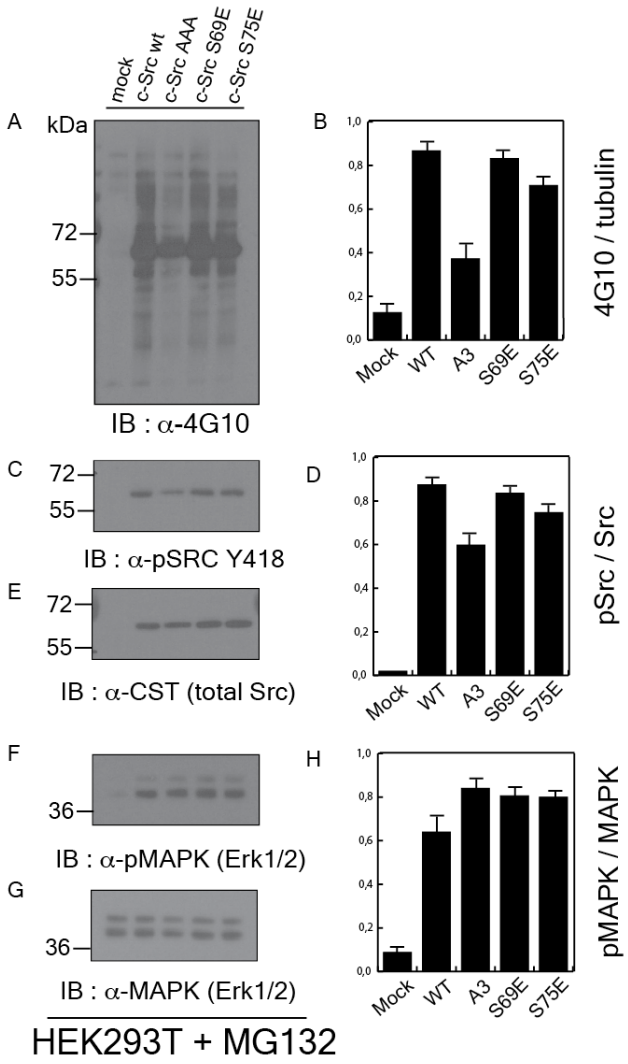


Figure II.3.3.Part I - Effect of MG132 on c-Src. (A) Western blot showing the tyrosine phosphorylation content of HEK-293T cells transfected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (B) Quantification of (A) as the ratio between A/P (4G10/tubulin). (C) Western blot showing c-Src activation levels. (D) Quantification of (C) as the ratio between C/E (pSrc/total Src). (E) Western blot showing Src level in transfected HEK-293T cells. (F-G) Western blotting analysis of phosphorylated and total MAPK (Erk1, Erk2). (E) Quantification of (F) as the ratio between F/G (pMAPK/total MAPK). Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

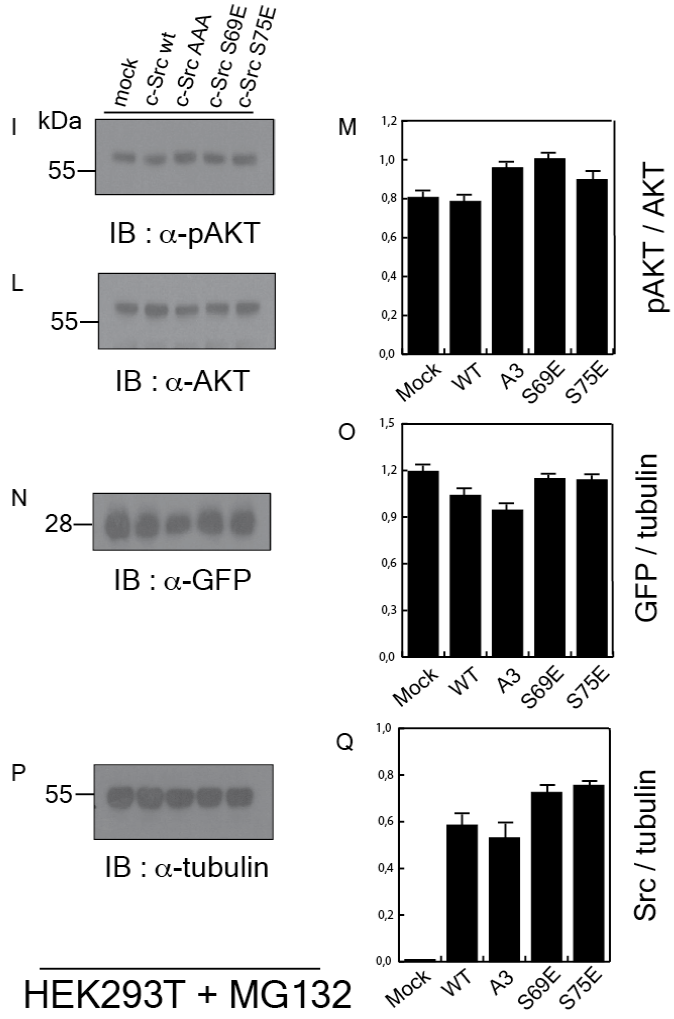


Figure II.3.3.Part II - Effect of MG132 on c-Src. (I-L) Western blot showing phosphorylated and total AKT in HEK-293T cells transfected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (M) Quantification of (I) as the ratio between I/L (pAKT/total AKT). (N) Western blot showing GFP expression levels. GFP is used as a report of the transfection yields (O) Quantification of (N) as the ratio between N/P (GFP/tubulin). (P) Western blot showing tubulin content in HEK-293T cells. Tubulin is used as a loading control. (Q) Quantification of (E) as the ratio between E/P (total Src/tubulin). Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

Creation of SW620 stable cell lines expressing c-Src Unique mutants

In order to study the effect of Unique mutants in cancer cells, four CRC SW620 stable cell lines overexpressing respectively c-Src WT, c-Src AAA, c-Src S69E and c-Src S75E were produced. SW620 cell line derives from Dukes' type C, colorectal carcinoma[12]. SW620 are transformed, metastatic cells where c-Src is highly deregulated even if they exhibit a moderate level of endogenous Src compared with the metastatic Colo205 cells[13]. Thus, it makes them a suitable system to study c-Src in a tumorigenic environment. The cDNAs encoding for the WT and Unique mutant forms (AAA or A3, S69E, S75E) of human c-Src were transiently co-transfected into HEK-293T cells with cDNAs encoding for gag/Pol and ENV proteins in order to allow retrovirus production. After 48 hrs of virus production the supernatant was recover and added onto CRC SW620 cells for virus infection. Positive cells were separated by fluorescence-activated cell sorting (FACS). Two different populations were put in culture after the FACS analysis: medium GFP-expressing cells and high GFP-expressing cells (Figure S5). After 1 week of recovering, high GFP-expressing cells started to die. Most probably, the huge expression of c-Src led to an increased transforming activity that caused cell death. A stable cellular line infected with the empty vector (mock) was also created as a negative control.

c-Src overexpression induces a clear morphological change in CRC SW620 cells. This phenotype was observed both for the WT and for the three Unique mutants (Figure II.3.4).

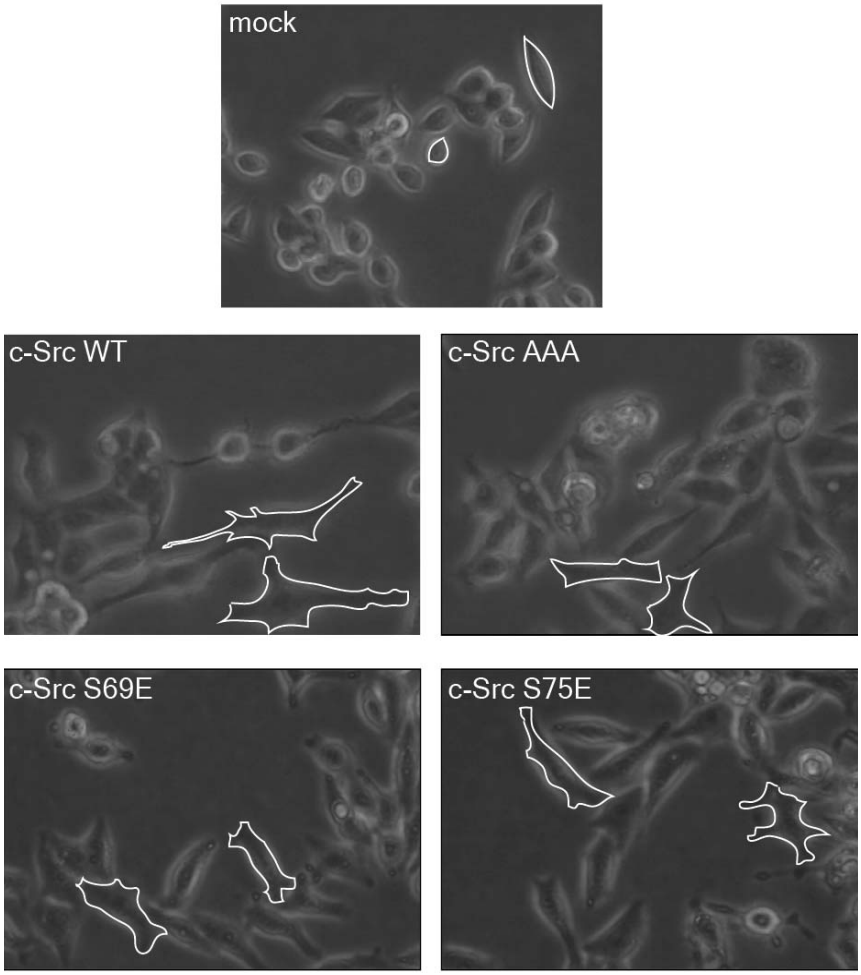


Figure II.3.4. The overexpression of c-Src promotes a morphological change in SW620 cancer cells. CRC SW620 cells present a rounded shape as in the case of the negative control (mock). When those cells overexpress c-Src, a global morphological change is observed (WT, AAA, S69E and S75E). The appearance of multiple protuberances from the plasma membrane is highlighted in white. An increased mobility is also observed and was previously described[7].

Biochemical assays in SW620 stable cell lines expressing c-Src Unique mutants

The created SW620 stable cell lines were used in order to study the effect of the Unique mutants on the signalling pathways regulated by c-Src. SW620 cancer cells overexpressing c-Src WT and mutant forms (AAA or A3, S69E, S75E) were grown until 80-90% of confluence was reached and then lysed in ice with RIPA buffer. The whole lysates were separated by SDS-PAGE and analysed by Western Blots (Figure II.3.5). A >5-fold increase in c-Src (WT and mutants, Figure II.3.5.P) protein level was obtained which induces a strong increase in tyrosine phosphorylation cell content (Figure II.3.5.A-B-C). Of all the transformed cell lines, the one expressing the S75E mutant had the lowest global activity (10-20% less than WT or AAA and S69E) although their cellular phosphotyrosine level was still much higher than in the negative control (Figure II.3.5.A-C). Interestingly, the tyrosine phosphorylation patterns of the three different Unique mutants and the WT form appeared slightly different (Figure II.3.5.B, red arrows). The level of c-Src activation was roughly the same for all conditions (Figure II.3.5.D-F). No changes were observed in MAPK (Figure II.3.5.G-H-I) or AKT signalling (Figure II.3.5.L-M-N) both found strongly activated by c-Src.

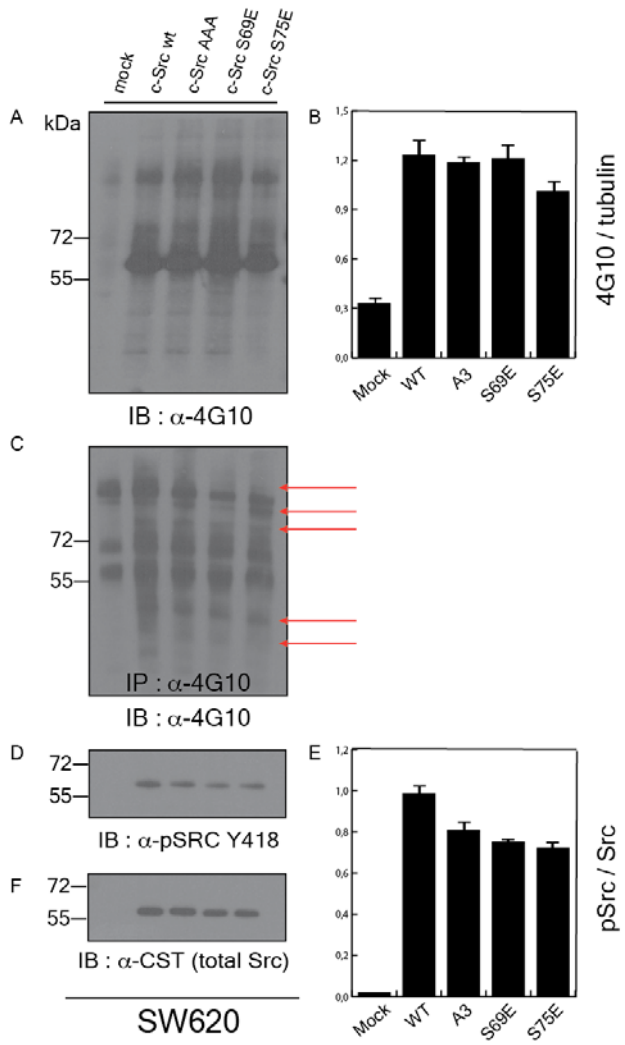


Figure II.3.5.Part I - c-Src in SW620 cancer cell. (A) Western blot showing the tyrosine phosphorylation content of SW620 stable cell line infected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (B) Western blot showing the tyrosine phosphorylation content that was immunoprecipitated (IP) from SW620 cell lysates using the indicated antibody (see Material and Methods) (C) Quantification of (A) as the ratio between A/O (4G10/tubulin). (D) Western blot showing c-Src activation levels. (E) Quantification of (C) as the ratio between D/F (pSrc/total Src). (F) Western blot showing Src level in infected SW620. Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

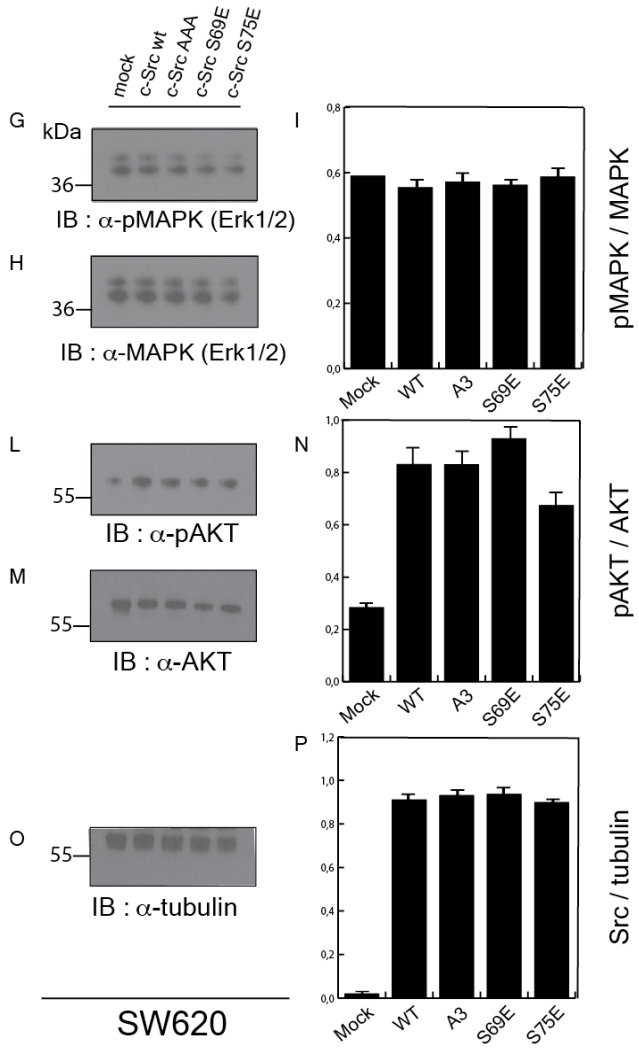


Figure II.3.5.Part II - c-Src in SW620 cancer cell. (G-H) Western blotting analysis of phosphorylated and total MAPK (Erk1, Erk2) in SW620 cells infected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (I) Quantification of (G) as the ratio between G/H (pMAPK/total MAPK). (L-M) Western blot showing phosphorylated and total protein level of AKT. (N) Quantification of (L) as the ratio between L/M (pAKT/total AKT). (O) Western blot showing tubulin content in SW620 cells. Tubulin is used as a loading control. (P) Quantification of (F) as the ratio between F/P (total Src/tubulin). Columns in quantification assays represent the mean of 3 independent experiments. Errors bars: standard deviation.

Unique mutants display reduced invasiveness in CRC SW620 cells

It had been previously demonstrated that the overexpression of the wild type form of human c-Src strongly increased cell invasiveness in CRC SW620 cells[7]. Cell invasion assays were carried out to test whether Unique mutants have an effect on the ability of SW620 cells to invade. Results are shown in Figure II.3.6. As described in Ref. 7, cells infected with the wild type form of c-Src displayed enhanced cell invasiveness in comparison with control cells (>30 fold higher). On the other hand, the invasion ability of Unique mutants was strongly affected. In particular, the AAA and S69E mutants showed a reduction of cell invasiveness higher than 50% with respect to wild type, while only a 20-30% of decrement was observed for the S75E Unique mutant.

The enhanced cell invasiveness induced by the overexpression of the wild type form of c-Src relies on its interaction with a cluster of tyrosine kinases required for cell invasion[7]. Indeed, Leroy *et al.* demonstrated that in SW620 cancer cells c-Src targets various substrates that were shown to be over-phosphorylated[7]. Biochemical assays were performed to test whether some of those described c-Src substrates involved in cellular invasion where differently regulated by Unique mutants in SW620 cells. CRC SW620 cells overexpressing c-Src WT and mutant forms (AAA or A3, S69E, S75E) were lysed in ice with RIPA buffer. Proteins were separated by SDS-PAGE and analysed by Western Blots (Figure II.3.7). Several differences in the phosphorylation level of previously identified substrates were observed. In particular, the AAA and S69E mutants were shown to phosphorylate the receptor tyrosine kinase Met (Figure II.3.7.A) to a higher level than the WT and S75E mutant, while cells overexpressing S75E mutant displayed a lower phosphorylation level of the cytoplasmic tyrosine kinase Syk (Figure II.3.7.D) as compared to WT

and the S69E and AAA mutants. No changes were observed for the pseudo-kinase Pragmin (Figure II.3.7.G).

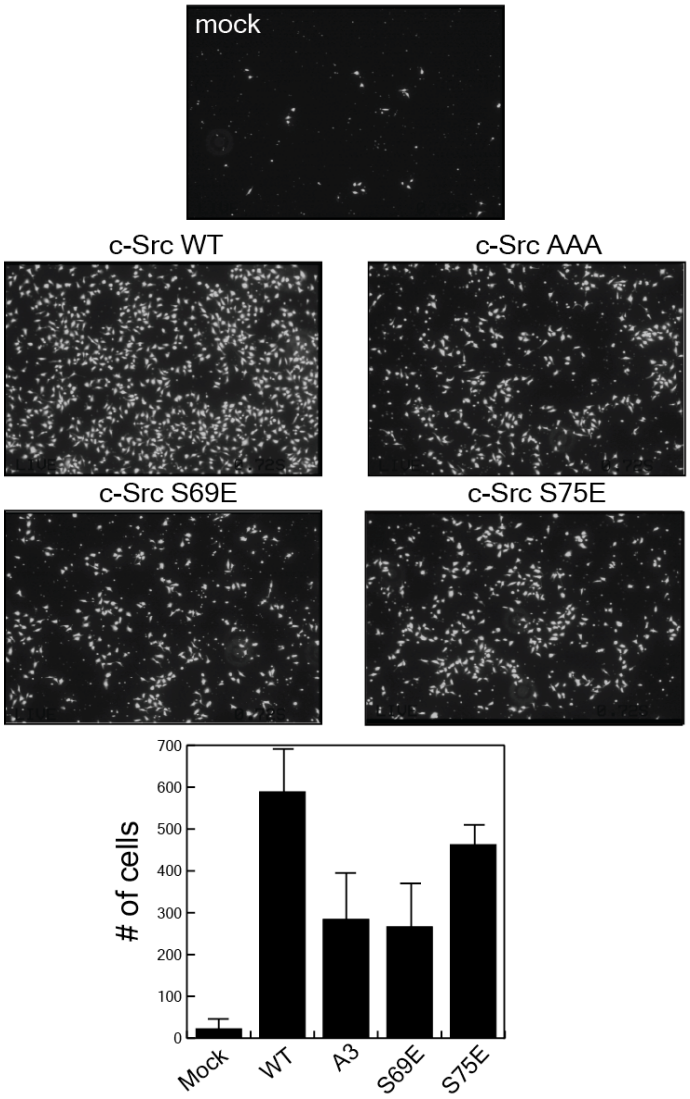


Figure II.3.6. Cell invasion assays of SW620 expressing WT and Unique mutants of c-Src. c-Src promotes cellular invasion through Matrigel of SW620 cells *in vitro*. Unique mutants (AAA, S69E and S75E) display reduced cell invasiveness compared with WT. *X-value* of the quantification plot refers to the number of cells that have invaded through Matrigel in Boyden chamber assays. Columns represent the mean of 3 independent experiments. Errors bars: standard deviation.

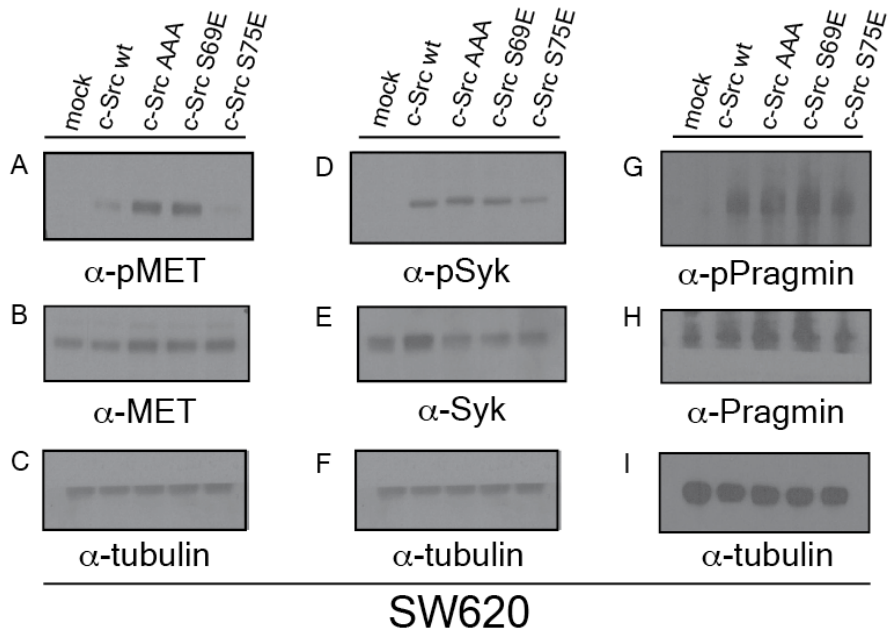


Figure II.3.7. Western blots of c-Src substrates involved in cellular invasion. (A-B) Western blot showing the phosphorylation and the protein levels of the tyrosine receptor kinase Met in SW620 stable cell line infected with: the empty vector (mock, 1st lane), c-Src WT (2nd lane), c-Src AAA (3rd lane), c-Src S69E (4th lane) and c-Src 75E (5th lane). (D-E) Western blot showing the phosphorylation level and the total protein content of the cytoplasmic tyrosine kinase Syk. (G-H) Western blot showing the phosphorylation and protein levels of the pseudo-kinase Pragmin. (C-F-I) Western blot showing tubulin content in SW620 cells. Tubulin is used as a loading control.

Src phosphosphoproteomics in CRC SW620 cells

We next investigated c-Src direct and indirect substrates in CRC SW620 cells by a semi-quantitative phosphoproteomic approach. The aim of the experiment was to identify differences in the cellular phosphorylation pattern between ULBR mutants and the wild type form of human c-Src. Four conditions were analyzed: i) SW620 cells infected with the empty vector (negative control); ii) SW620 cells infected with c-Src WT (positive control); iii) SW620 cells infected with

c-Src AAA mutant; iv) SW620 cells infected with c-Src S75E mutant. The S69E mutant was discarded for technical reasons. Cells were grown in standard conditions and next lysed when a confluence higher than 95% was reached. Proteins were analysed by SDS-PAGE (Figure S6), digested with trypsin, and phosphotyrosine-containing peptides and phosphoserine/threonine-containing peptides were purified from lysates using specific antibodies. After purification, samples were subjected to mass spectrometry analysis. A total of 1044 phosphorylated peptides belonging to 629 proteins were identified with this approach. 711 phosphopeptides (494 proteins) were found for the AAA mutant while 691 phosphopeptides (477 proteins) were encountered for the S75E mutant. Quantifications were made comparing the relative intensities of the obtained peptides for mutants and WT (Figure II.3.8, Figure II.3.9).

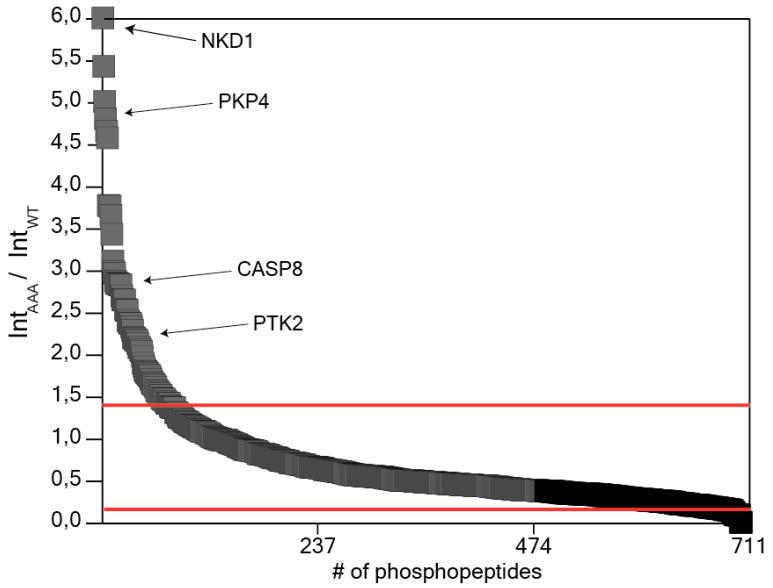


Figure II.3.8. Quantification of Src phosphoproteomic analysis in SW620 cells. The ratio between the relative intensities of the identified phosphopeptides for the AAA mutant and the WT form of c-Src is plotted (summary of fold change), with several individual proteins indicated. Intensity ratios higher than 1.45 (top red line) or lower than 0.18 (bottom red line) were considered significant.

