Angel Vizoso Vázquez

Mechanisms of transcriptional regulation mediated by *IXR1* in yeast

Facultade de Ciencias
Departamento de Bioloxía Celular e Molecular
Área de Bioquímica e Bioloxía Molecular

Directora

Mª Esperanza Cerdán Villanueva

2015

El presente trabajo, **Mechanisms of transcriptional regulation mediated by *IXR1* in yeast**, presentado por Don Ángel José Vizoso Vázquez para aspirar al grado de Doctor en Biología, ha sido realizado bajo mi dirección en el Departamento de Biología Celular y Molecular de la Universidad de A Coruña.

Revisado el texto, estoy conforme con su presentación para ser juzgado.

A Coruña, 28 de Septiembre de 2015

Vº Bº

LA DIRECTORA DEL TRABAJO

Dra. Mª Esperanza Cerdán Villanueva

Catedrática de Bioquímica y Biología Molecular

Parte de la investigación presentada en esta tesis ha sido realizada durante una estancia de investigación en la Universidad de Cambridge en el año 2014. El trabajo de investigación realizado durante la estancia ha sido supervisado por la Dra. Jean Olwen Thomas, Catedrática en Bioquímica de la Universidad de Cambridge, Master del *St Catharine's College*, Presidenta de la *Royal Society of Biology* y miembro electo de la EMBO (*European Molecular Biology Organisation*).

También se realizó una estancia en el Instituto Rocasolano (CSIC-Madrid) en el año 2013, bajo la supervisión de la Dra. Juliana Sanz Aparicio, para la puesta a punto de distintas técnicas de cristalización y difracción de rayos X.

Part of the research included in this PhD thesis was been carried out at the University of Cambridge in 2014. This research was under the supervision of Dr. Jean Olwen Thomas, Professor in Biochemistry at University of Cambridge, Master del *St Catharine's College*, President of the *Royal Society of Biology* elected Member of the EMBO (*European Molecular Biology Organisation*).

Other stay at Rocasolano Institute (CSIC-MAdird) was done in 2013, under the supervision of Dra. Juliana Sanz Aparicio, in order to learn crystallization and X-ray diffraction techniques.

A mis padres

Esta parte de la tesis la he dejado para el final porque es la más especial, una carta abierta a todas las personas que me han acompañado y ayudado durante todo este tiempo. Y también la parte más dura, un "colorín colorado, este cuento se ha acabado" que se hace difícil de pronunciar después de todos estos años. Han sido años irrepetibles donde me llevo un montón de recuerdos y experiencias imborrables.

En primer lugar quisiera ofrecer mi más sincera gratitud a mi directora de tesis, la Dra. Mª Esperanza Cerdán, la oportunidad que me ha brindado. Gracias por tus enseñanzas, consejos, ánimos y sobre todo por hacerme sentir colega de profesión en vez de aprendiz, de trabajar codo con codo y confiar en mi. Estaré siempre agradecido.

También quisiera dar las gracias al resto de miembros del Área de Bioquímica, Mª Isabel González, Ana Rodríguez, Esther Rodríguez, Mª Angeles (Marián) Freire y, especialmente, al Dr. Manuel Becerra el haberme acogido en el laboratorio de Bioquímica en mis primeros pasos. Y por supuesto, a la Dra. Mónica Lamas, siempre dispuesta a arrimar el hombro cuando haga falta y de pelearse con el comercial que se ponga por delante. El laboratorio no funcionaría sin ti. Sigue así!

A la Dra. Mª Fernanda Rodríguez de la Unidad de Biología Molecular del SAI-UDC y al Dr. Luís Lombardía de la Unidad de Diagnóstico Molecular del CNIO, por su colaboración y ayuda en la preparación, hibridación y análisis de datos de los arrays, así como al Dr. Julián Yañez por su ayuda para la captura de imágenes de microscopia electrónica.

A mis compañeros y amigos, camaradas de horas y horas de trabajo, los que saben lo que cuesta llegar hasta aquí y los que lo están descubriendo, esto no lo podría haber logrado sin vosotros. GRACIAS. Durante estos años he aprendido que la investigación es un trabajo de equipo, donde se comparten alegrías, bromas,

anécdotas, frustraciones y enfados, estrechando lazos para formar una gran FAMILIA. Mi gratitud a Rafa, que siempre me has ayudado, enseñado, animado y hasta acogido en tu casa, pero sobre todo, siempre serás un ejemplo, alguien en quien me gustaría verme reflejado por tu tesón, organización y amor por está profesión. GRACIAS AMIGO. A Angelito, siempre te admiré por tu constante buen humor, daba gusto llegar temprano al laboratorio y ver de frente tu cara de alegría (o a las tantas de la madrugada…). A la gente del laboratorio que seguro leerán estas líneas, a "los de ahora", gracias. A Agustín, mi primer "discípulo" (permíteme el vocablo) que no salió corriendo, gracias por todas las risas y los buenos debates de política, un verdadero placer. A Juanjo, una persona humilde y sencilla que estoy convencido que llegará lejos en esto. A mis chicas, María, Mariu, Aída e, incluso Olalla, que aunque no le gusten estas "cosas", sé que en el fondo me lo agradecerás. Sois geniales, hacéis que ir al laboratorio sea divertido y ameno, con un montón de quedadas y actividades "extraescolares". Y a otra muchísima gente que ha pasado a mi lado durante la tesis y que no quiero olvidarme de ella: Ana García, Raquel (DEP), Pili, Lisi, Saúl, Marcos, Silvia, María, Marta, Fátima, Susana, Gerald, Stefan, Bernadette… y un largo etcétera, ha sido un placer compartir con todos vosotros todos estos años.

También quería tener unas palabras de agradecimiento a la gente que he conocido durante las dos estancias que realicé durante esta tesis. Por un lado a la Dra. Julia Sanz y al grupo de fenómenos de los que está rodeada en el Rocasolano, gracias a Mercedes, María, Antonio, Rocio, Elsa… por haberme acogido como uno más desde el primer día. Por otro lado, a la Dra. Jean O. Thomas, del departamento de bioquímica de la Universidad de Cambridge, y al resto de su pequeño gran equipo, la Dra. Katherine Stott y Joyce Ratti, por mostrar siempre un enorme interés por mi trabajo y por aportar ideas que hoy se reflejan en esta tesis.

A mi familia y amigos lo primero que quiero es pedirles perdón. Perdón por mis ausencias, por perderme muchos y seguro que buenos momentos por

estar "atrapado en el labo". Y gracias, gracias porque aún así, siempre me habéis comprendido y apoyado. A mi hermana, que siempre me ha animado y ha sufrido conmigo los palos que a veces te da la vida. A mis padres, a los que le debo todo, la vida, lo que tengo y lo que soy. A mi madre, de la que he aprendido que sin la humildad, el respeto y el trabajo duro no conseguiría nada. Este logro es tanto tuyo como mío. Y a mi padre, te fuiste pronto, pero admirabas mi tesón y constancia, y eso lo convertí en mi máxima en la vida. Sé que si estuvieras conmigo estarías orgulloso.

Por último a Lucía, mi Lu, la que tantas veces me ha levantado cuando más lo necesitaba, la que me animaba cuando las cosas se torcían y la que más ha sufrido con esta tesis, estas palabras son de ETERNO AGRADECIMIENTO. Te admiro por tu honestidad, alegría, fuerza y ganas de vivir. Gracias por cuidar de mi, por tu cariño y hacerme mejor persona. Te quiero.

GRACIAS!!!