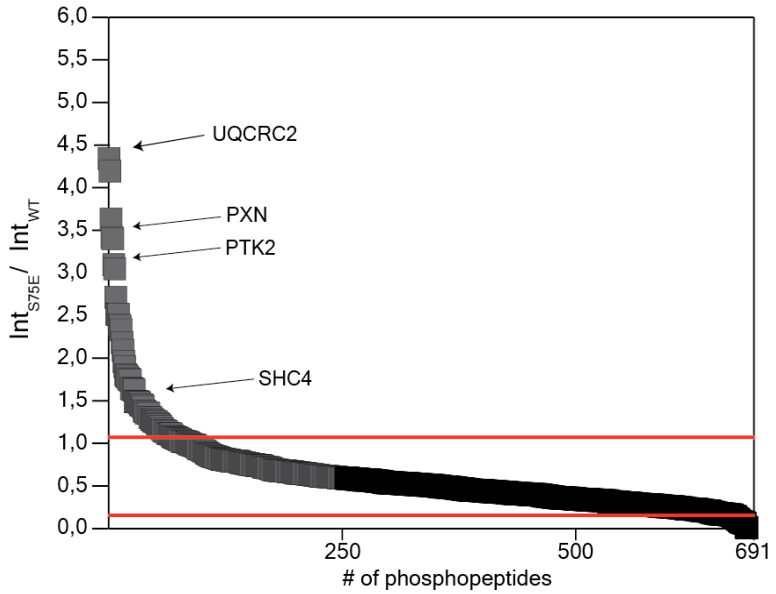


Figure II.3.9. Quantification of Src phosphoproteomic analysis in SW620 cells. The ratio between the relatives intensities of the identified phosphopeptides for the S75E mutant and the WT form of c-Src is plotted (summary of fold change), with several individual proteins indicated. Peptides with ratio >1.1 (top red line) are considered as increased and <0.2 (bottom red line) as decreased in their tyrosine/serine/threonine phosphorylation content.

Based on the distribution of intensity ratios between AAA and WT transformed cells (Figure II.3.8) 104 phosphopeptides (90 proteins) 26 of them exhibiting decreased phosphorylation with respect to the WT and 78 displaying an increased phosphorylation level were selected for further analysis (Table S3 and Table S4). Similarly, 107 phosphopeptides (87 proteins) were considered to have significative concentration differences in cells transformed with the S75E mutant and the WT (Table S5 and Table S6). Of these, 40 showed decrease phosphorylation while 77 exhibited increased phosphorylation (Figure II.3.9) compared to the WT. Genes codifying for the selected proteins (AAA and S75E conditions) were classified according to Gene Ontology (Figure II.3.10, Figure II.3.11).

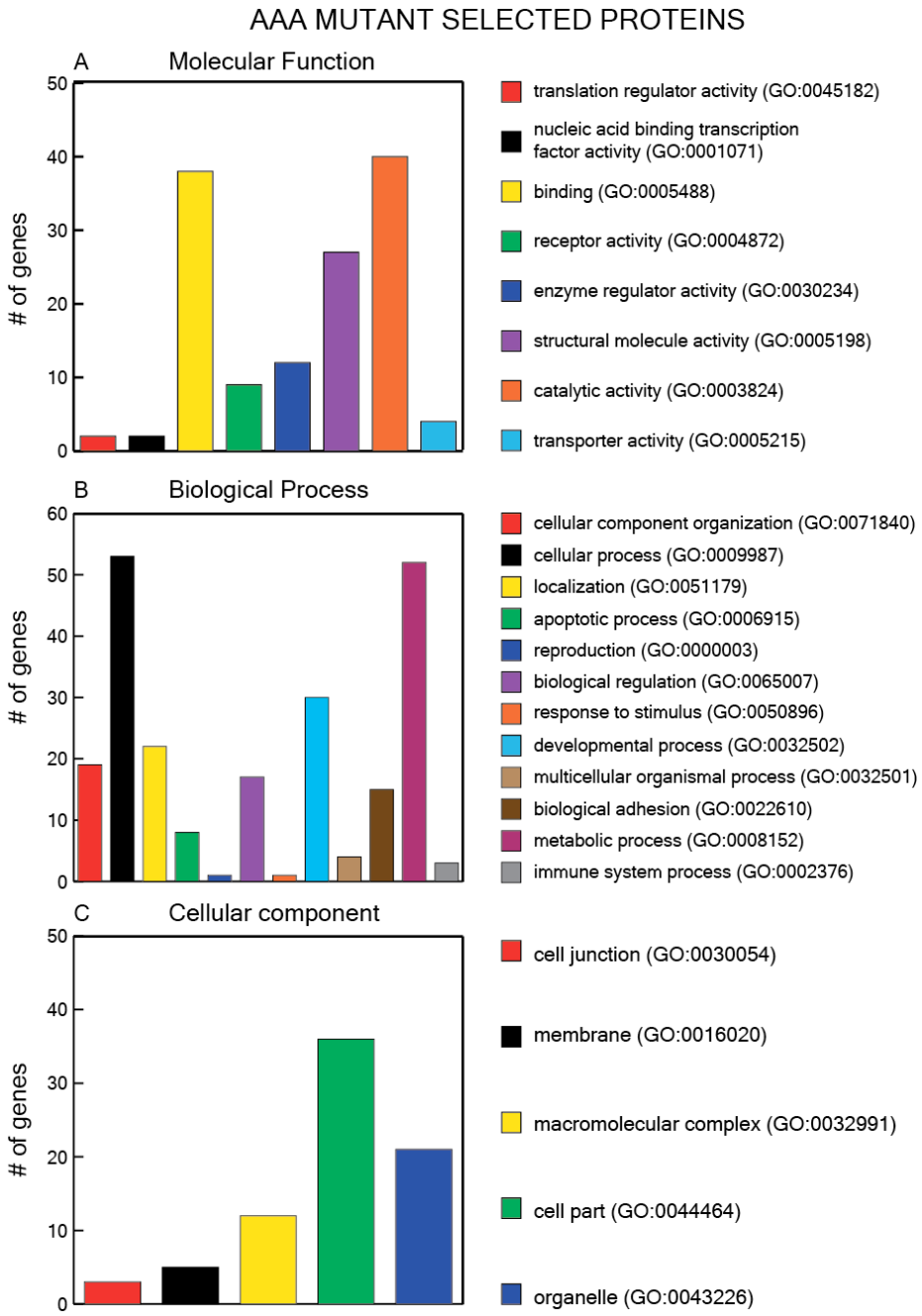


Figure II.3.10. Gene Ontology of AAA selected proteins. Genes encoding for the identified proteins were classified for (A) the molecular function, (B) biological process and (C) cellular compartment on the basis of the gene ontology annotation.

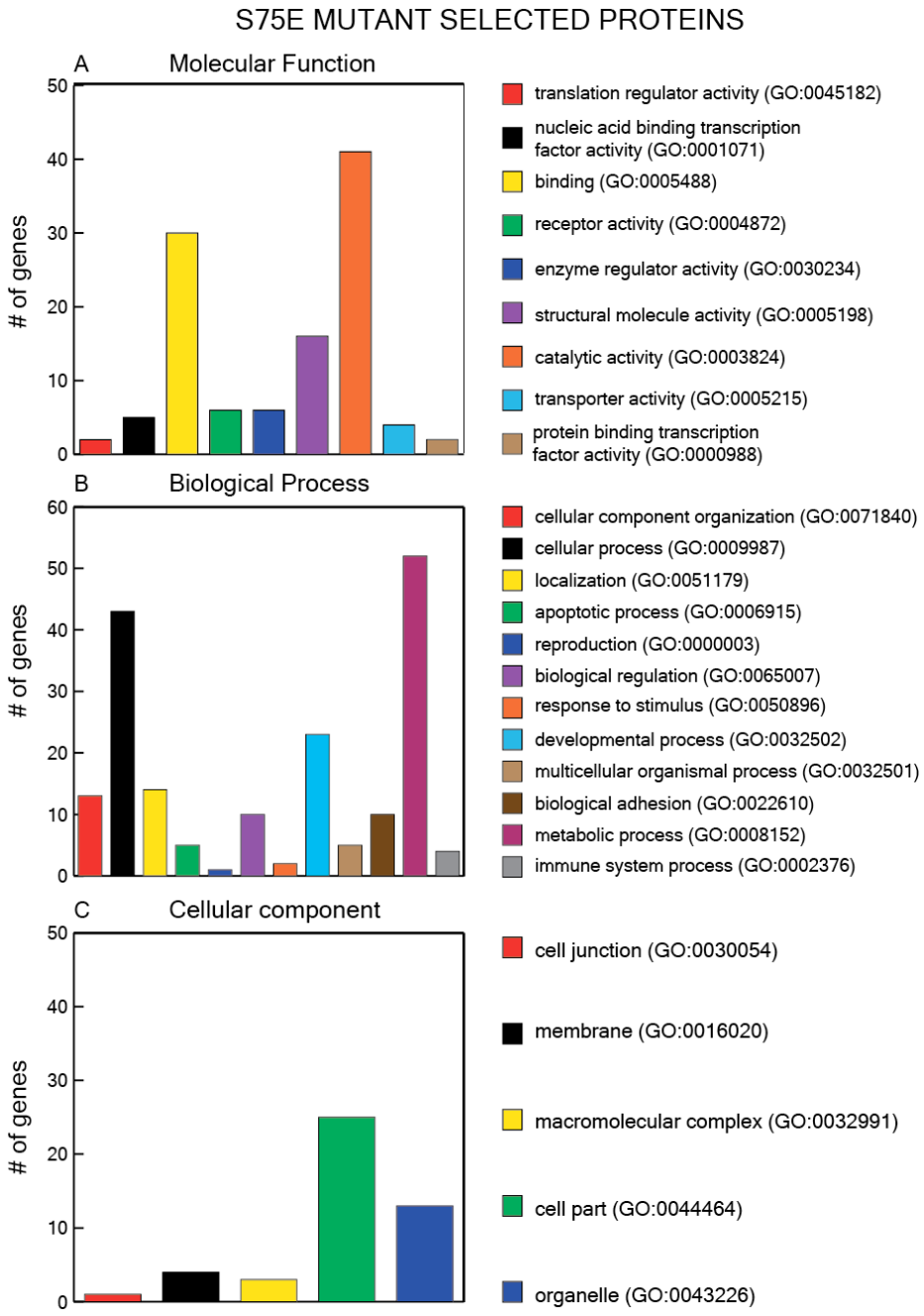


Figure II.3.11. Gene Ontology of AAA selected proteins. Genes encoding for the identified proteins were classified for (A) the molecular function, (B) biological process and (C) cellular compartment on the basis of the gene ontology annotation.

In both cases we observed enrichment (>40 genes) in those genes codifying for proteins involved in cellular processes such as signaling, cell communication and metabolic processes. The major part of the identified proteins (>30 genes) plays roles in binding or catalytic activity and localized at organelles, membranes or are part of macromolecular complexes. Cytoskeletal proteins and transferases represented the most abundant protein classes (Table S7). A summary of the pathways in which identified proteins are involved is shown in Table S7.

The levels of 39 selected phosphopeptides belonging to 39 different proteins were found to be significantly different with respect to WT (decreased/increased phosphorylation) for both AAA and S75E mutants (Table S8). Among them, *PTPRE* (Receptor-type tyrosine-protein phosphatase epsilon) levels were significantly reduced both in cells overexpressing AAA mutant and in cells infected with the S75E mutant. On the other hand, remarkable increased phosphorylation was identified in peptides belonging to Cytochrome b-c1 complex subunit 2 or *PTK2* (Focal adhesion kinase 1). A second group of phosphopeptides includes those found phosphorylated only for the WT but not for the AAA or S75E mutant, or the other way around (phosphorylated for AAA or S75E but not for the WT). The list is shown in Figure II.3.12. *PXN* (Paxillin, on Ser), *ANXA2* (Annexin A2, on Thr) and the C-terminal tyrosine of c-Src (Y527) were found to be strongly phosphorylated in cells overexpressing the WT form of human c-Src but not in those infected with AAA or S75E mutants. On the contrary, Transferrin receptor protein-1 and 6-phosphofructokinase type-C were significantly phosphorylated by both ULBR mutants but not by the WT. Interestingly, pY588 of Ephrin type-A receptor 2 was identified only in SW620 cells infected with AAA mutant.

Protein	Protein names	Gene names	Amino acid	Phospho (STY) Probabilities	Mass error [ppm]	Intensity A3	Intensity WT
P49023-2	Paxillin	PXN	S	FIHQQPQSSSSPVY(0.369)GS(0.13)S(0.501)AK	0.09149	0	17841000
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC	Y	KEPEERPTFEYLQAFLEDY(0.002)FT(0.023)S(0.023)T(0.024)EPQY(0.928)QPGENL	1.6568	0	17392000
P07355	Annexin A2;Annexin;	ANXA2	T	LSLEGDHS(0.014)T(0.785)PPS(0.176)AY(0.025)GSVK	0.60886	0	10300000
B0YJC4	Vimentin	VIM	Y	SLYASSPGGVY(0.993)AT(0.007)R	0.37643	0	5195700
G3V1A4	Cofilin-1;Cofilin-2	CFL1;CFL2	Y	YALYDAT(0.076)Y(0.902)ET(0.022)K	-0.14284	0	1317500
Q8N4X5-2	Actin filament-associated protein 1-like 2	AFAP1L2	S	VAQQPLSLVGCEVPDPS(0.687)PDHLY(0.233)S(0.081)FR	-3.6997	0	1250000
Q5T6L7	Argininosuccinate synthase	ASS1	Y	FAELVY(0.991)T(0.009)GFWHSPECEPVR	0.76096	0	1162200
Q562E7-4	WD repeat-containing protein 81	WDR81	Y	QAFVAGGAGGGEPPHVPHHIS(0.002)DVLS(0.003)DIT(0.059)Y(0.564)Y(0.178)VY(0.194)K	1.4795	0	1034500
F5H855	Breast cancer anti-estrogen resistance protein 1	BCAR1	T	DVPDGPLLREET(0.671)Y(0.329)DVPPAFAK	0.20523	0	1002600
Protein	**Protein names**	**Gene names**	**Amino acid**	**Phospho (STY) Probabilities**	**Mass error [ppm]**	**Intensity A3**	**Intensity WT**
Q8N4X5-2	Actin filament-associated protein 1-like 2	AFAP1L2	S	VAQQPLSLVGCEVPDPS(0.001)PDHLY(0.442)S(0.557)FR	0.68294	2993800	0
Q8NFJ5	Retinoic acid-induced protein 3	GPRC5A	T	AYSQEEIT(0.971)QGFEET(0.022)GDT(0.089)LY(0.697)APY(0.112)S(0.051)T(0.051)HFQLQNQPPQK	-0.25928	2324000	0
B1APP8	6-phosphofructokinase type C	PFKP	Y	NES(0.001)CS(0.001)ENYT(0.002)T(0.001)DFIY(0.993)QLY(0.003)SEEGK	0.37233	2153700	0
F8WBE5	Transferrin receptor protein 1;	TFRC	Y	SAFSNLFGGEPLS(0.015)Y(0.973)T(0.012)R	-0.15413	1472300	0
P29317	Ephrin type-A receptor 2	EPHA2	Y	T(0.173)Y(0.827)VDPHT(0.022)Y(0.978)EDPNQAVLK	0.3557	1314500	0
Q5T6L7	Argininosuccinate synthase	ASS1	T	FAELVY(0.46)T(0.532)GFWHS(0.008)PECEFVR	-0.43476	1225500	0
C9J9K3	40S ribosomal protein SA	RPSA;RPSAP58	Y	ADHQPLT(0.102)EAS(0.193)Y(0.705)VNLPTIALCNTDSPLR	0.56076	1130800	0

Protein	Protein names	Gene names	Amino acid	Phospho (STY) Probabilities	Mass error [ppm]	Int. S75E	Intensity WT
H0YDJ3	Serine/threonine-protein kinase PRP4 homolog	PRPF4B	Y	LCDFGSASHVADNDIT(0.004)PY(0.993)LVS(0.003)R	-0.13834	0	34035000
P49023-2	Paxillin	PXN	S	FIHQQPQSSSSPVY(0.369)GS(0.13)S(0.501)AK	0.09149	0	17841000
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC	Y	KEPEERPTFEYLQAFLEDY(0.002)FT(0.023)S(0.023)T(0.024)EPQY(0.928)QPGENL	1.6568	0	17392000
P07355	Annexin A2;Annexin	ANXA2	T	LSLEGDHS(0.014)T(0.785)PPS(0.176)AY(0.025)GSVK	0.60886	0	10300000
Q8N4X5-2	Actin filament-associated protein 1-like 2	AFAP1L2	Y	VAQQPLSLVGCEVPDPS(0.018)PDHLY(0.766)S(0.216)FR	-0.73294	0	10299000
G3V1A4	Cofilin-1;Cofilin-2	CFL1;CFL2	Y	YALYDAT(0.076)Y(0.902)ET(0.022)K	-0.14284	0	1317500
Q8N4X5-2	Actin filament-associated protein 1-like 2	AFAP1L2	S	VAQQPLSLVGCEVPDPS(0.687)PDHLY(0.233)S(0.081)FR	-3.6997	0	1250000
Q562E7-4	WD repeat-containing protein 81	WDR81	Y	QAFVAGGAGGGEPPHVPHHIS(0.002)DVLS(0.003)DIT(0.059)Y(0.564)Y(0.178)VY(0.194)K	1.4795	0	1034500
B4E022	Transketolase	TKT	Y	NMAEQIIQEIY(0.986)S(0.008)QIQS(0.006)K	-0.65701	0	1014300
Protein	**Protein names**	**Gene names**	**Amino acid**	**Phospho (STY) Probabilities**	**Mass error [ppm]**	**Int. S75E**	**Intensity WT**
B1APP8	6-phosphofructokinase type C	PFKP	Y	NES(0.001)CS(0.001)ENYT(0.002)T(0.001)DFIY(0.993)QLY(0.003)SEEGK	0.37233	1583100	0
F8WBE5	Transferrin receptor protein 1	TFRC	Y	SAFSNLFGGEPLS(0.015)Y(0.973)T(0.012)R	-0.15413	1514900	0
C9J287	Transmembrane protein 106B	TMEM106B	Y	NGDVS(0.001)QFPY(0.994)VEFT(0.005)GR	-0.0024115	1277200	0
E9PKD3	Thioredoxin reductase 1, cytoplasmic	TXNRD1	Y	KVVY(0.932)ENAY(0.068)GQFIGPHR	-0.73942	1037900	0

Figure II.3.12. List of the selected phosphopeptides sorted from phosphoproteomics analysis. Peptides that were found to be phosphorylated only in cells overexpressing WT form or ULBR mutants forms of c-Src are shown. Highlighted boxes refer to peptides shared between AAA and S75E mutants. Peptides were picked from an initial longer list according to these criteria: for WT peptides with intensities > 1,100,000 were selected; for AAA and S75E mutants peptides with intensities > 1,000,000 were selected.

DISCUSSION

We have demonstrated that the Unique domain of human c-Src is phosphorylated on Serine 17, Serine 69 and Serine 75 in several mammalian cell systems. Interestingly, the phosphorylation patterns vary between the different types of cells. The major changes were observed for Ser 69 and Ser 75 while pSer17 was found in all conditions. In COS-7 cells only pSer 75 was detected. This result is in agreement with previous published data in which Pan et *al.* demonstrated that phosphorylation of c-Src at Ser 75 by Cdk5 occurs in COS-7 cells[11].

In MEFs cells phosphorylation of Ser 75 or Ser 69 could only be detected in the presence of phosphatase inhibitors. Under these conditions the phosphorylation of Ser 69 was higher than that of Ser 75. Ser 75 (Ser 74 in mouse Src) was described to be phosphorylated in CCE cells (mouse embryonic stem cells – fibroblasts) by Cdk1 or a related kinase[14] while no data referring to pSer 69 is present in literature for this cell type.

in HeLa cancer cells phosphorylation of Ser 69 and Ser 75 was observed even in the absence of phosphatase inhibitors. Under these conditions the phosphorylation levels of the two sites were similar. In the presence of inhibitors, the phosphorylation of Ser 75 increased about 3-fold while that of Ser 69 remained similar. Phosphorylation of residue Ser 69 of c-Src had previously been detected only by MS in

cell extracts from cancer lines HCT116 and MDA-MB-435S[15], although its function is presently unknown. Thus, our data confirm that Ser 69 can be substantially phosphorylated, especially in cancer cell lines. Most probably the deregulation of cellular mechanisms in tumour cell lines activates a subset of kinases and/or inactivates phosphatases that allow phosphorylation of Ser 69.

The functional roles of the described phosphorylations together with ULBR mutations have been investigated in HEK293T cells and in the human colorectal SW620 cancer cell lines in the context of the full-length protein. Mutation to glutamic acid was used to simulate constituent phosphorylation. We have first shown that the AAA and S75E mutant display reduced activity and protein stability with respect to c-Src wt in HEK293T. The decreased expression level has been related to a degradation phenomenon involving the Ubiquitin-Proteasome system. Indeed, cell treatment with the proteasome inhibitor MG132 led to a full recover of the S75E mutant expression and, as a consequence, of its activity. Pan et *al.* already demonstrated that phosphorylation of c-Src at Ser 75 by Cdk5 is a crucial mechanism for the regulation of intracellular c-Src activity, as it is directly implicated in modulating the ubiquitin-dependent degradation of active c-Src[11]. The expression level of the AAA mutant was also restored after MG132 treatment. Nevertheless, its global activity remained lower than that of the WT. Our data suggest that this mutant is capable to activate itself (by autophosphorylation on Y418) but it display altered levels of activity towards other substrates.

Then, we have investigated the effect of USrc mutations in the colorectal SW620 cell system. USrc mutants as well as wild type induced a strong morphological change in CRC cells. The appearance of multi-protuberances sprouting from the plasma membrane has been observed and it relates with an increased mobility. Indeed, in SW620 cells c-Src targets several cytoplasmic and membrane proteins that are

involved in cellular growth, adhesion and motility[7]. No significant changes were detected in protein activity between mutants and WT. Interestingly, S75E mutant was not affected by degradation phenomena in SW620 cells suggesting that in this system a strong deregulation of the Ubiquitin-Proteasome machinery is also occurring.

c-Src overexpression induces high invasiveness of advanced cancer cells *in vitro*, as previously described by Leroy *et al*[7]. We have demonstrated that ULBR mutations strongly affect the capacity to invade of CRC SW620 cells. Indeed, ULBR mutants display reduced invasiveness (>20-50%) with respect to the WT. The observed effects can be attributable to the fact that mutants differently phosphorylate part of the cluster of substrates important for cell invasiveness. This hypothesis is reinforced by the fact that looking at the global phosphotyrosil cellular content several proteins were shown to display different patterns of phosphorylation between the distinct conditions. Semi-quantitative phosphoproteomic analysis further confirmed the preliminary analysis. More than 100 phosphopeptides were found to be over or under-phosphorylated in cells infected with ULBR mutants when compared to those overexpressing the WT form of human c-Src. Among all, 39 phosphopeptides were shared between the AAA and S75E mutants. Receptor-type tyrosine-protein phosphatase epsilon *(PTPRE)*, Annexin A2 *(ANXA2)* and tyrosine-protein phosphatase non-receptor type 6 *(PTPN6* – SHP-1*)* were some of the proteins showing decreased phosphorylation level in both mutants. Annexin A2 is a calcium regulated membrane binding-protein with high affinity for calcium greatly enhanced by anionic phospholipids such as phosphatidylinositol-4,5-biphosphate[16]. It mainly localized at plasma membrane and it is involved in several biological processes such as angiogenesis[17] or membrane raft assembly[16]. Tyr 24 of ANXA2 is a direct substrate of c-Src[18]. When c-Src phosphorylates this residue at the membrane level, Annexin A2 undergoes a conformational change

that might potentiate its insertion into the plasma membrane enhancing its affinity for the inner leaflet anionic phosphatidylserine[19]. Once phosphorylated, ANXA2 could modulate not only blood fluidity but also extravascular matrix remodeling and the invasion potential of multiple cell types[18]. ULBR mutants were incapable to phosphorylate Tyr 24 of ANXA2. Tyrosine-protein phosphatase SHP-1 can also be phosphorylated by c-Src on Tyr 536 and Tyr 564[20]. The tyrosil phosphorylation leads to the activation of the phosphatase that subsequently acts on several substrates at membrane level involved in many signaling pathways (e.g. EGFR signaling, JAK-STAT cascade). ULBR mutants were found to be unable to phosphorylate Tyr 564. Interestingly, this residue can also be modified by Lck[21] or Lyn[22], hence suggesting that in CRC SW620 c-Src mutants might differentially regulate signaling pathways involving other SFKs. Always in the same shared group we identified several proteins displaying an increased phosphorylation level in cells overexpressing ULBR mutants when compared to the WT. Among them, Focal Adhesion Kinase 1 *(PTK2)* and c-Src itself were the most interesting. PTK2 is a non-receptor protein-tyrosine kinase that plays an essential role in regulating cell migration, adhesion, spreading, reorganization of the actin cytoskeleton, formation and disassembly of focal adhesions and cell protrusions, cell cycle progression, cell proliferation and apoptosis[23]. It regulates numerous signaling pathways such as the AKT1 signaling cascade, the MAPK pathway but is also involved in signaling downstream of numerous growth factor receptors like G-protein coupled receptors (GPCR), EPHA2 (Ephrin-type A receptor-2), netrin receptors and LDL receptors. c-Src can phosphorylate PTK2 on Tyr 576 and Tyr 577 both located in the activation loop[23,24,25]. The phosphorylation events are required for the maximal kinase activity and upon them PTK2 can act as a key regulator of the numerous pathways in which is involved[23]. PTK2 can also phosphorylate c-Src

keeping it more active. ULBR mutants were found to strongly phosphorylate Tyr 577 with respect to WT form. Also c-Src was found to be over-phosphorylated on Tyr 379 in cells infected with ULBR mutants. This residue is located in the catalytic domain but its function is completely unknown and this phosphorylation was never detected until now.