5-FOA	*5-**F**luoro**O**rotic **A**cid*
μg	***microg**ram*
μL	***microL**iter*
μM	***microM**olar*
μmol	***micromol***
°C	***C**elsius grade*
AiBS	***A**NCHOR-**i**ndicated **B**inding **S**ites*
ALS	***A**myotrophic **L**ateral **S**clerosis*
Amp^R	***Amp**iciline **R**esistance*
Arg	***Arg**inine*
rRNA	***r**ibosomal **R**ibo**N**ucleic **A**cid*
ATP	***A**denosine **T**ri**P**hosphate*
BN-PAGE	***B**lue **N**ative- **P**oly**A**crylamide **G**el **E**lectrophoresis*
Bp	***B**ase-**p**air*
CD	***C**ircular **D**ichroism*
CDDP	***C**is-**D**iammine**D**ichloro**P**latinum (II)*
ChIP	***Ch**romatin **I**mmuno**P**recipitation*
CM	***C**omplete **M**edia*
Cys	***Cys**teine*
DAPI	*4',6-**DiA**midino-2-**P**henyl**I**ndole*
DEG	***D**ifferentially **E**xpressed **G**ene*
DISC	***D**eath-**I**nducing **S**ignaling **C**omplex*
DMSO	***D**i**M**ethyl **S**ulf**O**xide*
DNA	***D**eoxyribo**N**ucleic **A**cid*
dNTP	***d**esoxi**N**ucleotide **T**ri**P**hosphate*
DSF	***D**ifferential **S**canning **F**luorimetry*
DTT	***D**i**T**hio**T**hreitol*
EDTA	***E**thylene**D**iamine**T**etra**A**cetic acid*
EGFP	***E**nhanced **G**reen **F**luorescent **P**rotein*

Abbreviations

EMSA	*Electrophoretic Mobility Shift Assay*
FA	*Fluorescence Anisotropy*
FDA	*Food and Drug Administration*
FDR	*False Discovery Rate*
FE	*FRET Efficiency*
FPLC	*Fast Protein Liquid Chromatography*
FRET	*Förster Resonance Energy Transfer*
FTD	*FrontoTemporal lobar Degeneration*
g	*gram*
Gln	*Glutamine*
Gly	*Glycine*
GSH	*Glutathione reduced*
h	*hour*
HAS	*Human Serum Albumin*
HCA	*Hydrophobic Cluster Analysis*
HMG	*High Mobility Group domain*
HR	*Homology-directed DNA Repair*
HSQC	*Heteronuclear Single Quantum Coherence*
His	*Histidine*
ID	*IDentification number*
IDP	*Intrinsically Disorder Protein*
IDR	*Intrinsically Disorder Region*
IP	*InmunoPrecipitation*
IPTG	*IsoPropyl-β-D-1-ThioGalactopyranoside*
ITC	*Isothermal Tritation Calorimetry*
Kb	*Kilobase*
KDa	*KiloDalton*
L	*Litre*
LB	*Luria-Bertani*

Leu	***Leu**cine*
LIC	***L**igation **I**ndependent **C**lonning*
LN	***L**ewy **N**eurites*
Lys	***Lys**ine*
M	***M**olar*
MALDI-TOF	***M**atrix-**A**ssisted **L**aser **D**esorpsion **I**onization-**T**ime **O**f **F**light*
MEME	***M**ultiple **E**m for **M**otif **E**licitation*
mg	***m**ili**g**ram*
MG	***M**olten-**G**lobule*
min	***min**ute*
mL	***m**ili**L**iter*
mM	***m**ili**M**olar*
MMR	***M**is**M**atch **R**epair system*
MoRE	***Mo**lecular **R**ecognition **E**lements*
MRE	***M**ean **R**esidue **E**llipticity*
mRNA	***m**essenger **R**ibo**N**ucleic **A**cid*
MRP	***M**ultidrug **R**esistance-associated **P**roteins*
NAC	***N**on-**A**beta **C**omponent*
NER	***N**ucleotide **E**xcision **R**epair*
NFT	***N**euro**F**ibrillary **T**angle*
ng	***n**ano**g**ram*
NMR	***N**uclear **M**agnetic **R**esonance*
NO	***N**itric **O**xide*
NSS HMG-box	***N**on **S**equence **S**pecificity **H**igh **M**obility **G**roup **box** domain*
OD$_{600}$	***O**ptical **D**ensity at **600** nm*
ONPG	***O**rtho-**N**itro**P**henyl-β-**G**alactoside*
Ori	*replication **Ori**gin*

PCR	*Polymerase Chain Reaction*
PDB	*Protein Data Bank*
PEG	*PolyEthylene Glycol*
PFD	*Prion Forming Domain*
pH	*Hydrogen potential*
Phe	*Phenylalanine*
pI	*Isoelectric point*
PM	*Perfect Match*
PMG	*Pre-Molten Globule*
PMSF	*PhenylMethylSulfonyl Fluoride*
PrD	*Prion Domain*
PSSM	*Position-Specific Scoring Matrix*
Pt	*Platinum*
PTM	*Post-Translational Modification*
rpm	*revolutions per minute*
RC-like	*Random-Coil-like*
RIN	*RNA Integrity Number*
RNA	*RiboNucleic Acid*
ROS	*Reactive Oxygen Species*
RSAT	*Regulatory Sequence Analysis Tools*
S	*Svedberg unit*
SDD-AGE	*Semi-Denaturing Detergent-Agarose Gel Electrophoresis*
SDS-PAGE	*Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis*
SE-AUC	*Sedimentation Equilibrium-Analytical UltraCentrifugation*
sec	*second*
SEC	*Size Exclusion Chromatography*

SLC	*SoLute Carriers*
SS HMG-box	*Sequence Specificity High Mobility Group box domain*
SV-AUC	*Sedimentation Velocity-Analytical UltraCentrifugation*
TCEP	*Tris(2-CarboxyEthyl)Phosphine hydrochloride*
TCR	*Transcription-Coupled Repair*
TF	*Transcription Factor*
TFE	*2,2,2-TriFluoroEthanol*
Thr	*Threonine*
ThT	*Thioflavin T*
Tm	*melting Temperature*
Trp	*Tryptophan*
Tyr	*Tyrosine*
UCSC	*University of California Santa Cruz*
Ura	*Uracil*
UV	*UltraViolet light*
YEASTRACT	*Yeast Search for Transcriptional Regulators And Consensus Tracking*

Contents

Contents

Abstract

ABSTRACT

Ixr1 is a protein from *Saccharomyces cerevisiae* that was previously involved in the response to hypoxia and oxidative stress, as well as in cisplatin resistance. The work presented here allows expanding the knowledge regarding the regulatory mechanisms by which Ixr1 participates in these processes. Transcriptome analyses and experiments to identify the position and characteristics of Ixr1 binding sites along the genome were performed. They showed that Ixr1 participates in the regulation of sulphur assimilation, the metabolism of sulphur compounds, long chain amino acids and in diverse steps of ribosome biogenesis. The implication of Ixr1 on the control of cellular energetics is a crucial factor for explaining cellular adaptation to oxygen levels or cellular damage produced by cisplatin. Biochemical characterization of Ixr1, accompanied by functional analysis of different domains in the protein, also revealed that the binding to DNA is produced through its tandem HMG-box domains in a sequential model with positive cooperativity. Moreover, regions flanking the DNA binding domains show high structural disorder and they are prone to aggregate in amyloid fibrils, raising the possibility that Ixr1 could act as a prion.

RESUMEN

Ixr1 es una proteína de *Saccharomyces cerevisiae* previamente relacionada con la respuesta a hipoxia y a estrés oxidativo, así como en la resistencia a cisplatino. El trabajo presentado en esta Tesis permite ampliar el conocimiento de los mecanismos de regulación mediante los cuales Ixr1 participa en estos procesos. Se realizaron experimentos de transcriptómica y localización de los sitios de unión de Ixr1 a lo largo del genoma y sus características. Los resultados mostraron que Ixr1 participa en la regulación de la asimilación de azufre y en el metabolismo de compuestos que contienen azufre, de aminoácidos de cadena larga y en diversas etapas de la biogénesis de ribosomas. La implicación de Ixr1 en el control energético de la célula es un factor indispensable para explicar la adaptación de la célula a los niveles de oxígeno o al daño producido por cisplatino. La caracterización bioquímica de Ixr1, acompañada de análisis funcionales de distintos motivos de la proteína, mostró que su unión al ADN se produce mediante los HMG-box en tándem, según un modelo secuencial con cooperatividad positiva. Además, las regiones que flanquean a los dominos de unión a ADN presentan un alto grado de desorden estructural y con tendencia a la agregación en forma de fibras amiloides, lo que sugiere la posibilidad de que Ixr1 pueda actuar como prion.