The Ephrin-type A receptor 2 (EPHA2) was found over-phosphorylated on Tyr 588 only in SW620 cells overexpressing the AAA mutant. EPHA2 belongs to the cluster of protein tyrosine kinase regulating cell invasiveness[7]. Similarly, the phosphorylation level of Tyr 608 of EPHB3 (Ephrin-type B receptor-3) was increased in CRC SW620 cells infected with the S75E c-Src variant.

The changes in the phosphorylation levels observed in the whole cell experiments include both direct and indirect effects. A gene ontology (GO) analysis of the most perturbed genes clearly identifies cytoskeletal proteins as the major class, in the case of the AAA mutant and the second major class in the S75E mutant. This result is consistent with the important effects observed in cell invasion capacity.

All together, these data clearly indicate a regulatory role of the Unique domain affecting the overall kinase activity of c-Src as well to its selectivity towards different set of substrates.

Many of the identified proteins showing different pattern of phosphorylation among the analysed condition were targeted by c-Src at the plasma membrane level. c-Src is attached to membrane surfaces by its myristoylated SH4 domain and ULBR constitutes a second lipid binding region. Previous *in vitro* observations showed that phosphorylation of sites close to the ULBR modulated its lipid binding capacity[10,26]. We suggest that, at least some of the selective regulation effects may be mediated by the effective length spanned by the flexible Unique domain. By switching on and off lipid binding by the ULBR, the kinase domain or the domains helping to recognize the substrates are

either confined close to the membrane surface or can span a longer distance. The effective length of the UD introduces a certain degree of compartmentalisation in the direction perpendicular to the plane of the membrane that could prevent or modify the interaction between c-Src and other proteins, even if they are attached next to each other. Thus, modulation of the ULBR-lipid interaction, e.g. by phosphorylation as in the case of S75E mutant, provide a mechanism to change the effective length of the Unique domain that connects the folded domains of c-Src to the membrane, and therefore, the specificity of c-Src with respect to upstream and downstream signalling partners.

The observation of increased phosphorylation, by the AAA or S69E mutants, of known direct substrates of c-Src, such as the MET receptor but not other direct substrates of c-Src that are not a membrane anchored proteins, in spite of the overall decrease in kinase activity of these mutants, would be consistent with this model.

The presence of this additional regulatory mechanism ("positional regulation model", Figure II.3.13) will help to understand how c-Src effectively connects a wide variety of substrates and clarify for the first time the biological role of the intrinsically disordered Unique domain.

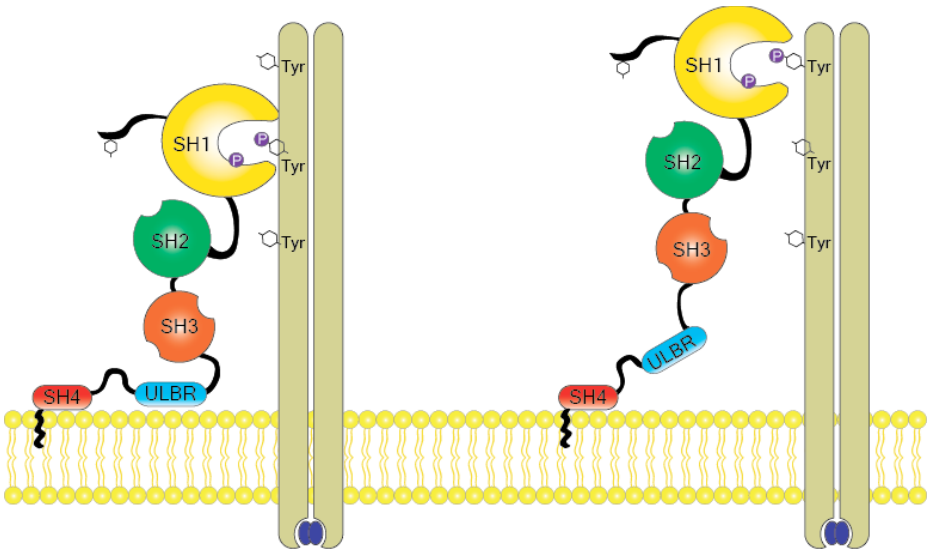


Figure II.3.13. "Positional regulation model" of c-Src. At plasma membrane c-Src interacts with several partners that can be phosphorylated. The Unique domain can modulate the effective length of c-Src in the direction perpendicular to the plane of the membrane through the regulation of the lipid binding (e.g. change of lipid composition; phosphorylation at Ser 75). This additional regulatory mechanism allows that c-Src can change its specificity, and therefore, the ability to control the various signaling pathways in which is involved.

MATERIAL AND METHODS

Cloning, Protein Expression and Purification

The cDNA encoding (1-85) human c-Src region with a Strep-tag in C-terminal position for purification purposes was cloned into a pET-14b vector (Novagen, UK) (Figure A.3.1). The glycine residue (G2) following the initial methionine was mutated to alanine to prevent myristoylation. The plasmid was transformed in *Escherichia coli* Rosetta^{TM} (DE3) pLysS cells (Novagen, UK) and cells were grown in M9 minimal medium supplemented with [^{15}N]H_4Cl (Cambridge Isotope Laboratories, UK). USrc protein was isolated using Strep-tactin Sepharose (IBA, Göttingen). After elution with 2.5 mM of desthiobiotin, it was further purified by size exclusion chromatography (Superdex 75 26/60, GE Healthcare, Spain) in phosphate buffer (50 mM phosphate, 0.2 mM EDTA, pH 7.0). Further details are described in II.1 – Material and Methods.

Mammalian cells extracts

COS-7, MEFs and HeLa cells were cultured in complete DMEM supplemented with 10% FBS in a humidified incubator at 37 °C and 5% CO_2. Cells were split into 4 culture dishes (165 cm^2) 24 hrs before preparing the cell extracts in order to reach 60–70% confluence at the time of extract preparation. Cells were first detached from dishes with cold PBS (10 mL per dish), collected into a 50-ml tube and sediment by low-speed centrifugation. Pellet was resuspended in ice with 50 μL of "NMR lysis buffer" (45 mM HEPES pH 7.2, 0.5 mM EDTA, 10 mM $MgCl_2$, 50 mM KCl, 3% (v/v) NP-40, 10% (v/v) glycerol, 1 tablet of complete protease inhibitor per 6.5 ml of lysis buffer, 1 tablet of PhosSTOP phosphatase inhibitor per 1.25 ml of lysis buffer, 1% (v/v) phosphatase inhibitor cocktails 2 and 3, and 3 mM DTT). Aliquots were

snap-frozen in liquid nitrogen and stored at -80 °C. The following products were used for "NMR lysis buffer": Complete protease inhibitor (Roche, cat. no. 11697498001), PhosSTOP phosphatase inhibitor (Roche, cat. no. 04906837001), Phosphatase inhibitor cocktail 2 (Sigma-Aldrich, cat. no. P5726), Phosphatase inhibitor cocktail 3 (Sigma-Aldrich, cat. no. P0044). Protocols are detailed in Ref. 27.

Western Blotting & Mass spectrometry

USrc was recovered from mammalian cell extracts by Strep-tag affinity purification after an over night incubation at room temperature. 0.5 µg of purified protein was separated by SDS-PAGE and transferred onto Hybond ECL nitrocellulose membrane (GE Healthcare). Membranes were blocked with 5% Milk in TBS-Tween and immunoblotted with the following antibodies: 1:1,000 anti-PhosphoSrc-Ser17 (#5473, Cell Signaling Technology), and 1:20,000 anti-Streptag antibody (IBA, Göttingen).

20 mg of ^{15}N-labeled USrc recovered from extracts, or unlabeled USrc treated in identical way, was digested by adding trypsin or chymotrypsin (2% w/v) and incubated at 37°C overnight. Digestions were stopped by adding formic acid to a final concentration of 1%. The resulting peptide mixtures were diluted in 1% FA and loaded in a nano-LC-MS/MS system connected to an Advion TriVersa NanoMate (Advion) fitted on an LTQ-FT Ultra mass spectrometer (Thermo Scientific). Further experimental details are given as supplementary information (see Addendum – A.4).

A database search was performed with Bioworks v3.1.1 SP1 and Proteome Discoverer software v1.3 (Thermo Scientific) by using the Sequest search engine and a home-made database in SwissProt format, which included USrc protein and the common repository of adventitious proteins (http://www.thegpm.org/crap/index.html). Search

parameters included no-enzyme specificity, methionine oxidation and phosphorylation in serine and threonine as dynamic modification and, depending on the sample, amino acids labeled with ^{15}N as static modification. Extracted ion chromatograms of MS or MS/MS ions were obtained using Xcalibur software vs 2.0SR2 (Thermo Scientific).

Cell culture, infections and invasion assays

Human c-Src was subcloned in pMX-pS-CESAR vector (Figure A.3.4). Mutations were introduced using the QuikChange site-directed mutagenesis kit (Stratagene). HEK-293T cells were maintained in culture with DMEM supplemented with 10% FBS in a humidified incubator at 37 °C and 5% CO_2. Transient transfections were performed using jetPEI reagent (Polyplus Transfection) according to manufacturer's procedures. SW620 cells (American Type Culture Collection) were grown, infected and selected as described in Ref. 28. c-Src expressers were isolated by fluorescence-activated cell sorting and maintained in culture with DMEM supplemented with 10% FBS in a humidified incubator at 37 °C and 5% CO_2. Cell invasion assays were performed in Boyden chambers (BD Bioscience) using 100,000 cells and in presence of 100 μL of Matrigel (0.5 mg/mL; BD Bioscience). After 28-32 hrs of invasion (37°C), cells were resuspended in PBS with Calcein for 30 minutes. After calcein staining images were acquired using an optical microscope.

Antibodies and reagents

The following antibodies were used in this study: anti-4G10 (phosphotyrosil content; 1:50, Serge Roche's lab), anti-phospho-Src-Y418 (1:1000, Cell Signaling), anti-Cst (total Src; 1:1000, Serge Roche's lab), anti-tubulin (1:1000, Serge Roche's lab), anti-phospho-

AKT (1:1000, Cell Signaling), anti-AKT (1:1000, Cell Signaling), anti-phospho-MAPK (phospho-ERK1/2; 1:1000, Cell Signaling), anti-MAPK (ERK1/2; 1:1000 Cell Signaling), anti-GFP (1:3000, Serge Roche's lab), anti-MET (1:1000, Santa Cruz Biotech), anti-phospho-MET (1:1000, Cell Signaling), anti-phospho-Syk (1:1000, Santa Cruz Biotech), anti-Syk (1:1000, Santa Cruz Biotech), anti-phospho-Pragmin (1:3000, Serge Roche's lab) and anti-Pragmin (1:400, Abnova). MG132 proteasome inhibitor (5 μg, final concentration) and Calcein were from Sigma (cat. no. C2211; C0875). HEK293T and CRC SW620 cells were lysed in ice with a modified version of the original "RIPA buffer"[29] (20 mM Tris pH 7.4, 150 mM NaCl 5M, 0,5%Triton-X 100, 0,5% DOX – deoxycholate, 60 mM Octyl glucoside, 1% Aproptin, 20 mM Leupeptin, 1 mM Orthovanadate, 1 mM NaF).

Semi-quantitative phosphoproteomics analysis of CRC SW620 cells

Phosphotyrosine immunoaffinity purification of digested peptides (using a mixture of a mixture of 4G10 and pY100 antibodies), was essentially performed as described in Ref. 30, except that cells ($2e10^8$) were previously treated with the tyrosine phosphatases inhibitor orthovanadate (1 mM) and the phosphatases inhibitor PhosSTOP (Roche, cat. no. 04906837001). Trypsin digestion was carried out after reduction/alkylation steps. A pSer/pThr enrichment was further performed using IMAC immunoaffinity purification (Sigma Aldrich, cat. no. I1408). Samples containing purified phosphopeptides were sent to Novartis, Switzerland for LC/MS/MS analysis. Results were then sorted at the FPP (Plate-forme de Protéomique Fonctionalle, Montpellier).

REFERENCES

1. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13,** 513-609 (1997).

2. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4,** 470-80 (2004).

3. Zhang, X. H. *et al.* Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell* **16,** 67-78 (2010).

4. Summy, J. M. & Gallick, G. E. Src family kinases in tumor progression and metastasis. *Cancer Metastasis Rev.* **22,** 337-58 (2003).

5. Irby, R. B. *et al.* Activating SRC mutation in a subset of advanced human colon cancers. *Nat. Genet.* **21,** 187-90 (1999).

6. Emaduddin, M., Bicknell, D. C., Bodmer, W. F. & Feller, S. M. Cell growth, global phosphotyrosine elevation, and c-Met phosphorylation through Src family kinases in colorectal cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 2358-62 (2008).

7. Leroy, C. *et al.* Quantitative phosphoproteomics reveals a cluster of tyrosine kinases that mediates SRC invasive activity in advanced colon carcinoma cells. *Cancer Res.* **69,** 2279-86 (2009).

8. Gluzman Y. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182 (1981).

9. Rahbari, R., Sheahan, T., Modes, V., Collier, P. & Macfarlane, C. A novel L1 retrotransposon marker for HeLa cell line identification. *Biotechniques* **46,** 277-284 (2009).

10. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391,** 136-48 (2009).

11. Pan, Q. *et al.* Cdk5 targets active Src for ubiquitin-dependent degradation by phosphorylating Src(S75). *Cell. Mol. Life Sci.* **68,** 3425-36 (2011).

12. Leibovitz A, *et al.* Classification of human colorectal adenocarcinoma cell lines. *Cancer Res.* **36**, 4562-4569 (1976).

13. Dehm, S., Senger, M. a & Bonham, K. SRC transcriptional activation in a subset of human colon cancer cell lines. *FEBS Lett.* **487,** 367–71 (2001).

14. Kato, G. & Maeda, S. Production of Mouse ES Cells Homozygous for Cdk5-Phosphorylated Site Mutation in c-src Alleles. *J. Biochem.* **133,** 563–569 (2003).

15. Oppermann, F. S. *et al.* Large-scale proteomics analysis of the human kinome. *Mol. Cell. Proteomics* **8,** 1751–64 (2009).

16. Drücker, P., Pejic, M., Galla, H.-J. & Gerke, V. Lipid segregation and membrane budding induced by the peripheral membrane binding protein annexin A2. *J. Biol. Chem.* **288,** 24764–76 (2013).

17. Aitkenhead, M. *et al.* Identification of endothelial cell genes expressed in an in vitro model of angiogenesis: induction of ESM-1, (beta)ig-h3, and NrCAM. *Microvasc. Res.* **63,** 159–71 (2002).

18. Deora, A. B., Kreitzer, G., Jacovina, A. T. & Hajjar, K. a. An annexin 2 phosphorylation switch mediates p11-dependent translocation of annexin 2 to the cell surface. *J. Biol. Chem.* **279,** 43411–8 (2004).

19. Montaville, P., Neumann, J.-M., Russo-Marie, F., Ochsenbein, F. & Sanson, A. A New Consensus Sequence for Phosphatidylserine Recognition by Annexins A New Consensus Sequence for Phosphatidylserine Recognition by Annexins *. *J. Biol. Chem.* **277,** 24684–93 (2002).

20. Frank, C. *et al.* Effective dephosphorylation of Src substrates by SHP-1. *J. Biol. Chem.* **279,** 11375–83 (2004).

21. Lorenz, U. *et al.* Lck-dependent tyrosyl phosphorylation of the phosphotyrosine phosphatase SH-PTP1 in murine T cells. *Mol. Cell. Biol.* **14,** 1824–34 (1994).

22. Yoshida, K., Kharbanda, S. & Kufe, D. Functional Interaction between SHPTP1 and the Lyn Tyrosine Kinase in the Apoptotic Response to DNA Damage. *J. Biol. Chem.* **274,** 34663–34668 (1999).

23. Parsons, J. T., Martin, K. H., Slack, J. K., Taylor, J. M. & Weed, S. a. Focal adhesion kinase: a regulator of focal adhesion dynamics and cell movement. *Oncogene* **19,** 5606–13 (2000).

24. Oh, M.-A. *et al.* Specific tyrosine phosphorylation of focal adhesion kinase mediated by Fer tyrosine kinase in suspended hepatocytes. *Biochim. Biophys. Acta* **1793,** 781–91 (2009).

25. Relou, I. a M., Bax, L. a B., van Rijn, H. J. M. & Akkerman, J.-W. N. Site-specific phosphorylation of platelet focal adhesion kinase by low-density lipoprotein. *Biochem. J.* **369,** 407–16 (2003).

26. Pérez, Y. *et al.* Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* **3,** 1295 (2013).

27. Theillet, F.-X. *et al.* Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat. Protoc.* **8,** 1416–32 (2013).

28. Sirvent, a, Boureux, a, Simon, V., Leroy, C. & Roche, S. The tyrosine kinase Abl is required for Src-transforming activity in mouse fibroblasts and human breast cancer cells. *Oncogene* **26,** 7313–23 (2007).

29. Harlow, E. and Lane, D. Antibodies a Laboratory Manual. *Cold Spring Harbor Press*, Cold Spring Harbor, NY, p. 449 (1988).

30. Amanchy, R., Kalume, D. E., Iwahori, A., Zhong, J. & Pandey, A. Phosphoproteome Analysis of HeLa Cells Using Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC). *J. Proteome Res.* **4,** 1661–1671 (2005).

Part III

Conclusions

III.1 – Concluding remarks

The present thesis addressed fundamental aspects of the biological role of the intrinsically disordered Unique domain of c-Src that until now was poorly understood. We have first characterized through a structural biology approach the effects of site-directed ULBR mutations on the previously described interactions with lipids and with the folded SH3 domain. Then, we have investigated the phosphorylation events taking place on the Unique domain by state-of-the-art in-cell NMR techniques. Finally, we have shown the functional role of USrc in the context of normal and cancer cell lines. All together, our results have demonstrated that the Unique domain of human c-Src actively participates in the regulation mechanism in which the tyrosine kinase c-Src is involved. This study represents a new essential milestone in the understanding of how c-Src can control a wide variety of signaling events inside cells. Furthermore, it potentially opens the way to the possibility of new applications for cancer therapy through the development of novel selective drugs targeting functional active residues locating in the Unique domain.

This thesis represents a “proof-of-concept” of how from the atomistic comprehension of a subject of interest (structural approach) it can be possible to understand and successively demonstrate the biological significance of the observed phenomena (functional approach).

III.2 – Bullet list of conclusion

- Mutations in the Unique Lipid Binding Region affect the lipid binding capability of USrc but not the inter-domain interaction with the folded SH3.
- The SH3 domain of human c-Src loses its ability to interact with lipids in presence of a polyproline peptide.
- Real-time NMR spectroscopy allowed the study of complex phosphorylation/dephosphorylation processes in USrc mediated by kinases and phosphatases in Xenopus laevis derived cells or cell extracts.
- Three different phosphorylation sites present in the Unique domain that were unequivocally assigned to Ser 17, Ser 69 and Ser 75 have been identified in Xenopus laevis, COS-7, MEFs and HeLa cell systems.
- It was the first time that phosphorylation of Ser 69 of c-Src was observed in those model systems.
- The biological relevance of the ULBR has been firstly demonstrated in Xenopus laevis oocytes in the context of the full-length c-Src.
- The functional roles of the described phosphorylation events together with ULBR mutations have been shown in HEK293T cells and in the human colorectal SW620 cancer cell lines in the context of the full-length protein.
- A “positional regulation model” has been proposed as a new c-Src selectivity regulation mechanism.

Part IV

Resum

Resum en català

Caracterització estructural i funcional del domini únic intrínsecament desplegat de c-Src humana.

Motivacions

Encara que c-Src s'hagi estudiat àmpliament en tot el món durant dècades, poc se sap sobre la funció del seu domini únic intrínsecament desordenat (USrc). En un treball recent s'ha descobert un ric repertori d'interaccions que impliquen residus específics del domini únic, el que suggereix un paper més específic que el d'un senzill espaiador.

Aquesta tesi té com a objectiu principal avançar en la caracterització estructural i funcional del domini intrínsecament desordenat de la tirosina cinasa humana c-Src.

c-Src també està implicada en diversos tipus de càncer. En contrast amb oncogens virals, el potencial oncogènic de c-Src no està associat a mutacions sinó a la desregulació de les xarxes de senyalització. Per tant, el descobriment d'un nou mecanisme de regulació que implica el domini únic podria conduïr al desenvolupament de nous fàrmacs específics que podrien inhibir l'activitat oncogènica de c-Src. Apuntar el domini únic té el potencial d'una selectivitat molt més gran i, potencialment, efectes més nets que els de la majoria dels fàrmacs existents que, actuant sobre el lloc d'unió d'ATP de la SFKs (família Src de cinases no receptores), pateixen de mala especificitat. Aquestes perspectives enforteixen encara més la importància biològica d'aquesta investigació.

ANTECEDENTS

c-Src tirosina cinasa

c-Src és el membre principal de la família Src de cinases no receptores (SFKs) i ha estat objecte d'intensa recerca durant dècades. Aquests estudis sorgeixen dels treballs sobre el virus del sarcoma de Rous, un virus tumoral de pollastre descobert el 1911 per Peyton Rous[1]. El gen que codifica per la proteïna src viral (v-src) va ser el primer oncogèn retroviral a ser identificat, i el seu homòleg cel·lular va ser el primer proto-oncogen que es va descobrir al genoma dels vertebrats[1].

Src està compost per 533 aminoàcids (Src de pollastre, 536 per a l'humà c-Src) i s'expressa de manera ubiqua. No obstant això, el cervell, els osteoclasts, i les plaquetes expressen aquesta proteïna a nivells entre 5-200 vegades mes alts que la majoria d'altres cèl·lules[2].

Tots els SFKs comparteixen un ordre dels dominis conservat i constituït per (Figura IV.1): una regió N-terminal seguida per un domini únic intrínsecament desordenat amb seqüències diferents entre els membres de la família; els dominis SH3 i SH2; el domini SH1 o cinasa i una cua C-terminal. Un domini d'ancoratge a membrana anomenat SH4 que conté llocs de miristoilació i a vegades de palmitoilació, forma la regió N-terminal. El domini SH3 que interacciona amb motius rics en prolina que contenen la seqüencia consens PXXP[3]. El domini SH2 pot unir-se a motius de fosfotirosina. El domini cinasa (SH1) és responsable de l'activitat catalítica: aquest domini té un bucle d'activació (A-loop) amb una tirosina de regulació positiva (Y418) que quan esta fosforilada permet la màxima activitat de la cinasa. La cua C-terminal conté una tirosina de regulació negativa (Y527) que, quan es troba fosforilada, interacciona amb el domini SH2[4,5]. La interacció intramolecular entre el domini SH3 amb el motiu ric en prolina[6] situat en el connector entre els dominis SH2 i SH1[4] contribueix encara més

en bloquejar la proteïna en una forma enzimàticament inactiva tancada.

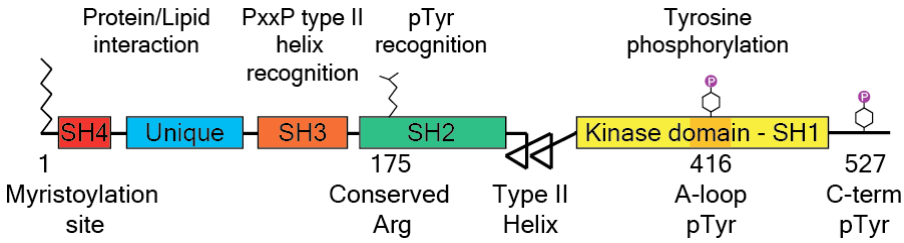


Figura IV.1. Estructura dels dominis de la família Src de cinases. L'arquitectura de la cinasa Src es compon de quatre dominis: la regió única, que varia entre els membres de la família, seguit pels dominis SH3, SH2 i el domini catalític SH1. El bucle d'activació (A-loop) del domini cinasa, els llocs de fosforilació (Tyr 416, Tyr 527), el residu conservat Arg 175 en el domini SH2 i l'hèlix de poliprolina de tipus II s'indiquen. Per convenció, els residus d'aminoàcids es numeren com en Src de pollastre.

Sent una proteïna citoplasmàtica, c-Src té un paper crític en la mediació de la transducció de senyals a través d'interaccions amb diferents proteïnes i complexos de proteïnes. Depenent de la seva localització cel·lular, c-Src fosforila diferents substrats en el citosol, a la cara interna de la membrana plasmàtica, a la matriu cel·lular o a les adherències cèl·lula-cèl·lula. La fosforilació de les tirosines pot directament afectar les funcions d'aquestes proteïnes. Alternativament, els residus fosforilats serveixen com llocs d'acoblament per a la unió de les proteïnes de senyalització que contenen dominis SH2. Aquests complexos de senyalització inicien vies que regulen la síntesi de proteïnes, l'expressió gènica, l'acoblament del citoesquelet i molts altres aspectes de la funció cel·lular, com la motilitat o la supervivència [2,7]. Alguns dels substrats regulats per c-Src es mostren a la Figura IV.2. Aquests inclouen la proteïna-cinasa Ras (Ras-MAPK), que regula l'expressió de *fos* través de la fosforilació del factor de transcripció Ets; la cinasa Akt (PI3K-Akt), que regula la iniciació de la traducció i la

supervivència cel·lular; i la via STAT3, que regula l'expressió de myc i de la ciclina D1[1,8].