RESUMO

Ixr1 é unha proteína de *Saccharomyces cerevisiae* anteriormente relacionada coa resposta á hipoxia e frente ó estrés oxidativo, así como na resistencia a cisplatino. O traballo presentado nesta Tese permite ampliar o noso coñecemento dos mecanismos de regulación polos cales Ixr1 participa nestes procesos. Experimentos de transcriptomica e localización dos sitios de unión de Ixr1 ao longo do xenoma e as súas características foron realizados. Os resultados mostraron que Ixr1 participa na regulación da asimilación de xofre e do metabolismo de compostos que conteñen xofre, aminoácidos de cadea longa e en varios estadios da bioxénese dos ribosomas. A implicación de Ixr1 no control da enerxética da célula é indispensable para explicar a adaptación da célula os niveis de osíxeno ou os danos por cisplatino. A caracterización bioquímica de Ixr1, acompañada polo análise funcional de varios motivos proteicos, mostrou que a súa unión ao ADN ocorre a través dos seus dominios HMG-box, seguindo un modelo secuencial con cooperatividade positiva. Ademais, as rexións que flanquean os dominos de unión ao ADN teñen un elevado grao de desorde estrutural e tendencia a agregar en forma de amiloides, suxerindo a posibilidade de que Ixr1 poida actuar coma un prion.

Introduction

Ixr1, object of this study, is an HMGB yeast protein that has been related to transcriptional regulation in response to stress signals like hypoxia or oxidants (Lambert *et al*. 1994; Bordineaud *et al*., 2000; Castro-Prego *et al*., 2010a; Castro-Prego *et al*., 2010b) and also participates in DNA repair after DNA damage induced by intra-strand crossing agents like cisplatin (Brown *et al*., 1993; Chow *et al*., 1994; McA'Nulty *et al*., 1996; McA´Nulty & Lippard, 1996; Huang *et al*., 2005; Rodriguez Lombardero *et al*., 2012). The amino acidic composition of Ixr1 is very peculiar, with large stretches of poly glutamine sequences, which condition disordered structures and promote aggregation. This characteristic composition allows us to hypothesize that this protein is prone to amyloid and prion formation, a process closely related to the regulation of gene expression.

1.- HMG proteins

Nucleosomes are fairly stable components of DNA packaging that constitutes the basic elements of chromatin. Nevertheless, the nucleosomal chromatin is surrounded by a highly mobile protein crowd that dynamically binds and modulates chromatin structure and function. The high mobility of nuclear proteins guarantees their rapid and permanent availability at diverse nuclear sites, including chromatin. Transient binding, interaction and competition create an overall shifting but stable network that regulates replication, DNA repair, gene transcription and chromatin remodelling. Among the crowd that constantly moves around and associates with nucleosomal chromatin are variants of the linker histone H1 family (Lever *et al*., 2000; Misteli *et al*., 2000; Kasinsky *et al*., 2001) and members of the high mobility group (HMG) protein superfamily (Bustin *et al*., 1999; Bianchi *et al*., 2005; Gerlitz *et al*., 2009).

1.1.- Classification of HMG proteins

The high-mobility group (HMG) proteins were discovered as nuclear factors more than 40 years ago and they were named for their high electrophoretic mobility in polyacrylamide gels (Goodwin *et al*., 1973). They are present in almost all metazoans and plants. Although HMG motifs are present in many nuclear proteins, the classification and nomenclature of the considered "canonical" HMG proteins is organized in 3 families named HMGA, HMGB, and HMGN, each one having a specific functional motif; the "AT-hook" in HMGA, the "HMG-box" in HMGB and the "nucleosomal binding domain" in HMGN (Bustin, 2001). In the nucleus, they act as non-histone architectural chromatin-proteins, but they also have other regulatory functions upon replication, transcription and DNA repair. Thus, HMG proteins are considered as architectural elements of chromatin. Recent findings support that HMG proteins may also play roles in epigenetics since their interaction with chromatin affects the level of histone modifications as was shown for HMGNs (Lim *et al*., 2004; Lim *et al*., 2005; Kim *et al*. 2009) and HMGBs (Lange *et al*., 2008). Moreover, they may be involved in nucleosome positioning and stability as is the case for HMGB1, which supports nucleosomal sliding (Bonaldi *et al*., 2002), and for HMGN1 and HMGN2, which stabilize nucleosomes by counterbalancing the action of ATP dependent chromatin remodelling machines (Rattner *et al*., 2009). Some HMG proteins have been related to extra-nuclear and extracellular functions during inflammation, cell differentiation, cell migration, and tumor metastasis (Hock, 2006).