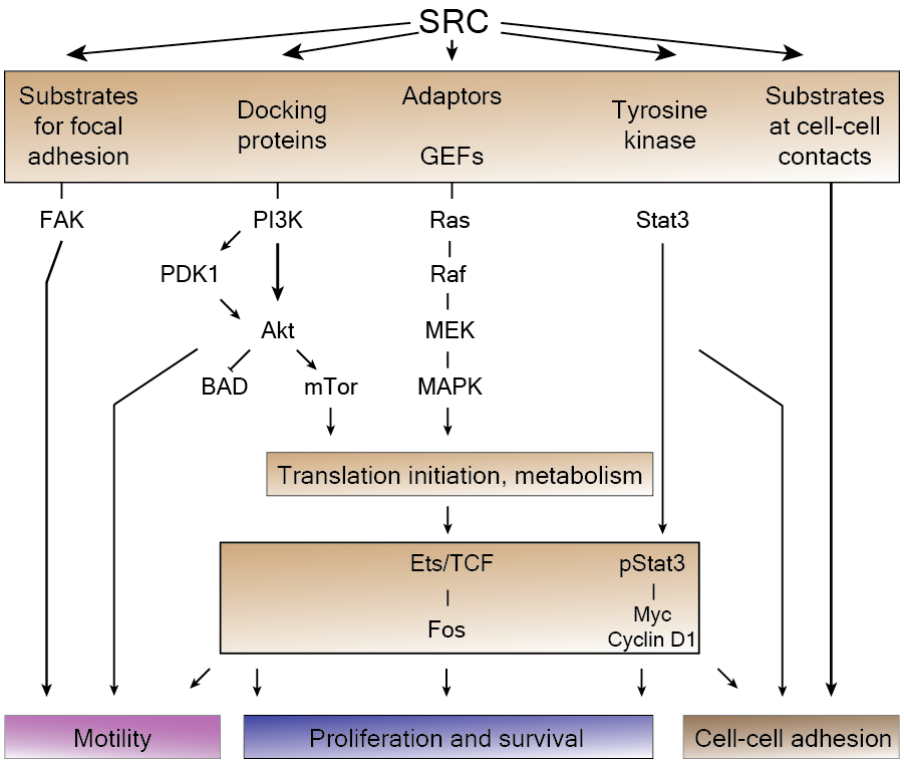


Figure IV.2. Vies de senyalització controlades per c-Src. Es mostren les vies de Ras (Ras-MAPK); la cinasa Akt (PI3K-Akt) via fosfatidilinositol-3; i la via de STAT3. (GEF, factors d'intercanvi de nucleòtids de guanina, mTOR, objectiu de rapamicina en mamífers, PDKs, cinases 3'-phosphoinositide-dependents.).

El domini únic de c-Src: un domini intrínsecament desordenat

Encara que c-Src va ser el primer oncogèn descobert, el paper de la seva regió N-terminal intrínsecament desordenat, anomenat el domini únic (USrc), ha romàs a les fosques durant molt de temps. Les estructures de raigs X obtingudes per c-Src[9,10] no contenen la regió N-terminal que impedeix la cristal·lització i és fàcilment degradat per les proteases[6]. Aquest fenomen està presumiblement relacionat amb el fet

que USrc és intrínsecament desordenat[11]. L'homologia de la seqüència dins de la família Src de cinases no receptores és bastant alt amb l'excepció de la regió N-terminal, on els únics aminoàcids clarament conservats són la Gly 2 i la Cys 3 que són llocs per a la miristoilació i palmitoilació, respectivament. No obstant això, el domini únic de cada membre individual de les SFKs està ben conservat entre els diferents organismes el que suggereix un paper més específic que el d'un senzill espaiador[12]. El domini SH4 i el domini únic de c-Src s'han relacionat amb la localització subcel·lular i transit[13]. La regió N-terminal de c-Src també es va demostrar que estava implicada en interaccions proteïna-proteïna[14,15,16]. Estudis recents han començat a donar pistes sobre possibles rols funcionals del domini únic de c-Src. En el 2009 Pérez et al.[11] van caracteritzar les propietats espectroscòpiques i l'espai conformacional mostrejat per USrc utilitzant tècniques de RMN. Es va identificar dins del fragment N-terminal intrínsecament desordenat una regió parcialment estructurada que comprèn els residus 60-75. A més, en el mateix estudi es van identificar in vitro esdeveniments de fosforilació que impliquen als residus Ser 17, Thr 37 i Ser 75. Més endavant el grup de Pons[17] va demostrar, per primera vegada, que c-Src humana conté dues regions addicionals d'unió a lípids: la regió d'unió a lípids del domini Únic (ULBR) i el domini SH3. La ULBR esta situada dins de la regió parcialment estructurada prèviament determinada per RMN. La capacitat de USrc d'interactuar amb els lípids es va demostrar que podia ser modulada in vitro per esdeveniments de fosforilació que involucren residus prèviament identificats (Ser 17, Thr 37 i Ser 75). El domini únic de c-Src humana esta també implicat en interaccions inter i intra-molecular. De fet, USrc interactua amb calmodulina en presència de calci, i també amb el domini SH3 veí. Aquesta interacció intra-molecular esta al·lostèricament inhibida per la unió al domini SH3 de c-Src d'un pèptid de poliprolina d'alta afinitat.

OBJECTIUS

Els objectius d'aquesta tesi són:

Objectiu 1 - Efecte de les mutacions que afecten la ULBR del domini únic de c-Src.

El primer objectiu d'aquesta tesi va ser la caracterització in vitro de l'efecte de les mutacions en la ULBR sobre les principals interaccions prèviament determinats del domini únic: la unió als lípids i el domini SH3.

Objectiu 2 - La fosforilació del domini únic de c-Src en oòcits de Xenopus laevis i extractes de cèl·lules de mamífers.

La fosforilació del domini únic de c-Src és un dels paràmetres de modulació que afecten la unió a lípids i anteriorment s'havia demostrat que podia afectar l'activitat de c-Src.

El segon objectiu va ser estudiar la fosforilació del domini únic aïllat en Xenopus laevis. Això va permetre l'estudi de la fosforilació en una cèl·lula viva i en extractes de cèl·lules i la seva posterior manipulació mitjançant modifiques de la xarxa de fosfatases / cinases.

Objectiu 3 - Els estudis funcionals del domini únic en el context de la proteïna completa.

L'últim objectiu va ser l'avaluació de la importància funcional de les modificacions investigats en el domini únic aïllat com objectius 1 i 2 en una línia cel·lular de càncer colorectal humà i en el context de la proteïna de longitud completa.

RESULTATS

Efecte de les mutacions en la ULBR del domini únic de c-Src

Estudis funcionals han demostrat que les mutacions que afecten la ULBR causen un fenotip aberrant en el procés de maduració dels oòcits de Xenopus laevis. Aquesta observació demostra que la ULBR juga un paper essencial en la regulació de c-Src. La ULBR i residus al seu voltant estan involucrats en almenys dos mecanismes importants: la interacció amb lípids la interacció inter-domini amb el SH3. S'ha demostrat que la capacitat d'unir lípids per la ULBR pot ser suprimida mitjançant la substitució dels residus ^{63}LFG^{65} per AAA (mutant AAA) o ^{63}LFGGFN^{68} per AAAEAE (mutant EAE) del domini únic de c-Src (Figura IV.3). Aquest efecte ha estat observat pel domini aïllat (USrc), així com per a la construcció lligada (USH3). No obstant això, la pèrdua de la interacció no afecta l'especificitat de lípids perquè els mutantes ULBR mostren el mateix patró de reconeixement de la forma de tipus salvatge.

D'altra banda, els residus que participen en la interacció domini Únic-SH3 encara mantenen aquesta interacció en el constructe USH3-AAA, demostrant que les mutacions en la ULBR no afecten a la interacció inter-domini amb el domini SH3. A mes, la interacció Únic-SH3 va ser confirmada pel fet que la modulació al·lostèrica amb un pèptid de poliprolina (VSL12), descrita anteriorment, en la interacció inter-domini encara es conserva.

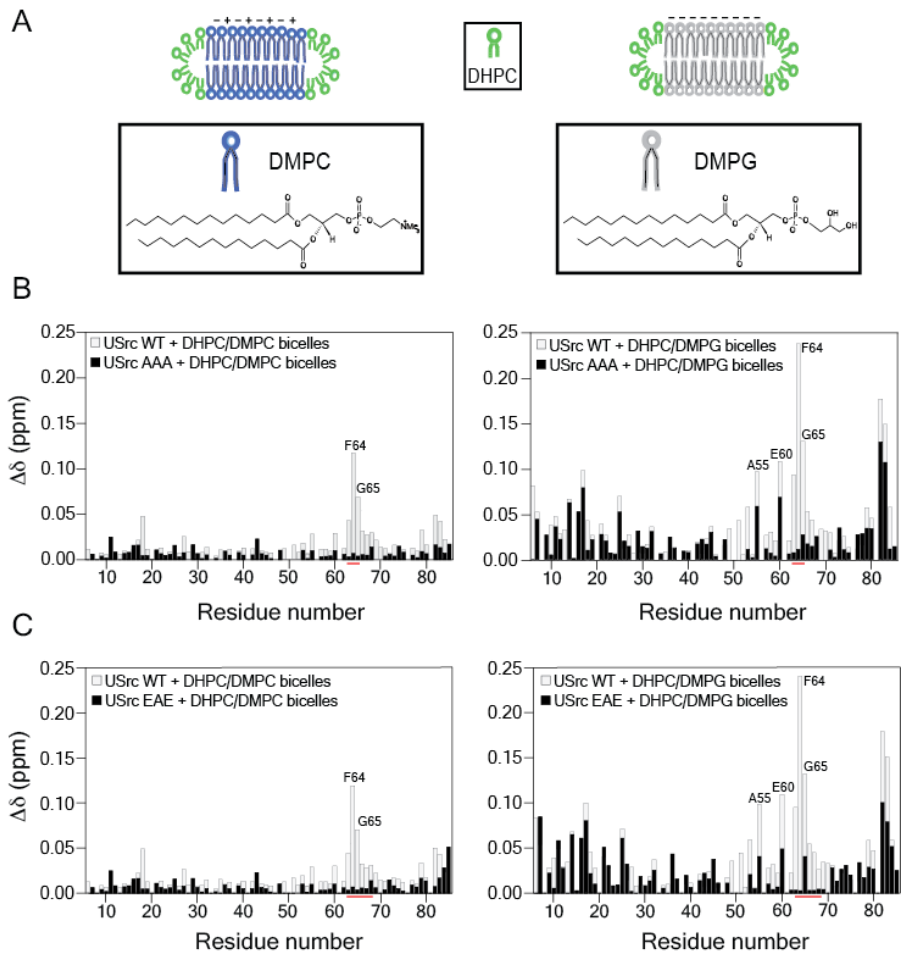


Figure IV.3. Interacció del domini únic de c-Src amb lipids. (A) Representació esquemàtica de l'estructura i la càrrega de les bicelas lípidiques utilitzades (DMPC, dimyristoyl phosphatidyl choline panell a la esquerre; DMPG, dimyristoyl phosphatidyl glycerol panell a la dret). (B-C) Valors absoluts combinats dels canvis químics ^{1}H-^{15}N induïts per la presencia de bicelas (DHPC / DMPC, panells de l'esquerra; DHPC / DMPG, panells de la dreta) en el mutant USrc AAA (B) i en el mutant USrc EAE (C). La concentració total de lípids va ser 8% w/v i la relació dels lípids de cadena llarga / DHPC (q) va ser 0,8. DHPC = dihexanoyl phosphatidylcholine.

La mutació AAA va ser dissenyada per eliminar parcialment el motiu FGGF present en la ULBR. Aquest motiu format per la seqüència FGG(F/V), seguit d'un tram de petits residus polars (N/D/S/T) és altament conservat en c-Src d'una sèrie d'espècies i es troba també en

el domini únic de Fyn i Yes, les dues cinases més similars a c-Src. El reemplaçament d'aquests aminoàcids provoca efectes locals importants que afecten a l'estructura de la ULBR. No obstant això, la modificació de l'estructura local no té cap efecte sobre les interaccions de llarg abast que tenen lloc en el domini únic de c-Src. Una d'aquestes interaccions és un contacte transitori amb residus de la regió SH4 observat per experiments de PREs (relaxació induïda per centres paramagnètics). L'estudi comparatiu del mutant AAA i de la forma salvatge mostren que la interacció ULBR – domini SH4 es conserva malgrat la mutació.

Per tant, hem estat capaços de discriminar entre les interaccions alternatives que impliquen el domini únic. De fet, la mutació AAA afecta principalment a la capacitat de la ULBR d'interaccionar amb lípids, mentre que no s'han observat efectes sobre les interaccions interdomini amb els dominis SH3 i SH4. La ULBR representa una segona regió d'unió de lípids dins del domini únic que podria cooperar amb el domini SH4 per tal de localitzar correctament la proteïna a la superfície de la membrana ("regulació posicional"). La posició relativa de c-Src a la membrana podria ser essencial per a la regulació de les diferents vies de senyalització que són controlades per c-Src. Suprimint la possibilitat de la ULBR de interaccionar amb lípids estem modulant el mecanisme amb el qual c-Src efectua sobre els diferents substrats a nivell de la membrana.

Una tercera regió important per l'unió a lípids està present en c-Src: el domini SH3. Els nostres resultats han demostrat que el domini SH3 de c-Src perd la seva capacitat d'interactuar amb lípids en presència d'un pèptid poliprolina (Figura IV.4). Aquesta observació suggereix que un mecanisme de regulació que implica la unió a lípids pot també modular el domini SH3. Els residus implicats en la unió a lípids es troben a la cara oposada del domini SH3 on és la escletxa que reconeix els motius de poliprolina.

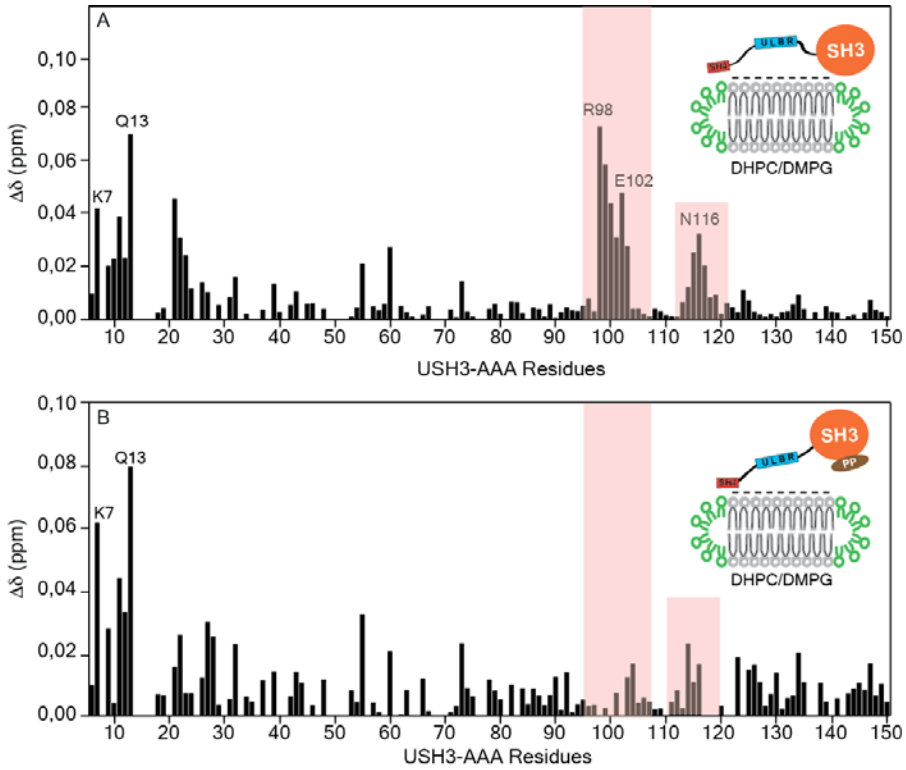


Figure IV.4. La interacció domini SH3-lípids és impedit per la unió d'un pèptid de poliprolina el domini SH3. (A) Valors absoluts combinats dels canvis químics ^{1}H-^{15}N induïts per la presencia de bicelas DHPC / DMPG (8% w/v de concentració total de lípids, q = 0,8) en el constructe USH3-AAA. (B) Valors absoluts combinats dels canvis químics ^{1}H-^{15}N entre USH3-AAA en presència del pèptid VSL12 (proporció de 1:1 molar) i en presència del pèptid VSL12 (proporció de 1: 1 molar) més bicelas DHPC / DMPG (8% w/v concentració total de lípids, q = 0,8). Tots els espectres es van misurar a 298 K i en un espectròmetre de 600 MHz. Les regions colorades en vermell representen els residus SH3 pertorbats en presència de bicelas.

Estudi del domini únic de c-Src en el sistema model de Xenopus laevis

Les modificacions post-traduccionals (PTMs) són crucials perquè les cèl·lules siguin capaços de regular ràpidament la funció de les proteïnes. En particular, la fosforilació, que es un procés reversible, constitueix un esdeveniment regulador molt important en el control de les vies de senyalització. La fosforilació de residus de serina i treonina és la més abundant en l'interior de les cèl·lules. Comunment, la mateixa proteïna pot contenir múltiples llocs de fosforilació que poden regular la seva funció o la seva capacitat per interactuar amb diferents partners[18].

El domini únic intrínsecament desordenat de c-Src conté diversos llocs susceptibles de fosforilació i desfosforilació.

Experiments de RMN en cèl·lules i mesurant en temps han permès l'estudi dels complexos processos de fosforilació / desfosforilació en el domini únic intrínsecament desordenat de c-Src humana mediats per cinases i fosfatases que estan presents en els oòcits de Xenopus laevis i extractes de cèl·lules. En primer lloc, hem observat que USrc es fosforila en la Ser 17 en oòcits vius de Xenopus. Resultats similars es van observar quan es va afegir USrc marcada amb ^{15}N als extractes que contenen la fracció soluble de lisats de oòcits en la etapa VI (Figura IV.5). La proteïna cinasa A (PKA) va ser identificada com el principal candidat per catalitzar aquesta reacció. De fet, la Ser 17 pot ser fosforilada in vitro per PKA[11], i aquesta cinasa és coneguda per ser activa en els oòcits de Xenopus[19]. La majoria de les ressonàncies de USrc que envolten la Ser 17 van ser significativament pertorbat en els extractes d'oòcits. El residu Glu 22 i altres residus carregats a prop de Ser 17, també va mostrar canvis de desplaçament químic per sobre de la desviació estàndard (0,13 ppm) de les variacions observades.

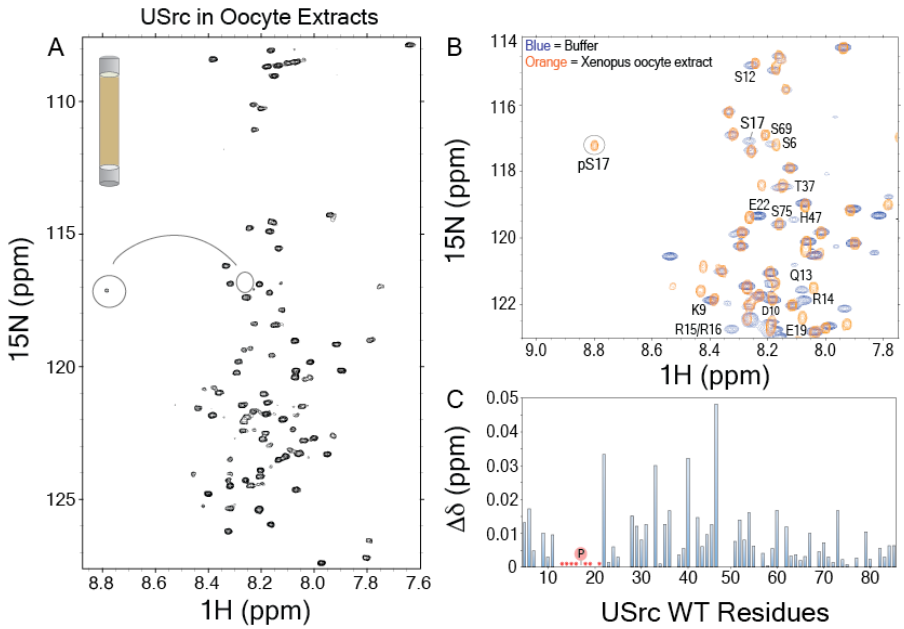


Figure IV.5. Fosforilació de USrc en lisats d'oòcits d'etapa VI. (A) ^{1}H-^{15}N espectra de RMN de USRC en extractes citoplasmàtic d'ovòcits. El cercle marca la posició de la fosfo Ser 17. La línia corba indica el moviment cap a l'esquerra des de la Ser 17 no modificada (oval) a la fosforilada. (B) Superposició de les expansions de ^{1}H-^{15}N espectres HSQC de USRC en tampó (blau) i en extracte de oòcits de Xenopus (taronja). El cercle marca la posició de pSer 17. (C) Els canvis de desplaçament químic entre les dues condicions (tampó vs. extracte) es representen. Els asteriscs vermells indiquen residus amb gran pertorbació (que impedeixin la seva inequívoca assignació en la mostra fosforilada). El lloc de fosforilació s'indica mitjançant un P.

A continuació vam estudiat USrc en els ous no fertilitzats de Xenopus laevis (extracte CSF) on vam estat capaços d'identificar 3 llocs de fosforilació diferents que van ser inequívocament assignats a la Ser 17, Ser 69 i Ser 75. Els principals candidats capaços de fosforilar USrc en els residus Ser 69 i Ser 75 van ser les CDKs (cinases dependents de ciclines). Es va prèviament demostrar que USrc pot ser fosforilada in vitro per CDK5/p25 en la Ser 75 i la Thr 37[11]. No obstant això, la fosforilació de Thr 37 no va ser detectada per espectrometria de masses després de la incubació de USrc amb extractes d'ou de

Xenopus. En contrast, l'espectrometria de masses (MS) va detectar de manera inequívoca Ser 69 fosforilada. La fosforilació del residu de Ser 69 de c-Src s'havia prèviament detectat només per MS en extractes cel·lulars de línies de càncer de HCT116 i MDA-MB-435S[20], encara que la seva funció és actualment desconeguda. L'addició de les CDKs o de PKA purificada, així com d'inhibidors de la cinasa PKA, va mostrar clarament la interrelació entre les activitats de les cinases i fosfatases en els extractes CSF. Així que, canviant les condicions experimentals de l'entorn cel·lular, hem estat capaços de descobrir un escenari molt complex d'esdeveniments de multi-fosforilació de USRC impulsats per cinases endògenes.

S'havia prèviament demostrat que el domini únic conté una regió d'unió de lípids-secundària, anomenada ULBR. L'observació dels espectres en ovòcit vius, així com en els diferents extractes suggereix que les interaccions que impliquen aquesta regió secundària són febles, almenys en absència de miristoilació.

Finalment hem demostrat la importància biològica de les interaccions que impliquen la ULBR en el context de la proteïna completa de c-Src, mitjançant la comparació dels efectes induïts per la proteïna de tipus salvatge i pels mutants ULBR (mutant AAA, mutant EAE) en la maduració d'oòcits de Xenopus laevis. La injecció d'ARNm que codifica una forma constitutivament activa de c-Src humana va causar una acceleració de la maduració dels ovòcits induïda per progesterona (Figura IV.6), que és molt similar a l'efecte ja observat per la forma constitutivament activa de c-Src de Xenopus o Src viral[21,22]. Aquest resultat valida el model heteròleg en el qual treballem. No obstant això, els ARNm que codifiquen per als mutantes ULBR causen que una part important (20% -25%) dels oòcits no acabin de madurar després del tractament amb progesterona. D'altra banda, la meitat dels ovòcits injectats amb ARNm dels mutants que van madurar, van mostrar símptomes d'apoptosi i despigmentació visible. Així, mentre que la

proteïna de tipus salvatge i les formes mutantes de c-Src humana són igualment actius en la iniciació de la maduració d'ovòcits, el fenotip letal només s'observa en oòcits madurats que expressen la c-Src mutada. Això demostra que la ULBR juga un paper essencial en la regulació de c-Src. Aquests resultats constitueixen una prova fonamental de la importància funcional de la ULBR i donen suport a la idea que les interaccions observades són de fet part d'un nou nivell de regulació de c-Src.

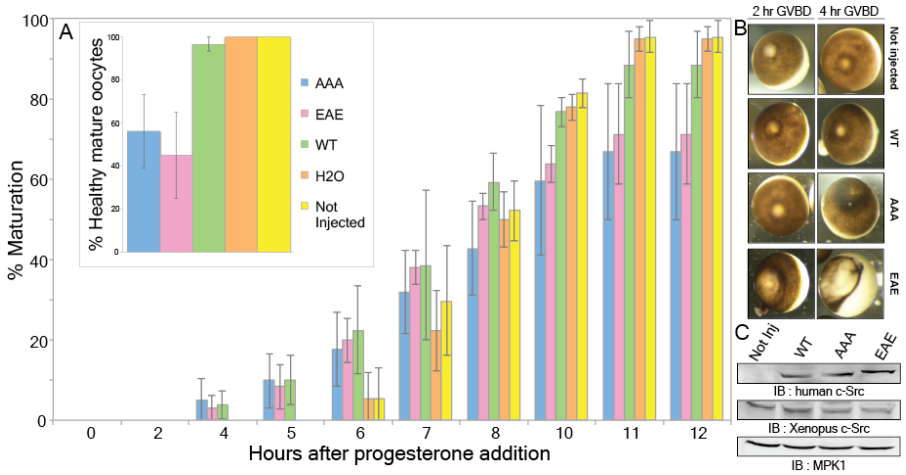


Figure IV.6. Efecte dels mutantes de c-Src en la maduració d'ovòcits de Xenopus. L'efecte de la injecció en oòcits amb ARNm que codifiquen per c-Src de tipus salvatge (verd), mutant AAA (blau), mutant EAE (magenta), o aigua (taronja) es compara amb la dels oòcits no injectats (groc). (A) Percentatge d'oòcits que va acabar al procés de maduració (GVBD), en diferents moments després de l'addició de la progesterona. El requadre mostra els percentatges d'oòcits morfològicament normals al final de l'experiment. (B) Aparença dels oòcits 2 i 4 h després de la GVBD. Aproximadament la meitat dels oòcits madurats que expressen mutantes AAA o EAE mostren despigmentació progressiva a partir de 2 hores després de GVBD. (C) Anàlisi dels nivells d'expressió de c-Src. Lisats d'oòcits van ser separades per SDS-PAGE i es van analitzar per immunotransferència amb diferents anticossos: anti- c-Src humà (panell superior), anti-c-Src de Xenopus (panell central) i anti-MPK1 (panell inferior) com a control de càrrega.

Hem demostrat que el domini únic de c-Src humana es pot fosforilar en serina 17, serina 69 i serina 75 també en diversos sistemes de cèl·lules de mamífers. Resulta interessant que els patrons de fosforilació varien entre els diferents tipus de cèl·lules. Els canvis mes importants es van observar per a la Ser 69 i Ser 75 mentre la pSer17 es va trobar en totes les condicions. En les cèl·lules COS-7 es va detectar només pSer 75. Aquest resultat està d'acord amb les dades publicades anteriorment en la que Pa et al. van demostrar que la fosforilació de c-Src a la Ser 75 per CDK5 es produeix en cèl·lules COS-7[23]. En les cèl·lules MEFs la fosforilació de la Ser 75 o Ser 69 només va ser detectada en la presència d'inhibidors de fosfatases. Sota aquestes condicions, la fosforilació de Ser 69 va ser més gran que la de la Ser 75. En cèl·lules de càncer HeLa la fosforilació de Ser 69 i Ser 75 es va observar fins i tot en absència d'inhibidors de la fosfatasa. Sota aquestes condicions, els nivells de fosforilació dels dos llocs van ser similars. En presència d'inhibidors, la fosforilació de Ser 75 va augmentar aproximadament 3 vegades mentre que la de Ser 69 va romandre similar. Per tant, les nostres dades confirmen que la Ser 69 pot ser substancialment fosforilat, i especialment en línies cel·lulars de càncer.

Estudis funcionals del domini únic de c-Src en el context de la proteïna de longitud completa

Els rols funcionals de les fosforilacions descrites juntament amb les mutacions de la ULBR s'han investigat en cèl·lules HEK293T i en la línia cel·lular humà de càncer colorectal (CRC) SW620 en el context de la proteïna de longitud completa. Per a simular la fosforilació constitutiva es va utilitzar la mutació dels residus de serina per àcid glutàmic.

Primer, hem mostrat que els mutant S75E i AAA tenen una activitat i una estabilitat mes baixa respecte a c-Src de tipus salvatge en

cèl·lules HEK293T. La menor d'expressió s'ha relacionat amb un fenomen que implica la degradació mitjançant el sistema ubiquitina-proteasoma. De fet, el tractament de cèl·lules amb l'inhibidor del proteasoma MG132 va conduir a una recuperació completa de l'expressió del mutant S75E i, com a conseqüència, de la seva activitat. Pa et al. ja havien demostrat que la fosforilació de c-Src a la Ser 75 per CDK5 és un mecanisme crucial per a la regulació de l'activitat intracel·lular de c-Src, ja que està directament implicat en la modulació de la degradació ubiqüitina-depenent de c-Src actiu[23]. El nivell d'expressió del mutant AAA també va ser restaurat després del tractament amb MG132. Tanmateix però, la seva activitat total es va mantenir més baixa que la de la WT. Les nostres dades suggereixen que aquest mutant és capaç d'activar-se (per autofosforilación en la Y418) però mostra nivells alterats d'activitat cap a altres substrats.

Després, hem investigat l'efecte de les mutacions de USrc en el sistema de cèl·lules SW620 de càncer colorectal. Els mutants de USrc, així com el tipus salvatge indueixen un fort canvi morfològic en les cèl·lules CRC. S'ha observat l'aparició de múltiples protuberàncies que brollen de la membrana plasmàtica i aquest fenomen es relaciona amb un augment de la mobilitat. En aquest sistema no es van detectar canvis significatius en l'activitat de la proteïna entre mutants i WT. Una observació interessant és que el mutant S75E no va ser afectat per fenòmens de degradació en les cèl·lules SW620 el que suggereix que aquest sistema presenta també una forta desregulació de la maquinària ubiquitina-proteasoma.

La sobre-expressió de c-Src indueix alta capacitat d'invassió de cèl·lules de càncer avançat in vitro, com s'havia descriu anteriorment per Leroy et al[24]. Hem demostrat que les mutacions de la ULBR afecten fortament la capacitat d'envair de les cèl·lules SW620 CRC. De fet, els mutants ULBR mostren una invasivitat reduïda (>20-50%) respecte al WT (Figura IV.7). Els efectes observats poden ser

atribuïbles al fet que els mutants fosforilen diferentment part del clúster de substrats importants per a la invasivitat d'aquestes cèl·lules. Aquesta hipòtesi es veu reforçada pel fet que, mirant el contingut cel·lular de fosfotirosines, diverses proteïnes es va mostrar tenir diferents patrons de fosforilació entre les diferents condicions. Els experiments de fosfoproteòmica semi-quantitativa van confirmar encara més l'anàlisi preliminar. De fet, moltíssimes proteïnes es van detectar fosforilades de manera diferent pels mutants ULBR.

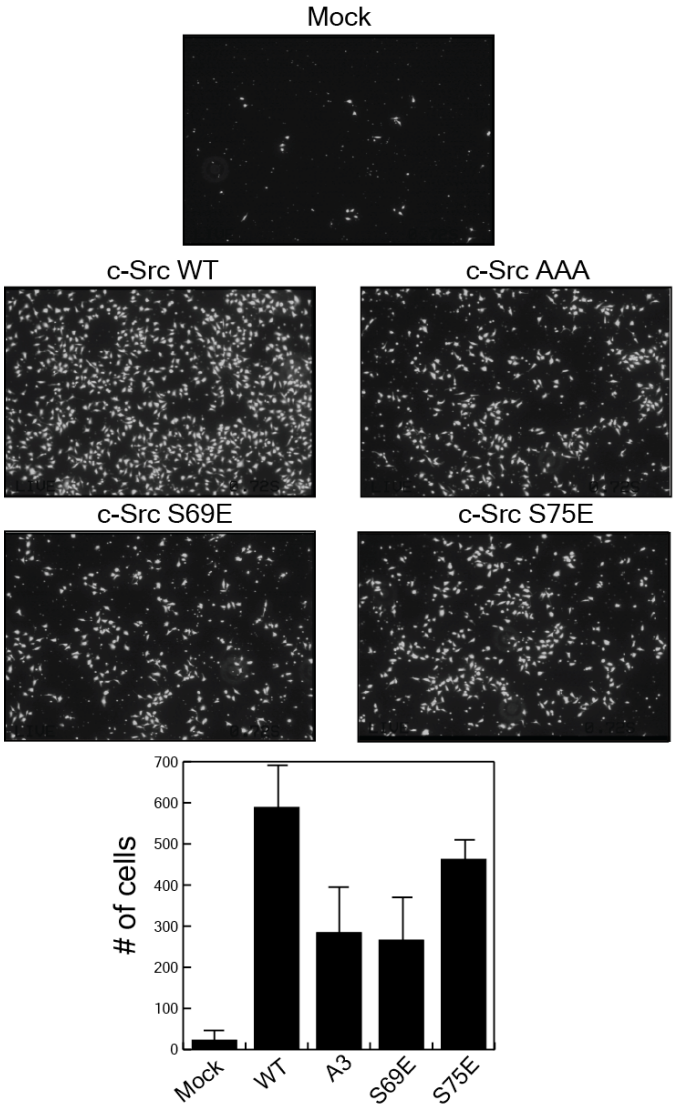


Figure IV.7. Assaig d'invasió cel·lular en CRC SW620 que expressen el WT i els mutants del domini únic de c-Src. c-Src promou la invasió cel·lular a través de Matrigel de les cèl·lules SW620 in vitro. Mutants del domini ùnic (AAA, S69E i S75E) mostrn una invasivitat reduïda en comparació amb el WT. El valor *X* del diagrama es refereix a la quantificació del nombre de cèl·lules que han envaït a través de Matrigel en assajos amb cambres de Boyden. Les columnes representen la mitjana de 3 experiments independents. Les barres d'error la desviació estàndard.

Tots plegats, aquestes dades indiquen clarament que el domini únic de c-Src té un paper regulador que afecta l'activitat global de la cinasa c-Src i també la seva selectivitat cap a un conjunt diferent de substrats. Moltes de les proteïnes identificades que mostren diferents patrons de fosforilació entre les condicions analitzades van ser modificades per c-Src a nivell de la membrana plasmàtica. c-Src s'uneix a la superfície de la membrana pel seu domini SH4 miristoilat i la ULBR constitueix una segona regió d'unió de lípids. Anteriorment, observacions in vitro havien mostrat que la fosforilació de llocs pròxims a la ULBR modula la seva capacitat d'unió a lípids[25]. Suggerim que, almenys alguns dels efectes de regulació selectiva de c-Src poden ser controlats per la longitud efectiva abastada pel domini únic. Així, modificant la unió a lípids pel ULBR es pot limitar la distància que el domini quinasa o els dominis que ajuden a reconèixer els substrats, poden allunyar-se de la superfície de la membrana. La longitud efectiva del domini únic introdueix així un cert grau de compartimentació en la direcció perpendicular al pla de la membrana que podria prevenir o modificar la interacció entre c-Src i altres proteïnes, fins i tot si estan units a la membrana un al costat de l'altre. Per tant, la modulació de la interacció ULBR-lípids, per exemple, mitjançant la fosforilació com en el cas del mutant S75E, proporciona un mecanisme per canviar la longitud efectiva del domini únic que connecta els dominis plegats de c-Src a la membrana, i per tant, l'especificitat de c-Src pels substrats.

La presència d'aquest mecanisme de regulació addicional ("model de regulació posicional", Figura IV.8) ajudarà a entendre com c-Src connecta amb eficàcia una àmplia varietat de substrats i pot aclarir per primera vegada el paper biològic del domini únic.

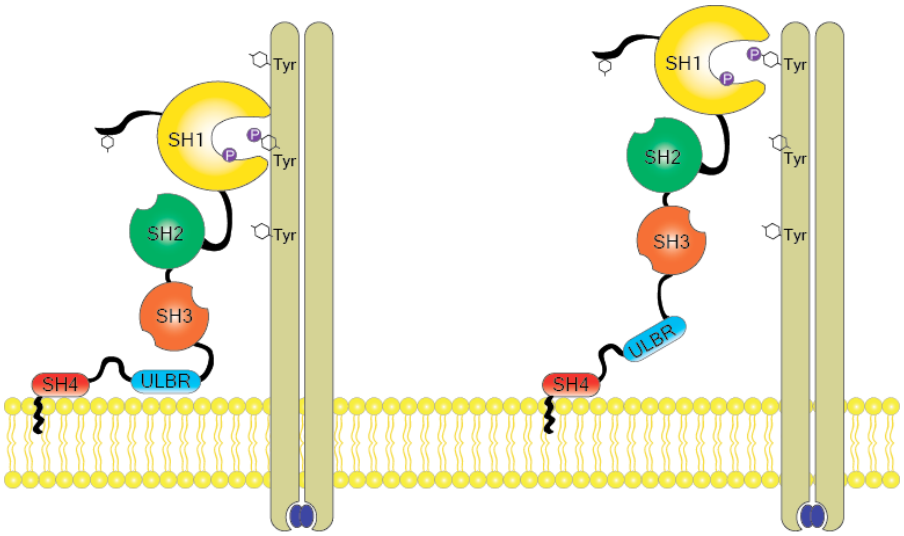


Figure IV.8. "Model de regulació posicional" de c-Src. A la membrana plasmàtica de c-Src interactua amb diversos substrats que poden ser fosforilats. El domini únic pot modular la longitud efectiva de c-Src en la direcció perpendicular al pla de la membrana a través de la regulació de la unió a lípids (per exemple, canvi de composició de lípids; fosforilació en Ser 75). Aquest mecanisme regulador addicional permet que c-Src pot canviar la seva especificitat, i per tant, la capacitat de controlar les diverses vies de senyalització en la qual està involucrat.

CONCLUSIONS

Les conclusions d'aquestes tesis estan resumides aquí:

• Les mutacions en la ULBR afecten la capacitat d'unió de lípids de USrc però no la seva interacció amb el dominis SH3.
• El domini SH3 de c-Src humana perd la seva capacitat d'interactuar amb els lípids en presència d'un pèptid poliprolina.
• La RMN en temps real permet l'estudi dels complexos processos de fosforilació i desfosforilació en USrc modulats per quinases i fosfatases en el sistema de Xenopus laevis.
• Tres diferents llocs de fosforilació presents en el domini únic, que van ser inequívocament assignats a Ser 17, Ser 69 i Ser 75, s'han identificat a Xenopus laevis, COS-7, MEFs i HeLa.
• Va ser la primera vegada que es va observar la fosforilació de la Ser 69 de c-Src en aquests sistemes model.
• La importància biològica de la ULBR s'ha demostrat primer en Xenopus laevis en el context de la proteïna de longitud completa.
• Els rols funcionals de les fosforilació descrites juntament amb les mutacions de la ULBR s'han demostrat en cèl·lules HEK293T i en la línia cel·lular de càncer colorectal SW620 en el context de la proteïna de longitud completa.
• Un "model de regulació posicional" ha estat proposat com un nou mecanisme de regulació de la selectivitat de c-Src.

AGRAIMENTS

Desitjo agrair a tots els col·laboradors, companys, amics, familiars i contactes per l'ajuda que m'han donat en realitzar aquesta tesi.

REFERÈNCIES

1. Martin, G. S. The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2,** 467–75 (2001).

2. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287,** 121–149 (1996).

3. Koch, C., Anderson, D., Moran, D., Ellis, V. & Pawson, T. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins, *Science* **252**, 668–674 (1991).

4. Xu, W., Doshi, a, Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3,** 629–38 (1999).

5. Cowan-Jacob, S. W. *et al.* The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13,** 861–71 (2005).

6. Bernadó, P., Pérez, Y., Svergun, D. I. & Pons, M. Structural characterization of the active and inactive states of Src kinase in solution by small-angle X-ray scattering. *J. Mol. Biol.* **376,** 492–505 (2008).

7. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13,** 513–609 (1997).

8. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4,** 470–80 (2004).

9. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385,** 595–602 (1997).

10. Sicheri, F. & Kuriyan, J. Structures of Src-family tyrosine kinases. *Curr. Opin. Struct. Biol.* **7,** 777–85 (1997).

11. Pérez, Y., Gairí, M., Pons, M. & Bernadó, P. Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. *J. Mol. Biol.* **391,** 136–48 (2009).

12. Amata, I., Maffei, M. & Pons, M. Phosphorylation of unique domains of Src family kinases. *Front. Genet.* **5,** 1–6 (2014).

13. Kasahara, K. *et al.* Rapid trafficking of c-Src, a non-palmitoylated Src-family kinase, between the plasma membrane and late endosomes/lysosomes. *Exp. Cell Res.* **313,** 2651–66 (2007).

14. Resh, M. D. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta* **1451,** 1–16 (1999).

15. Lee, H. *et al.* Palmitoylation of caveolin-1 at a single site (Cys-156) controls its coupling to the c-Src tyrosine kinase: targeting of dually acylated molecules (GPI-linked, transmembrane, or cytoplasmic) to caveolae effectively uncouples c-Src and caveolin-1 (TYR-14). *J. Biol. Chem.* **276,** 35150–8 (2001).

16. Gingrich, J. R. *et al.* Unique domain anchoring of Src to synaptic NMDA receptors via the mitochondrial protein NADH dehydrogenase subunit 2. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6237–6242 (2004).

17. Pérez, Y. et al. Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. Sci. Rep. 3, 1295 (2013).

18. Selenko, P. et al. In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. Nat. Struct. Mol. Biol. 15, 321–9 (2008).

19. Schmitt, A. & Nebreda, A. R. Signalling pathways in oocyte meiotic maturation. *J. Cell Sci.* **115,** 2457–9 (2002).

20. Oppermanna, F. S. *et al.* Large-scale proteomics analysis of the human kinome. *Mol. Cell. Proteomics* **8,** 1751–64 (2009).

21. Unger, T. F. & Steele, R. E. Biochemical and Cytological changes associated with Expression of deregulated pp60src In Xenopus Oocytes. Mol. Cell. Biol. 12, 5485–5498 (1992).

22. Tokmakov, A. et al. Regulation of Src kinase activity during Xenopus oocyte maturation. Dev. Biol. 278, 289–300 (2005).

23. Pan, Q. et al. Cdk5 targets active Src for ubiquitin-dependent degradation by phosphorylating Src(S75). Cell. Mol. Life Sci. 68, 3425–36 (2011).

24. Leroy, C. et al. Quantitative phosphoproteomics reveals a cluster of tyrosine kinases that mediates SRC invasive activity in advanced colon carcinoma cells. Cancer Res. 69, 2279–86 (2009).

25. Pérez, Y. et al. Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. Sci. Rep. 3, 1295 (2013).

Part V

Addenda

A.1 – Supporting Figures

II.1 - ULBR mutations modulate lipid binding by the Unique domain

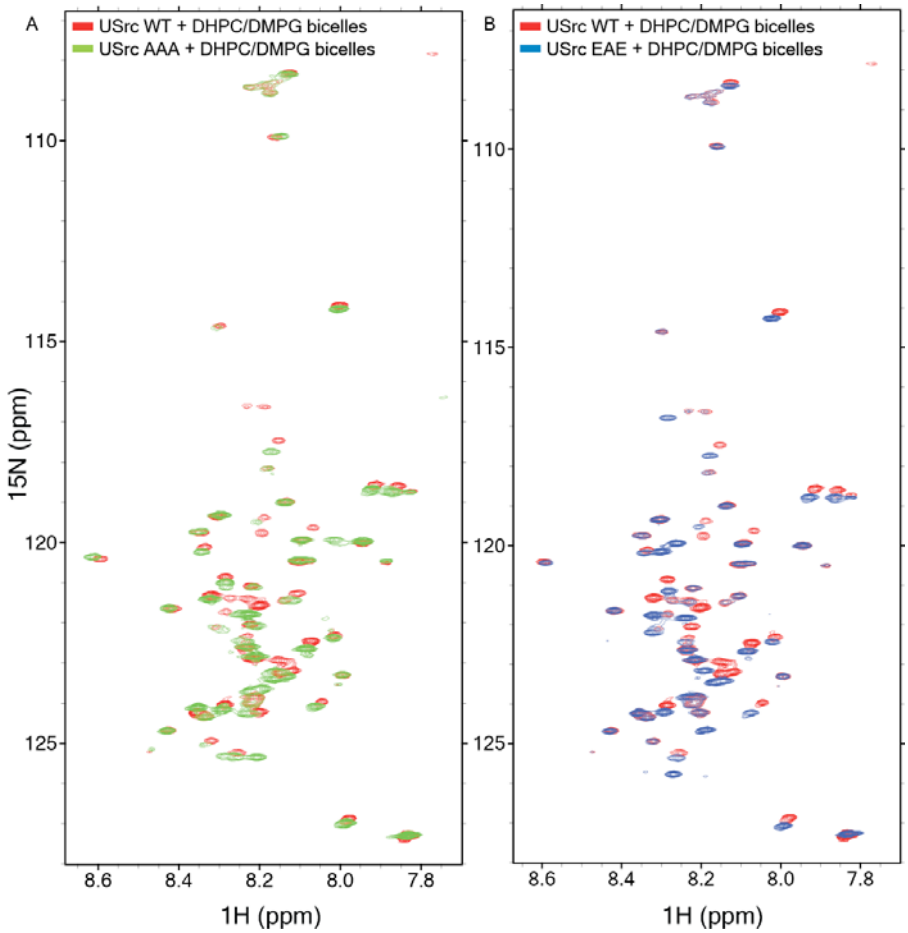


Figure S1. Lipid binding by the Unique domain of c-Src. (A) Overlay of ^{1}H-^{15}N HSQC spectra of USrc WT (red) and USrc-AAA (green) in presence of DHPC/DMPG bicelles. (B) Overlay of ^{1}H-^{15}N HSQC spectra of USrc WT (red) and USrc-EAE (blue) in presence of DHPC/DMPG bicelles. All spectra were recorded at 298 K and 600 MHz spectrometer.

II.1 - The SH3 domain capability of binding lipids is modulated by the interaction with a polyproline peptide

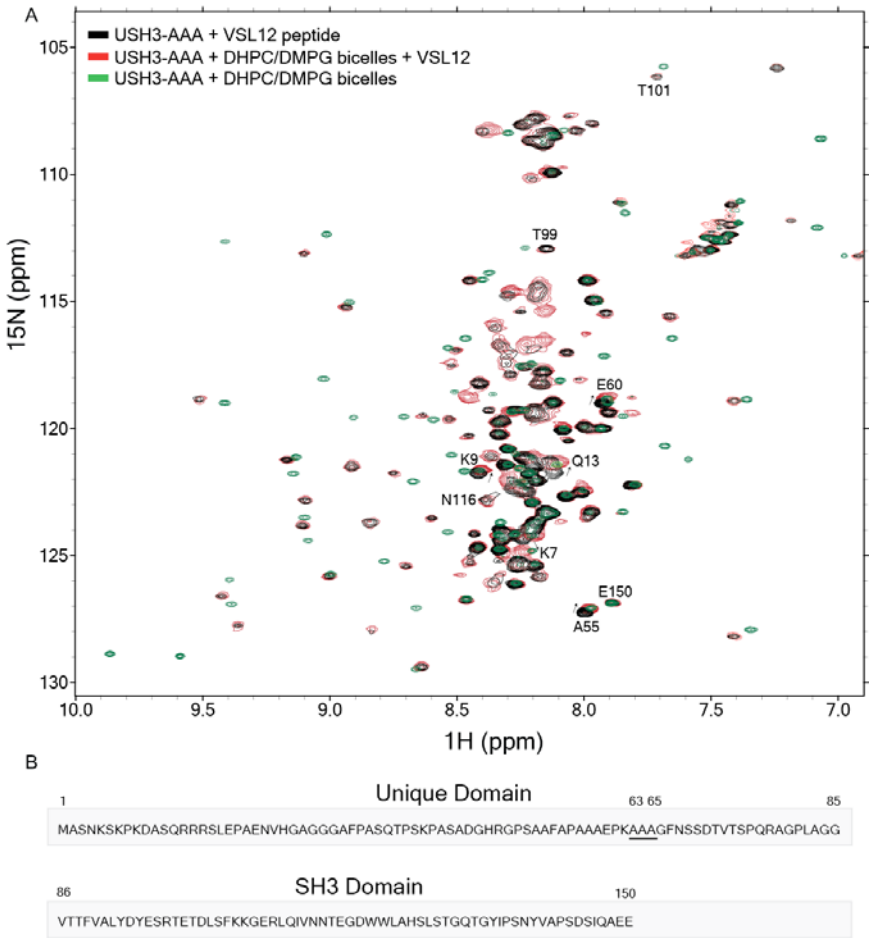


Figure S2. SH3 - lipid interaction is prevented by the binding of a polyproline peptide to the SH3 domain. (A) Overlay of ^{1}H-^{15}N HSQC spectra of human USH3-AAA construct + VSL12 peptide (1:1 molar ratio) (black), USH3-AAA construct + DHPC/DMPG + VSL12 peptide (1:1 molar ratio) (red) and USH3-AAA + DHPC/DMPG bicelles (green) measured at 298 K. SH3 residues (T99, T101, N116) normally perturbed by lipids did not interact with bicelles in presence of a polyproline peptide. Residues belonging to the SH4 domain (K7, K9, Q13) were affected by lipids even in presence of the VSL12 peptide. (B) Aminocids sequence of the USH3-AAA construct.

II.2 - Identification of USrc phosphorylation sites and active kinases in Xenopus egg extracts

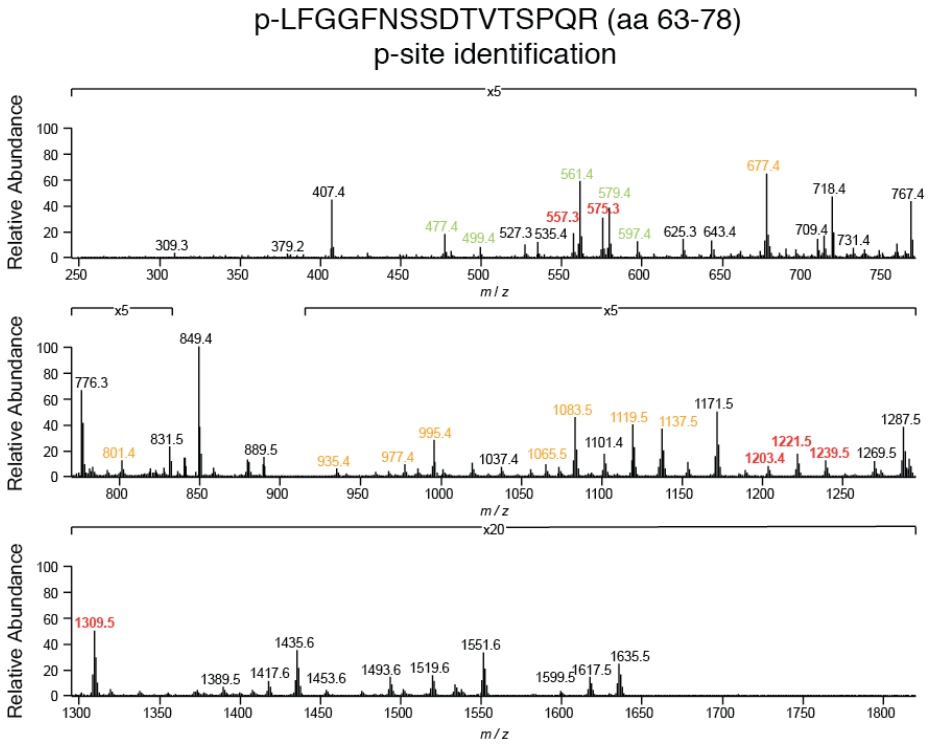


Figure S3. Phosphosite identification by mass-spectrometry. MS/MS spectrum of m/z 907.37 (^{15}N-labeled monophosphorylated p-LFGGFNSSDTVTSPQR $(M+2H)^{2+}$) and fragment peak assignment. Unique peaks that exclusively belong to p-site Ser 75 are marked in red and to Ser 69 in dark green. Orange and light green colors indicate fragments matching Ser 75 or Ser 69, but being not unique.

II.2 - Discussion: Phosphorylation of Ser 69

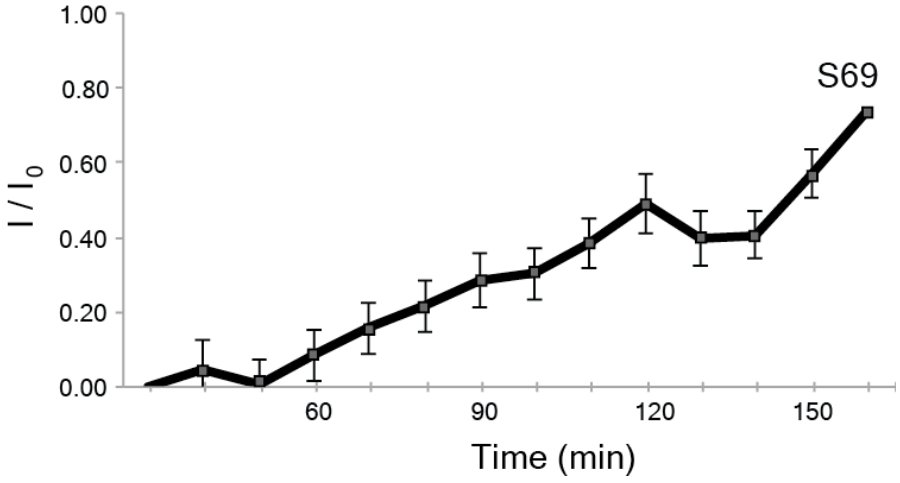
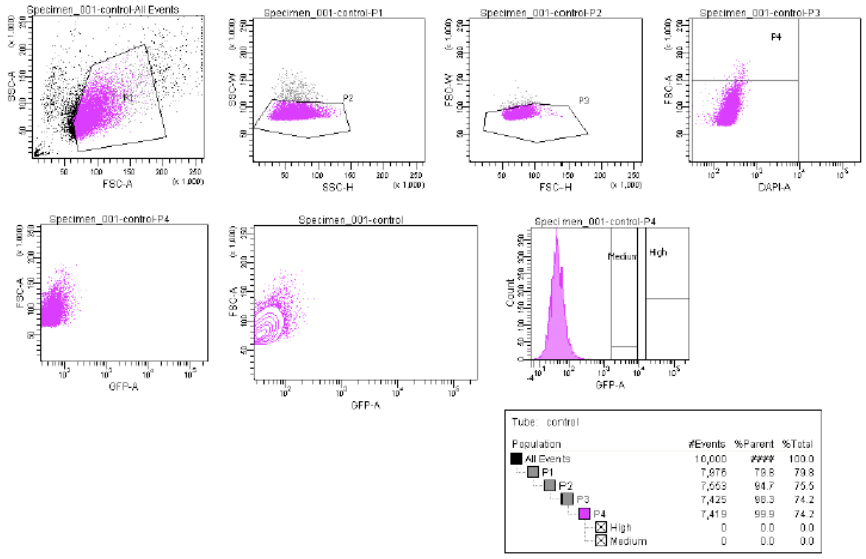


Figure S4. Time evolution of unphosphorylated USrc S69 in *Xenopus* egg extracts (CSF) followed by NMR. We interpret the intensity increase of Ser 69 signal as the result of the phosphorylation in other sites, Ser 75 and or Ser 17 causing a decrease in interactions with lipids of the Unique Lipid Binding Region, next to Ser 69. Intensities are normalized with respect to the maximum intensity of peak 1 (I_0) (see Figure II.2.6, Chapter II.2). Intensities were measured as running averages of two SOFAST-HMQC experiments of 5 minutes each. Error bars show standard deviations propagated from the variations of peaks unaffected by phosphorylation between experiments. Protein concentration was 50 μM.

II.3 - Creation of SW620 stable cell lines expressing c-Src Unique mutants



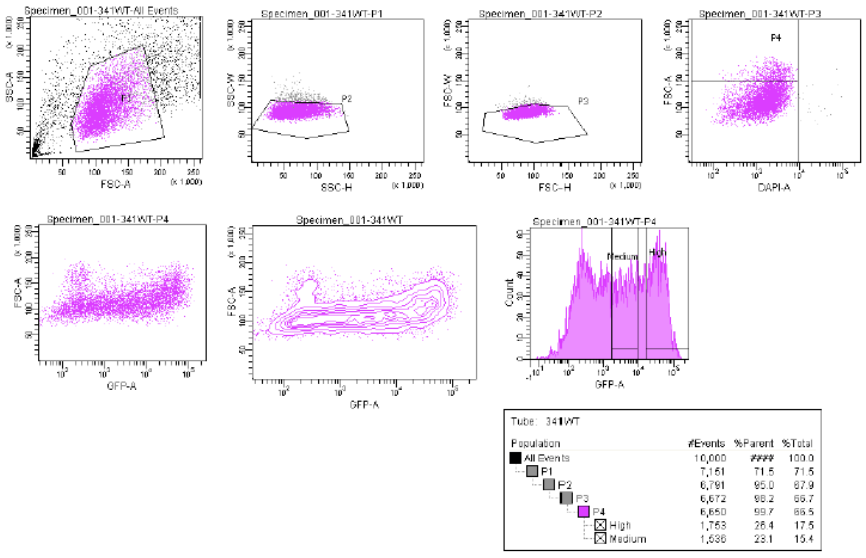


Figure S5. Fluorescent-activated cell sorting (FACS) of SW620 cells. Cells infected with empty vector (mock, top panels) and c-Src WT (bottom panels). The two selected population (medium GFP-expressing and high GFP-expressing) are shown in the last plot of WT panels. FACSs of Unique mutants are not shown.

II.3 - Src phosphosphoproteomics in CRC SW620 cells

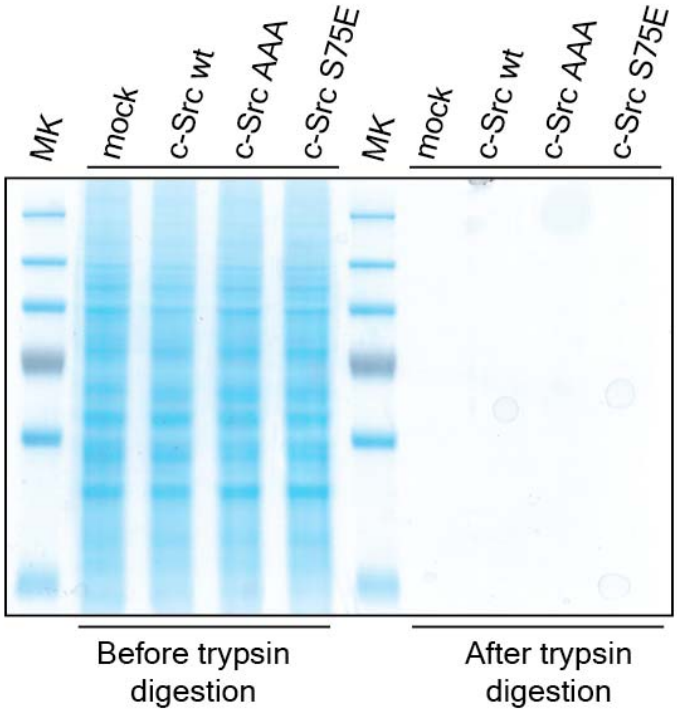


Figure S6. SDS-PAGE of cell lysates. Whole lysates were quantified and analysed on acrylamide gels before and after trypsin digestion. The loaded amount of proteins was identical for all conditions confirming that lysates were correctly quantified (left panel). Validation of digestion was verified by the disappearance of protein bands (right panel).

A.2 – Supporting Tables

II.2 – Phosphorylation of Serine 69

Digestion with trypsin and chymotrypsin

sequence	USrc WT	USrc S69E
A.AFAPAAAEPK.L		
A.FAPAAAEPK.L	1	2
A.GGGAFPASQTPSKPASADGHR.G		
A.GGSAWSHPQFE.K		1
A.GGSAWSHPQFEK.-		1
A.GPLAGGSAWSHPQFEK.-		
A.WSHPQFE.K		
A.WSHPQFEK.-		
D.GHRGPSAAFAPAAAEPK.L		
F.GGFNESDTVTSPQR.A		
F.GGFNSSDTVTSPQR.A		
G.FNSSDTVTSPQR.A		
G.GAFPASQTPSKPASADGHR.G		
G.GFNESDTVTSPQR.A		3
G.GFNSSDTVTSPQR.A	2	
G.GSAWSHPQFEK.-		
H.GAGGGAFPASQTPSKPASADGHR.G		
H.RGPSAAFAPAAAEPK.L		
K.LFGGFNES*DTVTSPQR.A		5
K.LFGGFNESDT*VTSPQR.A		5
K.LFGGFNESDTVT*SPQR.A		5
K.LFGGFNESDTVTS*PQR.A		5
K.LFGGFNESDTVTSPQR.A		2
K.LFGGFNS*SDTVTSPQR.A		
K.LFGGFNSS*DTVTSPQR.A		
K.LFGGFNSSDT*VTSPQR.A		
K.LFGGFNSSDTVT*SPQR.A		
K.LFGGFNSSDTVTS*PQR.A		
K.LFGGFNSSDTVTSP.Q		
K.LFGGFNSSDTVTSPQR.A	3	
K.PASADGHRGPSAAFAPAAAEPK.L		1
L.AGGSAWSHPQFEK.-		
N.VHGAGGGAFPA.S		
N.VHGAGGGAFPASQTPSKPASAD.G		
N.VHGAGGGAFPASQTPSKPASADGHR.G		
P.LAGGSAWSHP.Q		1
P.LAGGSAWSHPQFE.K		
P.SAAFAPAAAEPK.L	1	2
Q.TPSKPASADGHRGPSAAFAPAAAEPK.L		
R.AGPLAGGSAWSHP.Q	1	2
R.AGPLAGGSAWSHPQFE.K	1	2
R.AGPLAGGSAWSHPQFEK.-	2	3
R.GPSAAFAPAAAEPK.L	3	3
R.SLEPAENVH.G		
R.SLEPAENVHG.A		
R.SLEPAENVHGA.G		
S.AAFAPAAAEPK.L	1	2
S.AWSHPQFE.K		1
S.AWSHPQFEK.-		1
W.SHPQFEK.-		

Area

Sample	LFQGFNSSDTVTSPQR	S75-p-LFQGFNSSDTVTSPQR	S69-p-LFQGFNSSDTVTSPQR	LFQGFNESDTVTSPQR	S75-p-LFQGFNESDTVTSPQR
USrc WT	4114360,19	119266,88	19373,11	NF	NF
USrc S69E	NF	NF	NF	7956834,74	1385361,79

Normalized area

Sample	LFQGFNSSDTVTSPQR	S75-p-LFQGFNSSDTVTSPQR	S69-p-LFQGFNSSDTVTSPQR	LFQGFNESDTVTSPQR	S75-p-LFQGFNESDTVTSPQR
USrc WT	2,10819	0,06111	0,00993	#VALUE!	#VALUE!
USrc S69E	#VALUE!	#VALUE!	#VALUE!	1,18843	0,20692

Table S1. List of the digested fragments obtained from USrc WT and USrc S69E samples. pSer 75 was found in both conditions (WT, S69E). NF = not found. #value = not calculable.

II.2 - Discussion: p-sites identification

p-LFGGFNSSDTVTSPQR (63-78)
p-site identification

m/z	Ion	S69	S70	T72	T74	S75	p-site
477	Y4-H2O	x	x	x	x		S69, S70, T72, T74
499	b9-H2O 2+	x	x				S69, S70
557	**y4-H2O**					**x**	**S75**
561	y5-2H2O	x	x	x			S69, S70, T72
575	**y4**					**x**	**S75**
579	y5-H2O	x	x	x			S69, S70, T72
597	y5	x	x	x			S69, S70, T72
677	y5				x	x	T75, S75
801	b8-H2O			x	x	x	T72,T74,S75
935	b9			x	x	x	T72,T74,S75
977	y8-H2O			x	x	x	T72,T74,S75
985	**y9-H2O**	**x**					**S69**
995	y8			x	x	x	T72,T74,S75
1065	y9-H2O		x	x	x	x	S70, T72, T74, S75
1083	y9		x	x	x	x	S70, T72, T74, S75
1119	b10-H2O				x	x	T74, S75
1137	b10				x	x	T74, S75
1203	**b12-2H2O**					**x**	**S75**
1221	**b12-H2O**					**x**	**S75**
1239	**b12**					**x**	**S75**
1309	**b13-P**					**x**	**S75**

Table S2. Fragment ions that ascertain p-site location are summarized. Orange and light green colors indicate fragments matching S75 or S69, but that not provide unambiguous assignment to one of the two species. Instead, unique peaks that exclusively belong to p-site S75 and S69 are marked in red and in dark green, respectively. Considering that the isolated parent ion for the MS/MS experiments corresponds to a monophosphorylated form, MS/MS fragments with or without phosphorylation can both point to specific phosphorylations sites.

II.3 - Src phosphosphoproteomics in CRC SW620 cells

Protein	Protein names	Gene names	mino acd	Phospho (STY) Probabilities	Mass error [ppm]	Ratio A3/WT
P07108	Acyl-CoA-binding protein	DBI	Y	TKPSDEEMLFIY(0.083)GHY(0.917)K	-0,5041	0,013211423
P23469-3	Receptor-type tyrosine-protein phosphatase epsilon	PTPRE	S	VVQDFDIFS(0.841)DY(0.159)ANFK	2,4972	0,015374672
F8VZY9	Keratin, type I cytoskeletal 18	KRT18	Y	STFS(0.019)T(0.102)NY(0.879)R	0,63441	0,096940064
O60716-13	Catenin delta-1	CTNND1	S	APS(1)RQDVY(1)GPQPQVR	-0,48131	0,100632739
O60716-13	Catenin delta-1	CTNND1	Y	APS(1)RQDVY(1)GPQPQVR	-0,48131	0,100632739
Q9UHR4	ecific angiogenesis inhibitor 1-associated protein 2-like	BAIAP2L1	Y	KDY(1)DTLSK	0,45507	0,101194247
Q99704-3	Docking protein 1	DOK1	Y	VKEEGYELPY(0.001)NPAT(0.191)DDY(0.808)AVPPPR	1,0599	0,104701599
O75886	Signal transducing adapter molecule 2	STAM2	Y	LVNEAPVY(0.903)S(0.126)VY(0.874)S(0.097)K	0,38326	0,10849706
D6R9B4	Sialomucin core protein 24	CD164	Y	ERNY(1)HTL	0,22797	0,122558509
Q5TCQ9-3			Y	IEDPQY(0.88)GT(0.059)Y(0.047)Y(0.014)VDHLNQK	0,10429	0,13611179
Q99704-3	Docking protein 1	DOK1	Y	SHNSALY(0.988)S(0.012)QVQK	0,63565	0,151465483
P35221-3	Catenin alpha-1	CTNNA1	Y	LVY(1)DGIR	0,36867	0,156956587
B7Z941	Junctional adhesion molecule A	F11R	Y	KVIY(0.984)S(0.016)QPSAR	-0,1064	0,159061057
P29350-2	Tyrosine-protein phosphatase non-receptor type 6	PTPN6	Y	HKEDVY(1)ENLHTK	-0,039929	0,159694321
Q14289-2	Protein-tyrosine kinase 2-beta	PTK2B	Y	SPLT(0.008)PEKEVGY(0.991)LEFTGPPQKPPR	-0,86074	0,162909149
P07355	Annexin A2;Annexin;Putative annexin A2-like protein	NXA2;ANXA2	Y	LSLEGDHSTPPSAY(0.999)GSVK	-0,53323	0,163426557
Q9NVH2-4	Integrator complex subunit 7	INTS7	Y	SLEDPY(0.959)S(0.041)QQIR	1,5098	0,164905451
C9JFK9	BAG family molecular chaperone regulator 3	BAG3	T	T(0.006)HY(0.005)PAQQGEY(0.586)QT(0.731)HQPVY(0.673)HK	-0,040357	0,165110679
Q14451-2	Growth factor receptor-bound protein 7	GRB7	S	DASRPHVKVY(0.342)S(0.658)EDGACR	1,216	0,168098923
J3QSY7	Kinesin-like protein KIF16B	KIF16B	Y	QKIY(1)EVDGVQK	0,45105	0,169005795
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	Y	VGFQYEGT(0.133)Y(0.867)K	-0,90138	0,173943828
P11413	Glucose-6-phosphate 1-dehydrogenase	G6PD	Y	VQPNEAVY(0.615)T(0.385)K	-0,12239	0,174895206
K7ES69	Calponin-2;Calponin-3	CNN2;CNN3	Y	GPS(0.02)Y(0.952)GLS(0.028)AEVK	0,42715	0,177494873
G3V1A4	Cofilin-1	CFL1	T	EILVGDVGQTVDDPY(0.071)AT(0.929)FVK	0,31346	0,177572236
Q9Y2I9-3	TBC1 domain family member 30	TBC1D30	Y	QQQQVHQVY(1)IR	-1,5739	0,177875019
O95425-2	Supervillin	SVIL	Y	GMLREY(1)GETESKR	-0,21066	0,178333106

Table S3. List of the selected phosphopeptides from AAA condition exhibiting decreased phosphorylation (<0.18) with respect to the WT.

Protein	Protein names	Gene names	mho ac	Phospho (STY) Probabilities	Mass error [ppm]	Ratio A3/WT
H0YII7	Pre-mRNA-processing factor 40 homolog B	PRPF40B	T	FVCDS(0.077)AFEQIT(0.923)LES(0.001)ER	-1,9445	1,426973077
Q9H4L5-4	Oxysterol-binding protein-related protein 3	OSBPL3	S	WHES(0.615)IY(0.384)CGGGS(0.001)SSACVWR	0,3431	1,427473231
P63167	Dynein light chain 1, cytoplasmic	DYNLL1	Y	Y(1)NEKDIAAHIK	0,55972	1,430284459
O00571-2	hident RNA helicase DDX3X;ATP-dependent RNA helica	DX3X;DDX3	Y	KGADS(0.001)LEDFLY(0.999)HEGYACTSIHGDR	0,87497	1,434491712
B4E3S0	Coronin;Coronin-1C	CORO1C	Y	YFEITDES(0.004)PY(0.991)VHY(0.005)LNTFSSK	1,952	1,45077574
Q8NDI1-3	EH domain-binding protein 1	EHBP1	Y	VGNYETDT(0.001)NS(0.004)S(0.004)YDQEKFY(0.865)AELS(0.125)DLKR	-0,094311	1,475733142
P14618	Pyruvate kinase PKM	PKM	Y	EAEAAIY(1)HLQLFEELRR	0,92376	1,476508527
Q86YV9	Hermansky-Pudlak syndrome 6 protein	HPS6	Y	DLVFEEACGY(0.983)Y(0.017)QR	0,46068	1,48069603
Q14126	Desmoglein-2	DSG2	Y	VY(0.902)APAS(0.073)T(0.023)LVDQPY(0.001)ANEGTVVVTER	-0,15523	1,493244656
O15357-2	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase	INPPL1	Y	NSFNNPAY(0.846)Y(0.154)YLEGVPHQLLPPEPPSPAR	1,7055	1,521068522
Q15269	Periodic tryptophan protein 2 homolog	PWP2	Y	WDNLIY(0.854)Y(0.146)ALGGHK	-1,7461	1,523674634
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEYMLPVHFY(1)GR	-0,19278	1,547603166
P06239	Tyrosine-protein kinase Lck	LCK	Y	LGAGQFGEVWMGY(0.143)Y(0.839)NGHT(0.019)K	0,67083	1,553058416
Q14185	Dedicator of cytokinesis protein 1	DOCK1	Y	GS(0.037)VADY(0.96)GNLMENQDLLGS(0.003)PTPPPPPPHQR	-1,0861	1,571917669
Q99959-2	Plakophilin-2	PKP2	Y	AHY(0.822)T(0.174)HS(0.004)DYQYSQR	-0,44568	1,580938007
Q8N4X5-2	Actin filament-associated protein 1-like 2	AFAP1L2	Y	VAQQPLSLVGCEVVPDPS(0.018)PDHLY(0.766)S(0.216)FR	-0,73294	1,61976891
Q96CW5-2	Gamma-tubulin complex component 3	TUBGCP3	Y	WYDGELEDT(0.234)Y(0.765)HEFFVAS(0.001)DPTVK	0,17275	1,639594784
H0YHF9	Probable E3 ubiquitin-protein ligase DTX3	DTX3	Y	Y(0.804)GT(0.196)IVIQYVFPPGVQGAEHPNPGVR	1,7372	1,65578635
K7EL62	Guanine nucleotide-binding protein subunit alpha-11	GNA11	Y	ILY(0.722)S(0.212)HLVDY(0.066)FFEFDGPQR	0,4495	1,656161631
Q8WZ64	Rho-GAP domain, ANK repeat and PH domain-contain	ARAP2	T	SVAWQNSNEENS(0.001)S(0.002)S(0.001)IFPY(0.459)GET(0.537)FLFQR	-0,95356	1,682540941
Q15181	Inorganic pyrophosphatase	PPA1	Y	Y(1)KVPDGKPENEFAFNAEFK	-0,26919	1,707670802
Q9NRW3	DNA dC->dU-editing enzyme APOBEC-3C	APOBEC3C	Y	YCWENFVY(1)NDNEPFKPWK	-0,022029	1,737902043
O15027-2	Protein transport protein Sec16A	SEC16A	Y	ANHSSHQEDT(0.005)Y(0.995)GALDFTLSR	-0,32876	1,76124634
F8VR50	Actin-related protein 2/3 complex subunit 3	ARPC3	Y	DTDIVDEAIY(0.985)Y(0.015)FK	0,74546	1,764466826
Q96A22	Uncharacterized protein C11orf52	C11orf52	Y	SPGLMS(0.001)EDS(0.001)NLHY(0.912)ADIQVCS(0.086)RPHAR	0,24556	1,785347902
B4DX66	Epidermal growth factor receptor kinase substrate 8	EPS8	Y	RLSTEHSSVSEYHPADGYAFS(0.005)S(0.013)NIY(0.886)T(0.096)R	-0,97543	1,791609964
P15924-2	Desmoplakin	DSP	Y	EKS(0.033)AIY(0.967)QLEEEYENLLK	1,089	1,808695151
B4DX66	Epidermal growth factor receptor kinase substrate 8	EPS8	Y	LSTEHSSVS(0.016)EY(0.036)HPADGY(0.933)AFS(0.005)S(0.005)NIY(0.448)T(0.548)R	-0,43997	1,821020926
P27797	Calreticulin	CALR	Y	FY(1)ALSASFEPFSNK	-0,25248	1,863335444
B4DX66	Epidermal growth factor receptor kinase substrate 8	EPS8	T	LSTEHSSVSEYHPADGYAFS(0.033)S(0.098)NIY(0.229)T(0.64)R	-0,13948	1,90553219
Q9BXJ9-4	N-alpha-acetyltransferase 15, NatA auxiliary subunit	NAA15	Y	LKIY(1)EEAWTK	0,37691	1,987199648

Protein	Protein names	Gene names	mino acid	Phospho (STY) Probabilities	Mass error [ppm]	Ratio A3/WT
Q9Y333	U6 snRNA-associated Sm-like protein LSm2	LSM2	Y	MLFY(0.981)S(0.019)FFK	-0,34705	2,048229222
Q9Y6N7-6	Roundabout homolog 1	ROBO1	Y	THLIQEDILPY(0.995)CRPT(0.005)FPTSNNPR	-1,2543	2,062989422
K7ENJ4	ATP synthase subunit alpha, mitochondrial	ATP5A1	Y	KLY(0.997)CIY(0.003)VAIGQK	-0,45946	2,092647229
V9GY89	Xaa-Pro dipeptidase	PEPD	Y	VGMKEY(0.675)ELES(0.325)LFEHYCYSR	1,016	2,093514052
D6RER8	Putative sodium-coupled neutral amino acid transporter	SLC38A9	Y	ALIAPDHVVPAPEECY(0.012)VY(0.832)S(0.139)PLGS(0.012)AY(0.005)K	1,02774	2,140478886
I3L404	40S ribosomal protein S2	RPS2	Y	AFVAIGDY(1)NGHVGLGVK	1,4351	2,170782066
G3V1A4	Cofilin-1	CFL1	Y	EILVGDVGQTVDDPY(0.956)AT(0.044)FVK	0,42734	2,185569065
Q71U36-2	alpha-3C/D chain;Tubulin alpha-4A chain;Tubulin alpha-C;TUBA4A;TU		Y	AY(0.932)HEQLS(0.068)VAEITNACFEPANQMVK	0,97153	2,185977011
Q6JPS3	beta-4B chain;Tubulin beta-4A chain;Tubulin beta-2B c;TUBB4A;TU		T	LT(0.018)T(0.052)PT(0.501)Y(0.427)GDLNHLVS(0.001)ATMSGVTTCLR	1,2424	2,205556335
Q6RI18	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	S	AKS(0.999)PQPPVEEEDEHFDDT(0.001)VVCLDTYNCDLHFK	-0,4083	2,209367562
Q14204	Cytoplasmic dynein 1 heavy chain 1	DYNC1H1	Y	HVPVVY(0.001)VDYPGPAS(0.004)LT(0.014)QIY(0.969)GT(0.011)FNR	-0,84197	2,295668393
O75083-3	WD repeat-containing protein 1	WDR1	Y	AHDGGIY(0.997)AIS(0.003)WSPDSTHLLSASGDK	0,66898	2,336582818
H0YB16	Focal adhesion kinase 1	PTK2	Y	YMEDS(0.015)T(0.015)Y(0.96)Y(0.937)KAS(0.072)K	0,70082	2,375912409
O60506-5	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	Y	LKDY(1)AFIHFDER	-0,31638	2,380106749
B7ZM68	FYVE, RhoGEF and PH domain-containing protein 5	FGD5	T	HLFLMNDVLLY(0.211)T(0.705)Y(0.084)PQK	-0,40204	2,394935866
Q14126	Desmoglein-2	DSG2	Y	VY(0.885)APAS(0.058)T(0.058)LVDQPY(0.998)ANEGT(0.002)VVVTER	0,61307	2,415764844
Q9BQL6-4	Fermitin family homolog 1	FERMT1	Y	AFKQY(1)WFIFK	0,23024	2,533858622
C9JME6	Exocyst complex component 7	EXOC7	S	TLS(0.002)CLDHVIS(0.539)Y(0.427)Y(0.031)HVASDTEK	-0,84964	2,533869139
B4E3P0	ATP-citrate synthase	ACLY	Y	TTDGVY(1)EGVAIGGDR	-0,11222	2,547817558
J3QQI0	ARF GTPase-activating protein GIT1	GIT1	Y	LQPFHSTELEDDAIY(0.931)S(0.069)VHVPAGLYR	-1,0157	2,656137277
Q9UP95-3	Solute carrier family 12 member 4	SLC12A4	Y	DLAVFLY(1)HLR	-0,43214	2,659877142
J3QR64	Eukaryotic initiation factor 4A-I	EIF4A1	Y	GFKDQIY(1)DIFQK	-0,0456O1	2,660341119
J9JIE0	G-protein coupled receptor family C group 5 member C	GPRC5C	S	RPVS(0.061)PY(0.432)S(0.499)GY(0.007)NGQLLT(0.131)S(0.625)VY(0.232)QPT(0.012)EMALMHK	0,73892	2,706782557
Q9H4G0-4	Band 4.1-like protein 1	EPB41L1	Y	IRPGEY(1)EQFESTIGFK	-0,18358	2,834535363
Q71U36-2	Chain;Tubulin alpha-4A chain;Tubulin alpha-1B chain;TUBA4A,TUBA		Y	IHFPLAT(0.056)Y(0.943)APVISAEK	0,32006	2,835812056
Q006I0-2	Clathrin heavy chain 1	CLTC	Y	SVNESLNNLFIT(0.04)EEDY(0.96)QALR	-0,32843	2,84308445
H7C0E2	Caspase-8;Caspase-8 subunit p18;Caspase-8 subunit p1	CASP8	Y	GIIY(0.002)GT(0.007)DGQEAPIY(0.99)ELT(0.002)SQFTGLK	1,6242	2,849041215
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEY(1)MLPVHFYGR	-0,27082	2,852890594
Q9Y5X1	Sorting nexin-9	SNX9	Y	EKIPIIVGDY(1)GPMWVYPTSTFDCVVADPR	0,041657	2,857088761
F5H3F0	Ubiquitin thioesterase OTUB1	OTUB1	Y	VYLLY(0.992)RPGHY(0.003)DILY(0.005)K	0,71732	2,883025202
G3V2W4	Alpha-actinin-1;Alpha-actinin-4	ACTN1;ACTN	Y	AIMTYVSSFY(0.999)HAFSGAQK	0,0019376	2,974626985
V9GYF0	Rho guanine nucleotide exchange factor 2	ARHGEF2	Y	S(0.001)LPAGDALY(0.992)LS(0.007)FNPPQPSR	0,242	2,995001315
P00338	L-lactate dehydrogenase A chain	LDHA	Y	DQLY(1)NLLK	0,65406	3,120469451
P06737-2	Glycogen phosphorylase, liver form;Phosphorylase	PYGL	Y	SRPEFMLPVHFY(1)GK	0,27787	3,438447024
Q9H400-2	Lck-interacting transmembrane adapter 1	LIME1	Y	ALDVDSGPLENVY(1)ESIR	0,10154	3,66474664
Q006I0-2	Clathrin heavy chain 1	CLTC	Y	LAELEEFINGPNNAHIQQVGDRCY(1)DEK	0,53334	3,772013624
Q5TBN5	Plastin-2	LCP1	Y	VDTDGNGY(1)ISFNELNDLFK	1,1814	3,781138982
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC	Y	LPQLVDMAAQIAS(0.031)GMAY(0.969)VER	0,32253	3,782236902
C9J1S9	Mitochondrial chaperone BCS1	BCS1L	Y	TVMYTAVGS(0.003)EWRPFGY(0.997)PR	0,82827	4,58520929
Q3669G9	Protein naked cuticle homolog 1	NKD1	Y	NHYLDLAGIENY(0.998)T(0.002)SGFGPGSPSVAQK	0,66554	4,666208414
Q995G9-2	Plakophilin-4	PKP4	Y	NNY(1)ALNTT(0.001)AT(0.041)Y(0.956)AEPY(0.001)RPIQYR	-0,38388	4,814242913
H3BSJ9	Cytochrome b-c1 complex subunit 2, mitochondrial	UQCRC2	Y	VTSEELHY(0.999)FVQNHFTSAR	1,5346	5,018965002
Q5SXB7	Rab GDP dissociation inhibitor beta	GDI2	Y	NPY(0.991)Y(0.009)GGESASITPLEDLYKR	1,02	5,436221209
Q3669G9	Protein naked cuticle homolog 1	NKD1	Y	RNHY(0.995)LDLAGIENY(0.545)T(0.19)S(0.255)QFGPGS(0.014)PS(0.001)VAQK	-0,49779	6,011392744

Table S4. List of the selected phosphopeptides from AAA condition exhibiting increased phosphorylation (>1.45) with respect to the WT.

Protein	Protein names	Gene names	Amino aci	Phospho (STY) Probabilities	Mass error [ppm]	Ratio S75E/WT
P23469-3	Receptor-type tyrosine-protein phosphatase epsilon	PTPRE	S	VVQDFIDIFS(0.841)DY(0.159)ANFK	2,4972	0,009734919
P07355	Annexin A2;Annexin;Putative annexin A2-like protein	ANXA2;ANXA2P2	Y	LSLEGDHSTPPSAY(0.999)GSVK	-0,53323	0,014283792
P07355	Annexin A2;Annexin;Putative annexin A2-like protein	ANXA2;ANXA2P2	S	LSLEGDHSTPPS(0.001)AY(0.426)GS(0.573)VK	0,79319	0,063075556
O14828	Secretory carrier-associated membrane protein 3	SCAMP3	S	KLS(0.817)PT(0.182)EPKNY(0.086)GS(0.184)Y(0.155)S(0.52)T(0.046)QAS(0.01)AAAATAELLK	0,55098	0,068339041
O14828	Secretory carrier-associated membrane protein 3	SCAMP3	S	KLS(0.817)PT(0.182)EPKNY(0.086)GS(0.184)Y(0.155)S(0.52)T(0.046)QAS(0.01)AAAATAELLK	0,55098	0,068339041
Q99704-3	Docking protein 1	DOK1	Y	KKPLY(1)WDLY(1)EHAQQQLLK	-1,2093	0,073244743
Q99704-3	Docking protein 1	DOK1	T	LT(0.616)DPKEDPIY(0.384)DEPEGLAPVPPQGLY(1)DLPR	0,25826	0,080053311
Q99704-3	Docking protein 1	DOK1	Y	LT(0.606)DPKEDPIY(0.394)DEPEGLAPVPPQGLY(1)DLPR	-0,043702	0,080053311
D6R9B4	Sialomucin core protein 24	CD164	Y	ERNY(1)HTL	0,22797	0,090054796
F8VZY9	Keratin, type I cytoskeletal 18	KRT18	Y	STFS(0.019)T(0.102)NY(0.879)R	0,63441	0,105610584
Q99704-3	Docking protein 1	DOK1	Y	VKEEGYELPY(0.001)NPAT(0.191)DDY(0.808)AVPPPR	1,0599	0,107923065
Q8NFJ5	Retinoic acid-induced protein 3	GPRC5A	Y	AYSQEEITQGFEETGDT(0.001)LY(0.996)APY(0.849)S(0.077)T(0.077)HFQLQNQPPQK	-0,1399	0,120307443
Q9H3Z4-2	DnaJ homolog subfamily C member 5	DNAJC5	Y	Y(1)HPDKNPDNPEAADKFK	0,60742	0,132365548
P98172	Ephrin-B1	EFNB1	Y	TTENNY(1)CPHY(1)EK	0,86648	0,142180413
P35221-3	Catenin alpha-1	CTNNA1	Y	LVY(1)DGIR	0,36867	0,14268039
Q8WVV4-3	Protein POF1B	POF1B	Y	GSHFFPGNNVIY(1)EK	0,35565	0,14646372
B4DRS7	SH2 domain-containing protein 3A	SH2D3A	S	T(0.483)PS(0.517)FELPDAS(0.002)ERPPT(0.421)Y(0.577)CELVPR	0,93387	0,15870304
B4DRS7	SH2 domain-containing protein 3A	SH2D3A	Y	T(0.483)PS(0.517)FELPDAS(0.002)ERPPT(0.421)Y(0.577)CELVPR	0,93387	0,15870304
B4DPY6	Hepatocyte growth factor receptor	MET	Y	DMYDKEY(1)Y(0.998)S(0.002)VHNK	-0,25546	0,160597279
O15357-2	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2	INPPL1	S	NS(0.001)FNNPAY(0.778)Y(0.221)VLEGVPHQLLPPEPPS(1)PAR	0,014883	0,161115728
B7Z941	Junctional adhesion molecule A	F11R	Y	KVIY(0.984)S(0.016)QPSAR	-0,1064	0,165791014
Q6IAA8	Ragulator complex protein LAMTOR1	LAMTOR1	Y	IAAY(1)AY(0.9)S(0.1)ALSQIR	0,0043656	0,166706059
Q6IAA8	Ragulator complex protein LAMTOR1	LAMTOR1	Y	IAAY(1)AY(0.9)S(0.1)ALSQIR	0,0043656	0,166706059
Q07666-3	KH domain-containing, RNA-binding, signal transduction-associated protein 1	KHDRBS1	Y	GAY(1)REHPY(0.512)GRY(0.488)	0,10589	0,167063359
P35221-3	Catenin alpha-1;Catenin alpha-2	CTNNA1;CTNNA2	Y	EVKEY(1)AQVFR	0,31867	0,168708146
Q13546-2	Receptor-interacting serine/threonine-protein kinase 1	RIPK1	Y	VSHDPFAQQRPY(0.914)ENFQNT(0.086)EGK	0,56687	0,169165511
P29350-2	Tyrosine-protein phosphatase non-receptor type 6	PTPN6	Y	HKEDVY(1)ENLHTK	-0,039929	0,170148697
O75886	Signal transducing adapter molecule 2	STAM2	Y	LVNEAPVY(0.903)S(0.126)VY(0.874)S(0.097)K	0,38326	0,171728661
Q99704-3	Docking protein 1	DOK1	Y	VKEEGY(0.259)ELPY(0.733)NPAT(0.336)DDY(0.672)AVPPPR	0,54614	0,172200402
Q14247	Src substrate cortactin	CTTN	S	LPS(0.003)S(0.997)PVY(1)EDAASFK	-0,26324	0,173896131
E9PF63	Rho-associated protein kinase 2	ROCK2	Y	IY(1)ESIEEAK	-0,14336	0,173897608
F8VZY9	Keratin, type I cytoskeletal 18	KRT18	Y	PVSSAASVY(1)AGAGGSGSR	-0,011492	0,174996803
Q969G9	Protein naked cuticle homolog 1	NKD1	Y	LEQS(0.005)GCY(0.995)HHCVDENIER	-0,38201	0,177036957
Q9P258	Protein RCC2	RCC2	Y	VFS(0.002)WGFGGY(0.998)GR	0,31699	0,181471549
Q99704-3	Docking protein 1	DOK1	Y	SHNSALY(0.988)S(0.012)QVQK	0,63565	0,182277617
P13639	Elongation factor 2	EEF2	Y	LMEPIYLVEIQCPEQVVGGIY(1)GVLNR	0,94925	0,182621609
P26599	Polypyrimidine tract-binding protein 1	PTBP1	Y	EGQEDQGLT(0.001)KDY(0.997)GNS(0.003)PLHR	-0,010754	0,188937738
E9PMS6	LIM domain only protein 7	LMO7	Y	KAQS(0.001)NPY(0.412)Y(0.587)NGPHLNLK	0,027843	0,191700927
Q8NFJ5	Retinoic acid-induced protein 3	GPRC5A	S	AY(0.142)S(0.851)QEEITQGFEETGDT(0.032)LY(0.683)APY(0.239)S(0.031)T(0.014)HFQLQNQPPQK	-1,1307	0,192095264
F5H855	Breast cancer anti-estrogen resistance protein 1	BCAR1	Y	GLPPSNHHAVY(0.853)DVPPS(0.066)VS(0.081)K	-0,060972	0,192321595

Table S5. List of the selected phosphopeptides from S75E condition exhibiting decreased phosphorylation (<0.2) with respect to the WT.

Protein	Protein names	Gene names	Amino aci	Phospho (STY) Probabilities	Mass error [ppm]	Ratio S75E/WT
P00505	Aspartate aminotransferase, mitochondrial	GOT2	Y	NLDKEY(1)LPIGGLAEFCK	-0,21514	1,110390676
J3KRI6	T-complex protein 1 subunit zeta-2;T-complex protein 1 subunit zeta	CCT6B;CCT6A	Y	TEVNS(0.001)GFFY(0.999)K	0,56071	1,111126386
H0YII7	Pre-mRNA-processing factor 40 homolog B	PRPF40B	T	FVCDS(0.077)AFEQIT(0.923)LES(0.001)ER	-1,9445	1,112452937
A8MT37	Glycogen synthase kinase-3 beta;Glycogen synthase kinase-3 alpha	GSK3A;GSK3B	Y	GEPNVS(0.011)Y(0.989)ICSR	-0,045272	1,115898696
Q8NFJ5	Retinoic acid-induced protein 3	GPRC5A	Y	AYSQEEITQGFEETGDT(0.001)LY(0.996)APY(0.849)S(0.077)T(0.077)HFQLQNQPPQK	-0,1399	1,119113207
B4DHN5	Syntenin-1	SDCBP	Y	VIQAQTAFSANPANPAILSEASAPIPHDGNLY(0.999)PR	0,75748	1,138477753
S4R435	40S ribosomal protein S10	RPS10	Y	IAIY(1)ELLFK	-0,079921	1,142735043
E5RIU6	Cyclin-dependent kinase 1; kinase 2; kinase 3	CDK1;CDK2;CDK3	Y	IGEGT(0.001)Y(0.999)GVVYK	-0,051021	1,14591955
S4R3Q6	Vacuolar protein sorting-associated protein 26A	VPS26A	Y	EYDLIVHQLAT(0.02)Y(0.977)PDVNNS(0.003)IK	0,85997	1,148366518
Q5RI18	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	S	AKS(0.999)PQPPVEEEDEHFDDT(0.001)VVCLDTYNCDLHFK	-0,4083	1,168650119
Q9H792	Pseudopodium-enriched atypical kinase 1	PEAK1	T	FNS(0.003)Y(0.002)NNAGMPPFPIIIHDEPT(0.803)Y(0.193)AR	0,97131	1,172033285
K7EL62	Guanine nucleotide-binding protein subunit alpha-11	GNA11	Y	ILY(0.722)S(0.212)HLVDY(0.066)FPEFDGPQR	0,4495	1,195602024
C9JLN0	Cysteine--tRNA ligase, cytoplasmic	CARS	Y	LHLY(0.677)NS(0.32)LT(0.003)R	-0,12488	1,204775812
Q14511-3	Enhancer of filamentation 1;Enhancer of filamentation 1 p55	NEDD9	Y	HQSLSPNHPPPQLGQS(0.001)VGS(0.04)QNDAY(0.958)DVPR	-0,49371	1,222334803
P11171-4	Protein 4.1	EPB41	Y	LDGENIY(1)IR	-0,088506	1,230572276
P00519	Tyrosine-protein kinase ABL1;Abelson tyrosine-protein kinase 2	ABL1;ABL2	Y	LMT(0.002)GDT(0.275)Y(0.693)T(0.03)AHAGAK	0,10353	1,233693627
O15197-2	Ephrin type-B receptor 6	EPHB6	Y	Y(0.5)Y(0.5)IDPSTYEDPCQAIR	-0,67553	1,254543379
O15197-2	Ephrin type-B receptor 6	EPHB6	Y	Y(0.5)Y(0.5)IDPSTYEDPCQAIR	-0,67553	1,254543379
Q9Y617-2	Phosphoserine aminotransferase	PSAT1	Y	AS(0.02)LY(0.98)NAVTIEDVQK	-0,75571	1,262420495
B4DX66	Epidermal growth factor receptor kinase substrate 8	EPS8	Y	LSTEHS(0.004)S(0.004)VS(0.016)EY(0.036)HPADGY(0.933)AFS(0.005)S(0.005)NIY(0.448)T(0.548)R	-0,43997	1,264267597
E9PBS1	Phosphoribosylaminoimidazole-succinocarboxamide synthase	PAICS	Y	T(0.002)KEVY(0.998)ELLDSPGK	0,20414	1,278307151
Q96L35	Ephrin type-B receptor 4;Ephrin type-B receptor 3	EPHB4;EPHB3	Y	VY(1)IDPFTYEDPNEAVR	-0,29376	1,278887304
F5GYU9	Pleckstrin homology domain-containing family B member 1	PLEKHB1	Y	VYSPY(0.018)QDY(0.216)Y(0.766)EVVPPNAHEATYVR	-1,1547	1,300583491
H7C0E2	Caspase-8;Caspase-8 subunit p18;Caspase-8 subunit p10	CASP8	Y	GIIY(0.002)GT(0.007)DGQEAPIY(0.99)ELT(0.002)SQFTGLK	1,6242	1,30635586
O15197-2	Ephrin type-B receptor 6	EPHB6	Y	YYIDPS(0.001)T(0.004)Y(0.995)EDPCQAIR	-0,29388	1,353584835
F5H3F0	Ubiquitin thioesterase OTUB1	OTUB1	Y	VYLLY(0.992)RPGHY(0.003)DILY(0.005)K	0,71732	1,354380076
P46108	Adapter molecule crk	CRK	Y	RVPNAY(1)DK	0,70147	1,35573273
Q09472	Histone acetyltransferase p300	EP300	Y	VYKY(0.963)CS(0.037)K	0,49343	1,373714822
O00571-2	ATP-dependent RNA helicase DDX3X;ATP-dependent RNA helicase DDX3Y	DDX3X;DDX3Y	Y	KGADS(0.001)LEDFLY(0.999)HEGYACTSIHGDR	0,87497	1,379568604
Q9H4G0-4	Band 4.1-like protein 1	EPB41L1	Y	IRPGEY(1)EQFESTIGFK	-0,18358	1,428635028
F5H855	Breast cancer anti-estrogen resistance protein 1	BCAR1	Y	SQVTHY(0.978)S(0.021)NLLCDLLR	-0,35697	1,447145367
O60506-5	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	Y	LKDY(1)AFIHFDER	-0,31638	1,451638649
P27797	Calreticulin	CALR	Y	FY(1)ALSASFEPFSNK	-0,25248	1,451787462
Q14126	Desmoglein-2	DSG2	Y	VY(0.885)APAS(0.058)T(0.058)LVDQPY(0.998)ANEGT(0.002)VVVTER	0,61307	1,474553963
B4DXV1	Elongator complex protein 3	ELP3	Y	NLHDALSGHTSNNIY(1)EAVK	-0,41259	1,48421371
E7ETQ7	Vacuolar protein sorting-associated protein VTA1 homolog	VTA1	Y	Y(0.961)AGS(0.033)ALQY(0.005)EDVSTAVQNLQK	0,25551	1,487845913
F5H5M1	SHC-transforming protein 4	SHC4	S	Y(0.005)HMDNS(0.842)LPIIS(0.553)S(0.555)GS(0.04)EVS(0.005)LK	0,36947	1,488594268
F5H5M1	SHC-transforming protein 4	SHC4	S	Y(0.005)HMDNS(0.842)LPIIS(0.553)S(0.555)GS(0.04)EVS(0.005)LK	0,36947	1,488594268
F5H5M1	SHC-transforming protein 4	SHC4	S	Y(0.005)HMDNS(0.842)LPIIS(0.553)S(0.555)GS(0.04)EVS(0.005)LK	0,36947	1,488594268
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEYMLPVHFY(1)GR	-0,19278	1,623199039
F5H6V2	Phosphoglucomutase-2	PGM2	Y	IKY(0.5)Y(0.5)AELCAPPGNSDPEQLKK	-0,51683	1,634714791
F5H6V2	Phosphoglucomutase-2	PGM2	Y	IKY(0.5)Y(0.5)AELCAPPGNSDPEQLKK	-0,51683	1,634714791
Q8N1I0-2	Dedicator of cytokinesis protein 4	DOCK4	Y	KEY(1)AMVDPEDISITELYR	0,16829	1,646950902
B4E3S0	Coronin;Coronin-1C	CORO1C	Y	YFEITDES(0.004)PY(0.991)VHY(0.005)LNTFSSK	1,952	1,740761636
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC	Y	LPQLVDMAAQIAS(0.031)GMAY(0.969)VER	0,32253	1,750791784
Q9Y6M7-11	Sodium bicarbonate cotransporter 3	SLC4A7	Y	AVY(1)IGVHVPFSK	-0,87918	1,757381475
P63261	Actin, cytoplasmic 2;Actin, cytoplasmic 1;	ACTG1;ACTB	Y	TTGIVMDSGDGVT(0.002)HT(0.002)VPIY(0.99)EGY(0.006)ALPHAILR	0,29791	1,777112293
V9GY89	Xaa-Pro dipeptidase	PEPD	Y	VGMKEY(0.675)ELES(0.325)LFEHYCYSR	1,016	1,791343331
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEY(1)MLPVHFYGR	-0,27082	1,80575032
Q9Y6N7-6	Roundabout homolog 1	ROBO1	Y	THLIQEDILPY(0.995)CRPT(0.005)FPTSNNPR	-1,2543	1,897272977
Q9UP95-3	Solute carrier family 12 member 4	SLC12A4	Y	DLAVFLY(1)HLR	-0,43214	1,958454575
B3KWE1	Histidine--tRNA ligase, cytoplasmic	HARS	Y	LIY(1)DLKDQGGELLSLR	0,21026	2,092497106
Q9Y333	U6 snRNA-associated Sm-like protein LSm2	LSM2	Y	MLFY(0.981)S(0.019)FFK	-0,34705	2,181348994
I7HJJ0	ADP/ATP translocase 3;ADP/ATP translocase 1;ADP/ATP translocase 2;ADP/ATP translocase 4	SLC25A6;SLC25A4	Y	Y(1)FPTQALNFAFK	1,065	2,331030464
Q99959-2	Plakophilin-2	PKP2	Y	AHYTHSDY(0.995)QY(0.005)SQR	-0,24586	2,372724815
Q5TBN5	Plastin-2	LCP1	Y	VDTDGNGY(1)ISFNELNDLFK	1,1814	2,390569491
G3V2W4	Alpha-actinin-1;Alpha-actinin-4	ACTN1;ACTN4	Y	AIMTYVSSFY(0.999)HAFSGAQK	0,0019376	2,514401979
Q5SX87	Rab GDP dissociation inhibitor beta	GDI2	Y	NPY(0.991)Y(0.009)GGESASITPLEDLYKR	1,02	2,517128493
Q5SX87	Rab GDP dissociation inhibitor beta	GDI2	Y	VLHMDRNPY(0.5)Y(0.5)GGESASITPLEDLYK	1,0411	2,517128493
P00338	L-lactate dehydrogenase A chain	LDHA	Y	DQLIY(1)NLLK	0,65406	2,714489127
Q9HC38-3	Glyoxalase domain-containing protein 4	GLOD4	T	TMVGFGPEDDHFVAELT(0.637)Y(0.356)NY(0.007)GVGDYK	0,24044	3,048911168
H0YB16	Focal adhesion kinase 1	PTK2	Y	YMEDS(0.015)T(0.015)Y(0.96)Y(0.937)KAS(0.072)K	0,70082	3,095952223
P06737-2	Glycogen phosphorylase, liver form;Phosphorylase	PYGL	Y	SRPEFMLPVHFY(1)GK	0,27787	3,401280805
P49023-2	Paxillin	PXN	Y	FIHQQPQSSSPVY(0.91)GS(0.045)S(0.045)AK	0,070504	3,416795528
Q8NDI1-3	EH domain-binding protein 1	EHBP1	Y	VGNYETDT(0.001)NS(0.004)S(0.004)VDQEKFY(0.865)AELS(0.125)DLKR	-0,094311	3,632222279
B4E3P0	ATP-citrate synthase	ACLY	Y	TTDGVY(1)EGVAIGGDR	-0,11222	4,194282912
H3BSJ9	Cytochrome b-c1 complex subunit 2, mitochondrial	UQCRC2	Y	VTSEELHY(0.999)FVQNHFTSAR	1,5346	4,341107631

Table S6. List of the selected phosphopeptides from S75E condition exhibiting increased phosphorylation (>1.1) with respect to the WT.

	PATHWAYS AAA proteins	N of genes		
1	Axon guidance mediated by Slit/Robo (P00008)	1	0.9%	1.4%
2	Apoptosis signaling pathway (P00006)	2	1.8%	2.8%
3	Angiogenesis (P00005)	4	3.5%	5.6%
4	Alzheimer disease-presenilin pathway (P00004)	1	0.9%	1.4%
5	5HT2 type receptor mediated signaling pathway (P04374)	1	0.9%	1.4%
6	Interferon-gamma signaling pathway (P00035)	1	0.9%	1.4%
7	Alpha adrenergic receptor signaling pathway (P00002)	1	0.9%	1.4%
8	Integrin signalling pathway (P00034)	8	7.1%	11.3%
9	Insulin/IGF pathway-protein kinase B signaling cascade (P00033)	1	0.9%	1.4%
10	Inflammation mediated by chemokine and cytokine signaling pathway (P00031)	4	3.5%	5.6%
11	Nicotine pharmacodynamics pathway (P06587)	1	0.9%	1.4%
12	Huntington disease (P00029)	4	3.5%	5.6%
13	Arginine biosynthesis (P02728)	1	0.9%	1.4%
14	terotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway (P000	2	1.8%	2.8%
15	terotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway (P000	3	2.7%	4.2%
16	Wnt signaling pathway (P00057)	3	2.7%	4.2%
17	VEGF signaling pathway (P00056)	2	1.8%	2.8%
18	Glycolysis (P00024)	1	0.9%	1.4%
19	Thyrotropin-releasing hormone receptor signaling pathway (P04394)	1	0.9%	1.4%
20	FGF signaling pathway (P00021)	1	0.9%	1.4%
21	T cell activation (P00053)	1	0.9%	1.4%
22	FAS signaling pathway (P00020)	1	0.9%	1.4%
23	Oxytocin receptor mediated signaling pathway (P04391)	1	0.9%	1.4%
24	ATP synthesis (P02721)	1	0.9%	1.4%
25	Endothelin signaling pathway (P00019)	1	0.9%	1.4%
26	Parkinson disease (P00049)	2	1.8%	2.8%
27	Cytoskeletal regulation by Rho GTPase (P00016)	3	2.7%	4.2%
28	PI3 kinase pathway (P00048)	2	1.8%	2.8%
29	Histamine H1 receptor mediated signaling pathway (P04385)	1	0.9%	1.4%
30	Cadherin signaling pathway (P00012)	3	2.7%	4.2%
31	Dopamine receptor mediated signaling pathway (P05912)	1	0.9%	1.4%
32	B cell activation (P00010)	1	0.9%	1.4%
33	Muscarinic acetylcholine receptor 1 and 3 signaling pathway (P00042)	1	0.9%	1.4%
34	Metabotropic glutamate receptor group I pathway (P00041)	1	0.9%	1.4%
35	Cortocotropin releasing factor receptor signaling pathway (P04380)	1	0.9%	1.4%
36	Pyruvate metabolism (P02772)	1	0.9%	1.4%
37	Gonadotropin releasing hormone receptor pathway (P06664)	6	5.3%	8.5%
	PROTEIN CLASS AAA proteins	**N of genes**		
1	protease (PC00190)	3	2.7%	2.2%
2	cytoskeletal protein (PC00085)	25	22.1%	18.5%
3	transporter (PC00227)	4	3.5%	3.0%
4	transmembrane receptor regulatory/adaptor protein (PC00226)	2	1.8%	1.5%
5	transferase (PC00220)	10	8.8%	7.4%
6	oxidoreductase (PC00176)	3	2.7%	2.2%
7	lyase (PC00144)	1	0.9%	0.7%
8	cell adhesion molecule (PC00069)	8	7.1%	5.9%
9	ligase (PC00142)	3	2.7%	2.2%
10	nucleic acid binding (PC00171)	12	10.6%	8.9%
11	enzyme modulator (PC00095)	13	11.5%	9.6%
12	calcium-binding protein (PC00060)	3	2.7%	2.2%
13	defense/immunity protein (PC00090)	1	0.9%	0.7%
14	hydrolase (PC00121)	14	12.4%	10.4%
15	transfer/carrier protein (PC00219)	2	1.8%	1.5%
16	membrane traffic protein (PC00150)	4	3.5%	3.0%
17	phosphatase (PC00181)	4	3.5%	3.0%
18	transcription factor (PC00218)	2	1.8%	1.5%
19	chaperone (PC00072)	1	0.9%	0.7%
20	cell junction protein (PC00070)	4	3.5%	3.0%
21	structural protein (PC00211)	2	1.8%	1.5%
22	kinase (PC00137)	5	4.4%	3.7%
23	receptor (PC00197)	9	8.0%	6.7%

	PATHWAYS S75E proteins	N of genes		
1	Axon guidance mediated by Slit/Robo (P00008)	2	2.0%	2.5%
2	Apoptosis signaling pathway (P00006)	2	2.0%	2.5%
3	Angiogenesis (P00005)	5	4.9%	6.3%
4	Alzheimer disease-presenilin pathway (P00004)	2	2.0%	2.5%
5	5HT2 type receptor mediated signaling pathway (P04374)	1	1.0%	1.3%
6	Interferon-gamma signaling pathway (P00035)	1	1.0%	1.3%
7	Alpha adrenergic receptor signaling pathway (P00002)	1	1.0%	1.3%
8	Integrin signalling pathway (P00034)	8	7.8%	10.0%
9	Insulin/IGF pathway-protein kinase B signaling cascade (P00033)	2	2.0%	2.5%
10	Inflammation mediated by chemokine and cytokine signaling pathway (P00031)	3	2.9%	3.8%
11	Asparagine and aspartate biosynthesis (P02730)	1	1.0%	1.3%
12	Pentose phosphate pathway (P02762)	1	1.0%	1.3%
13	Nicotine pharmacodynamics pathway (P06587)	2	2.0%	2.5%
14	Huntington disease (P00029)	3	2.9%	3.8%
15	p53 pathway (P00059)	1	1.0%	1.3%
16	Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway (P00027)	1	1.0%	1.3%
17	PLP biosynthesis (P02759)	1	1.0%	1.3%
18	Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway (P00026)	3	2.9%	3.8%
19	Wnt signaling pathway (P00057)	4	3.9%	5.0%
20	VEGF signaling pathway (P00056)	2	2.0%	2.5%
21	Transcription regulation by bZIP transcription factor (P00055)	1	1.0%	1.3%
22	Thyrotropin-releasing hormone receptor signaling pathway (P04394)	1	1.0%	1.3%
23	Vitamin B6 metabolism (P02787)	1	1.0%	1.3%
24	Ras Pathway (P04393)	1	1.0%	1.3%
25	FGF signaling pathway (P00021)	1	1.0%	1.3%
26	TGF-beta signaling pathway (P00052)	1	1.0%	1.3%
27	FAS signaling pathway (P00020)	1	1.0%	1.3%
28	Oxytocin receptor mediated signaling pathway (P04391)	1	1.0%	1.3%
29	Tyrosine biosynthesis (P02784)	1	1.0%	1.3%
30	Endothelin signaling pathway (P00019)	1	1.0%	1.3%
31	Parkinson disease (P00049)	1	1.0%	1.3%
32	Cytoskeletal regulation by Rho GTPase (P00016)	3	2.9%	3.8%
33	PI3 kinase pathway (P00048)	2	2.0%	2.5%
34	PDGF signaling pathway (P00047)	1	1.0%	1.3%
35	Histamine H1 receptor mediated signaling pathway (P04385)	1	1.0%	1.3%
36	Cadherin signaling pathway (P00012)	3	2.9%	3.8%
37	Nicotinic acetylcholine receptor signaling pathway (P00044)	1	1.0%	1.3%
38	Serine glycine biosynthesis (P02776)	1	1.0%	1.3%
39	Dopamine receptor mediated signaling pathway (P05912)	2	2.0%	2.5%
40	B cell activation (P00010)	1	1.0%	1.3%
41	Muscarinic acetylcholine receptor 1 and 3 signaling pathway (P00042)	1	1.0%	1.3%
42	Metabotropic glutamate receptor group I pathway (P00041)	1	1.0%	1.3%
43	Cortocotropin releasing factor receptor signaling pathway (P04380)	1	1.0%	1.3%
44	Gonadotropin releasing hormone receptor pathway (P06664)	5	4.9%	6.3%
	PROTEIN CLASS S75E proteins	**N of genes**		
1	protease (PC00190)	3	2.9%	2.3%
2	cytoskeletal protein (PC00085)	15	14.7%	11.4%
3	transporter (PC00227)	4	3.9%	3.0%
4	transmembrane receptor regulatory/adaptor protein (PC00226)	1	1.0%	0.8%
5	transferase (PC00220)	20	19.6%	15.2%
6	oxidoreductase (PC00176)	4	3.9%	3.0%
7	lyase (PC00144)	2	2.0%	1.5%
8	cell adhesion molecule (PC00069)	7	6.9%	5.3%
9	ligase (PC00142)	4	3.9%	3.0%
10	nucleic acid binding (PC00171)	13	12.7%	9.8%
11	signaling molecule (PC00207)	2	2.0%	1.5%
12	enzyme modulator (PC00095)	8	7.8%	6.1%
13	calcium-binding protein (PC00060)	3	2.9%	2.3%
14	defense/immunity protein (PC00090)	1	1.0%	0.8%
15	hydrolase (PC00121)	10	9.8%	7.6%
16	transfer/carrier protein (PC00219)	2	2.0%	1.5%
17	membrane traffic protein (PC00150)	3	2.9%	2.3%
18	phosphatase (PC00181)	4	3.9%	3.0%
19	transcription factor (PC00218)	5	4.9%	3.8%
20	chaperone (PC00072)	2	2.0%	1.5%
21	cell junction protein (PC00070)	2	2.0%	1.5%
22	structural protein (PC00211)	1	1.0%	0.8%
23	kinase (PC00137)	9	8.8%	6.8%
24	receptor (PC00197)	6	5.9%	4.5%
25	isomerase (PC00135)	1	1.0%	0.8%

Table S7. Gene Ontology of the selected AAA and S75E proteins. Genes encoding for the identified proteins were classified for pathways in which they are involved and protein class at which they belong on the basis of the gene ontology annotation.

Protein	Protein names	Gene names	Amino acid	Phospho (STY) Probabilities	Mass error [ppm]	Ratio A3/WT	Ratio S75E/WT
P23469-3	Receptor-type tyrosine-protein phosphatase epsilon	PTPRE	S	VVODFIDIFS(0.841)DY(0.159)ANFK	2,4972	0,0153747	0,009734919
P07355	Annexin A2;Annexin;Putative annexin A2-like protein	ANXA2;ANXA2P2	Y	LSLEGDHSTPPSAY(0.999)GSVK	-0,53323	0,1634266	0,014283792
D6R9B4	Sialomucin core protein 24	CD164	Y	ERNY(1)HTL	0,22797	0,1225585	0,090054796
F8VZY9	Keratin, type I cytoskeletal 18	KRT18	Y	STFS(0.019)T(0.102)NY(0.879)R	0,63441	0,0969401	0,105610584
Q99704-3	Docking protein 1	DOK1	Y	VKEEGYELPY(0.001)NPAT(0.191)DDY(0.808)AVPPPR	1,0539	0,1047016	0,107923065
P35221-3	Catenin alpha-1	CTNNA1	Y	LVY(1)DGIR	0,36867	0,1569566	0,14268039
B7Z941	Junctional adhesion molecule A	F11R	Y	KVIY(0.984)S(0.016)QPSAR	-0,1064	0,1590611	0,165791014
P29350-2	Tyrosine-protein phosphatase non-receptor type 6	PTPN6	Y	HKEDVY(1)ENLHTK	-0,039929	0,1596943	0,170148697
O75886	Signal transducing adapter molecule 2	STAM2	Y	LVNEAPVY(0.903)S(0.126)VY(0.874)S(0.097)K	0,38326	0,1084971	0,171728661
Q99704-3	Docking protein 1	DOK1	Y	SHNSALY(0.988)S(0.012)QVQK	0,63565	0,1514655	0,182277617
P26599	Polypyrimidine tract-binding protein 1	PTBP1	Y	EGQEDQGLT(0.001)KDY(0.997)GNS(0.003)PLHR	-0,010754	0,1922567	0,188937738
H0YII7	Pre-mRNA-processing factor 40 homolog B	PRPF40B	T	FVCDS(0.077)AFEQIT(0.923)LES(0.001)ER	-1,9445	1,4269731	1,112452937
Q6RI18	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	S	AKS(0.999)PQPPVEEEDEHFDDT(0.001)VVCLDTYNCDLHFK	-0,4083	2,2093676	1,168650119
K7EL62	Guanine nucleotide-binding protein subunit alpha-11	GNA11	Y	ILY(0.722)S(0.212)HLVDY(0.066)FPEFDGPQR	0,4495	1,6561616	1,195602024
B4DX66	Epidermal growth factor receptor kinase substrate 8	EPS8	Y	LSTEHSSVS(0.016)EY(0.036)HPADGY(0.933)AFSSNIY(0.448)T(0.548)R	-0,43997	1,8210209	1,264267597
H7C0E2	Caspase-8;subunit p18;subunit p10	CASP8	Y	GIIY(0.002)GT(0.007)DGQEAPIY(0.99)ELT(0.002)SQFTGLK	1,6242	2,8490412	1,30635586
F5H3F0	Ubiquitin thioesterase OTUB1	OTUB1	Y	VYLLY(0.992)RPGHY(0.003)IDILY(0.005)K	0,71732	2,8830252	1,354380076
O00571-2	ATP-dependent RNA helicase DDX3X	DDX3X	Y	KGADS(0.001)LEDFLY(0.999)HEGYACTSIHGDR	0,87497	1,4344917	1,379568604
Q9H4G0-4	Band 4.1-like protein 1	EPB41L1	Y	IRPGEY(1)EQFESTIGFK	-0,18358	2,8345354	1,428635028
O60506-5	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	Y	LKDY(1)AFIHFDER	-0,31638	2,3801067	1,451638649
P27797	Calreticulin	CALR	Y	FY(1)ALSASFEPFSNK	-0,25248	1,8633354	1,451787462
Q14126	Desmoglein-2	DSG2	Y	VY(0.885)APAS(0.058)T(0.058)LVDQPY(0.998)ANEGT(0.002)VVVTER	0,61307	2,4157648	1,474553963
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEYMLPVHFY(1)GR	-0,19278	1,5476032	1,623199039
B4E3S0	Coronin;Coronin-1C	CORO1C	Y	YFEITDES(0.004)PY(0.991)VHY(0.005)LNTFSSK	1,962	1,4507757	1,740761636
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC	Y	LPQLVDMAAQIAS(0.031)GMAY(0.969)VER	0,32253	3,7822369	1,750791784
V9GY89	Xaa-Pro dipeptidase	PEPD	Y	VGMKEY(0.675)ELES(0.325)LFEHYCYSR	1,016	2,0935141	1,791343331
P11216	Glycogen phosphorylase, brain form	PYGB	Y	ARPEY(1)MLPVHFYGR	-0,27082	2,8528906	1,80575032
Q9Y6N7-6	Roundabout homolog 1	ROBO1	Y	THLIQEDILPY(0.995)CRPT(0.005)FPTSNNPR	-1,2543	2,0629894	1,897272977
Q9UP95-3	Solute carrier family 12 member 4	SLC12A4	Y	DLAVFLY(1)HLR	-0,43214	2,6598771	1,958454575
Q9Y333	U6 snRNA-associated Sm-like protein LSm2	LSM2	Y	MLFY(0.981)S(0.019)FFK	-0,34705	2,0482292	2,181348994
Q5TBN5	Plastin-2	LCP1	Y	VDTDGNGY(1)ISFNELNDLFK	1,1814	3,781139	2,390569491
G3V2W4	Alpha-actinin-1;Alpha-actinin-4	ACTN1;ACTN4	Y	AIMTYVSSFY(0.999)HAFSGAQK	0,00119376	2,974627	2,514401979
Q5SX87	Rab GDP dissociation inhibitor beta	GDI2	Y	NPY(0.991)Y(0.009)GQESASITPLEDLYKR	1,02	5,4362212	2,517128493
P00338	L-lactate dehydrogenase A chain	LDHA	Y	DQLIY(1)NLLK	0,65406	3,1204695	2,714489127
H0YB16	Focal adhesion kinase 1	PTK2	Y	YMEDS(0.015)T(0.015)Y(0.96)Y(0.937)KAS(0.072)K	0,70082	2,3759124	3,095952223
P06737-2	Glycogen phosphorylase, liver form;Phosphorylase	PYGL	Y	SRPEFMLPVHFY(1)GK	0,27787	3,438447	3,401280805
Q8NDI1-3	EH domain-binding protein 1	EHBP1	Y	VGNYETDT(0.001)NS(0.004)S(0.004)YDQEKFY(0.865)AELS(0.125)DLKR	-0,094311	1,4757331	3,632222279
B4E3P0	ATP-citrate synthase	ACLY	Y	TTDGVY(1)EGVAIGGDR	-0,11222	2,5478176	4,194282912
H3BSJ9	Cytochrome b-c1 complex subunit 2, mitochondrial	UQCRC2	Y	VTSEELHY(0.999)FVQNHFTSAR	1,5346	5,018955	4,341107631

Table S8. List of shared selected phosphopeptides between AAA and S75E mutants.

A.3 – Supporting Materials

Material and Methods: Vector Maps

pET-14b sequence landmarks	
T7 promoter	646-662
T7 transcription start	645
His•Tag coding sequence	554-571
Multiple cloning sites	
(*Nde* I -*Bam*H I)	510-526
T7 terminator	404-450
pBR322 origin	2845
bla coding sequence	3606-4463

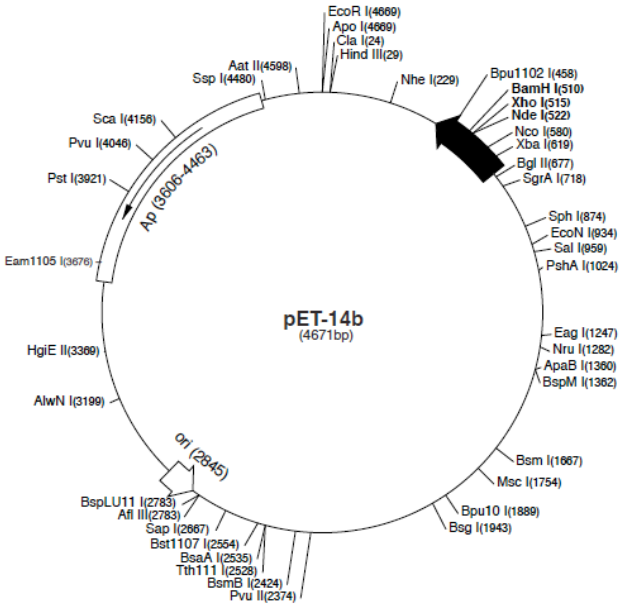


Figure A.3.1. pET-14b vector map (Novagen, UK). Human USrc was cloned into this vector using the *Nco I* and *Nde I* restriction sites. A Strep-tag was added to the C-terminus for purification proposal.

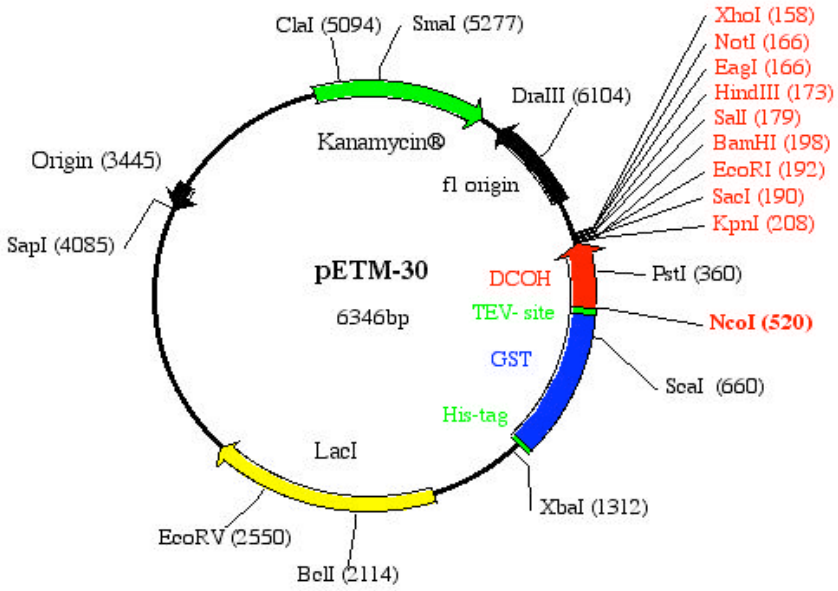


Figure A.3.2. pETM-30 vector map (from EMBL). cDNAs encoding for human USH3 and SH3 proteins were cloned into this vector using the *Nco I* and *Xho I* restriction sites.

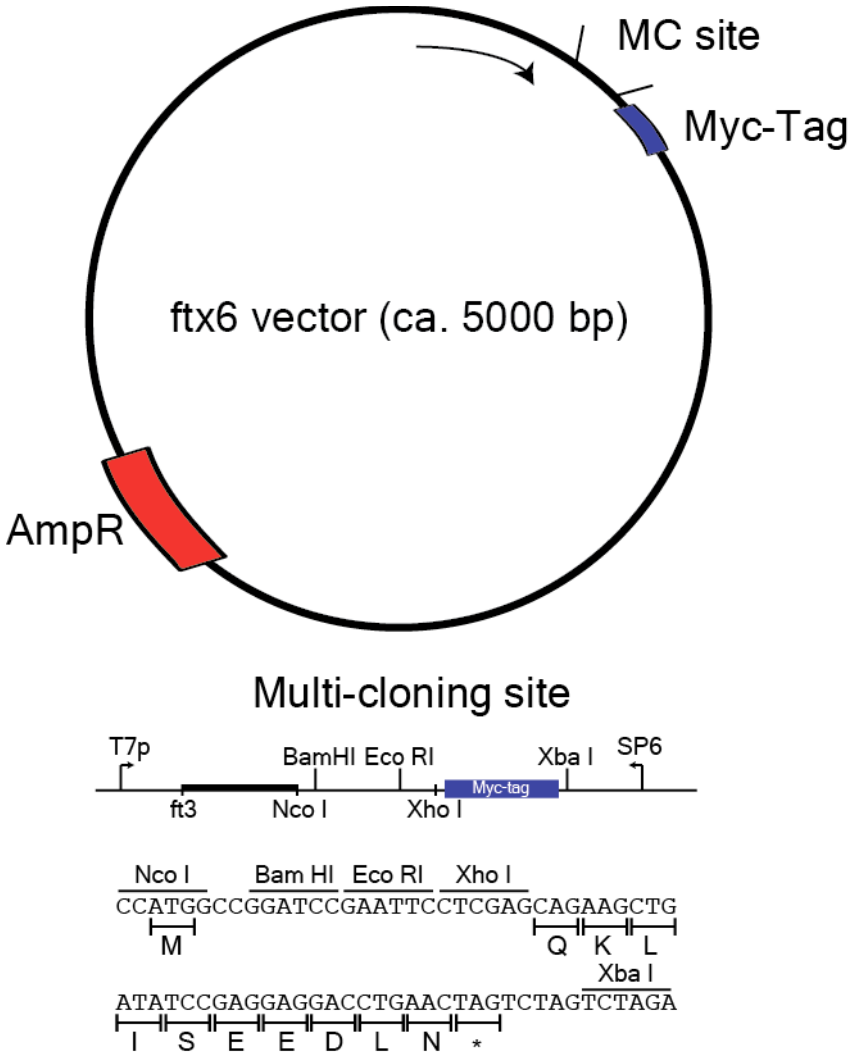


Figure A.3.3. ftx6 vector map. cDNA encoding for full-length human c-Src protein was subcloned into this vector using *Nco I* and *Eco RI* restriction sites. The plasmid was a gift from Nebreda's lab. The vector derives from the original pGEX6P1 and was modified for experimental purposes.

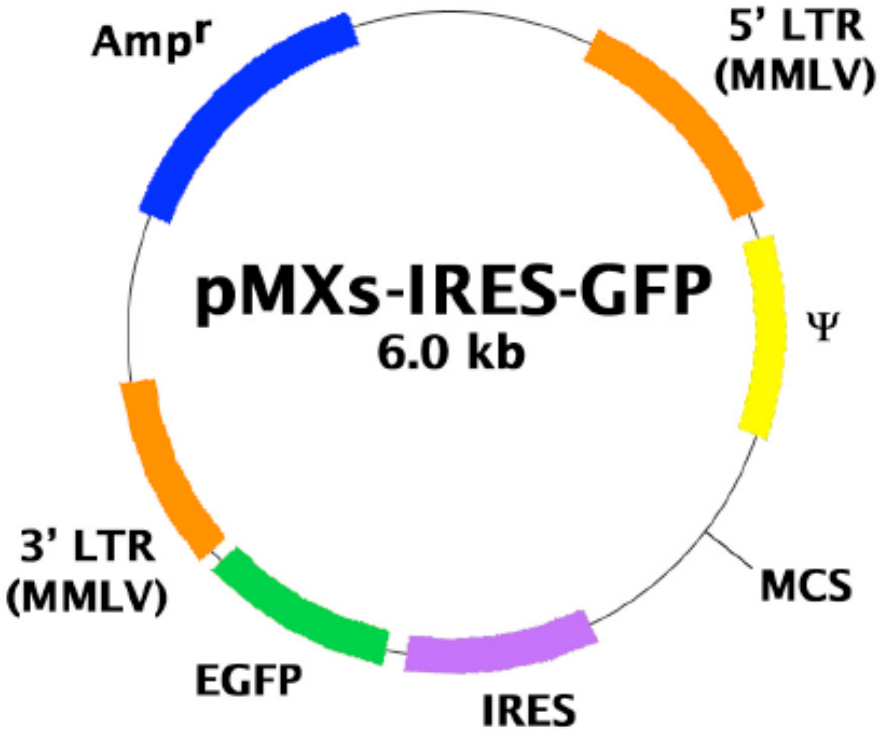


Figure A.3.4. pMX-pS-Cesar vector map. cDNA encoding for full-length human c-Src protein was subcloned into this vector using *Bam HI* and *Xho I* restriction sites. The plasmid was a gift from Roche's lab. The vector derives from the original pMXs-IRES-GFP and was modified for experimental purposes.

A.4 – Supporting Information

II.2 – Material and Methods: Mass Spectrometry

Proteins (20 μg) were digested by adding trypsin or chymotrypsin (2% w/v) and incubated at 37°C overnight. Digestions were stopped by adding formic acid to a final concentration of 1%. The resulting peptide mixtures were diluted in 1% FA and loaded in a nano-LC-MS/MS system. The nano-LC-MS/MS set up was as follows. Samples were loaded to a 180 μm × 2 cm C18 Symmetry trap column (Waters) at a flow rate of 15 μl/min using a nanoAcquity Ultra Performance LC™ chromatographic system (Waters Corp., Milford, MA). Peptides were separated using a C18 analytical column (BEH130™ 75 μm × 25 cm, 1.7 μm, Waters Corp.) with a 90 min run, comprising three consecutive steps with linear gradients from 1 to 35% B in 60 min, from 35 to 50% B in 5 min, and from 50 % to 85 % B in 3 min, followed by isocratic elution at 85 % B in 10 min and stabilization to initial conditions (A= 0.1% FA in water, B= 0.1% FA in CH3CN). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an LTQ-FT Ultra mass spectrometer (Thermo Scientific). Spray voltage in the NanoMate source was set to 1.70 kV. Capillary voltage and tube lens on the LTQ-FT were tuned to 40 V and 120 V. The spectrometer was working in positive polarity mode. At least two blank runs before each analysis were performed in order to ensure the absence of cross contamination from previous samples. Mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the FT with the resolution (defined at 400 m/z) set to 100,000. Up to six of the most intense ions per scan were fragmented and detected in the linear ion trap. The ion count target

value was 1,000,000 for the survey scans and 50,000 for the MS/MS scan. Target ions already selected for MS/MS were dynamically excluded for 30 s.

A database search was performed with Bioworks v3.1.1 SP1 and Proteome Discoverer software v1.3 (Thermo Scientific) using Sequest search engine and SwissProt database, which included USrc protein and the common repository of adventitious proteins (http://www.thegpm.org/crap/index.html). Search parameters included no-enzyme specificity, methionine oxidation and phosphorylation in serine and threonine as dynamic modification and, depending on the sample, amino acids labeled with 15N as static modification. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.8 Da. For Bioworks searches, peptides with a Sf higher than 0.85 and a peptide probability lower than 1E-3 were considered as positive identifications. For Proteome Discoverer searches, peptides with a q-value lower than 0.1 were considered as positive identifications. Extracted ion chromatograms of MS or MS/MS ions were obtained using Xcalibur software vs 2.0SR2 (Thermo Scientific).