All HMG proteins are characterized to share a common modular structure with defined DNA/chromatin binding motifs flanked by positively charged amino acids connected to a highly acidic tail that are involved in interactions with other proteins and regulate HMGs binding affinity (Catez & Hock, 2010). Binding of HMGs to their targets is indifferent to DNA sequences and, in general, all HMG proteins tend to recognize local peculiarities of chromatin rather than the underlying DNA

sequence. Nevertheless, some favorite binding sites, described for each HMG family, compartmentalize their distribution along the chromatin fiber (Catez & Hock, 2010). However, as dynamically binding and mobile proteins, all HMG proteins are able to spread along the DNA fiber outside their favorite binding sites. Whether they bind to restricted high-affinity binding sites or spread out among the chromatin is most likely dependent on their cellular amounts. HMG proteins are among the most abundant chromatin binding proteins in undifferentiated proliferating cells, are differentially expressed during differentiation and absent in terminally differentiated cells (Muller *et al*., 2004; Hock *et al*., 2007).

1.2.- Characteristics of HMGB family

The structures of various HMG-boxes are well studied and it is known that their folding is far more conserved than the amino acid sequences. In general, the HMG-box domain contains 65-85 amino acids and has a characteristic L-shaped fold formed by three α-helices with an angle of ≈80° between the two arms. The long arm or minor wing is composed by the extended N-terminal strand and third α-helix, while first and second α-helix form the short arm, or major wing (Figure 1a).

There are two broad subfamilies of HMG-box containing proteins, based on structural and phylogenetic studies. One class includes those that bind to distorted DNA with low or without sequence specificity (*Non Sequence Specificity*, NSS, HMG-box domains) (Grosschedl *et al*., 1994; Bustin and Reeves, 1996) and have, in general, two or more in tandem arranged HMG-box domains. Examples of proteins without sequence specificity are the mammalian HMGB1-4 and UBF proteins, HMGD from *Drosophila* or Nhp6 from *Saccharomyces cerevisiae*. Their role is related to chromatin modification, participating in many different functions such as co-activation or silencing of transcription and V(D)J junction recombination. A second class of HMG-box containing proteins bind to DNA by recognizing a

specific DNA sequence (*Sequence Specificity*, SS, HMG-box domains) (Grosschedl *et al*., 1994; Bustin and Reeves, 1996) and they usually contain a single HMG-box domain. They generally function as transcription factors, only expressed in a few cell types, and they also contain other regulatory associated domain. The determinants for DNA sequence specificity lie mainly in the minor wing of the HMG-box. Examples of this kind of HMGB proteins are the mammalian lymphoid enhancer factor (Lef-1), the sex determining factor (Sry) and the Sry-related HMG-box (SOX) family; or the hypoxic gene repressor (Rox1) from *Saccharomyces cerevisiae*.

Despite these differences, both subfamilies of HMG-box proteins are able to bind to B-form DNA through the minor groove with high affinity and they induce a large DNA bending, ultimately forming complexes of rather similar structure. They use their concave surface to intercalate one or two (SS and NSS HMG-box domains, respectively) bulky hydrophobic amino acids between base-pairs in the minor groove. Other extensive protein-DNA contacts, not sequence-specific, as well as hydrogen bonds are made with the phosphate backbone. As a result of these interactions, the DNA is bent and under-wound with positive roll angles, widening the minor groove and compressing the major groove (Figure 1b).

Saccharomyces cerevisiae has seven genes expressing HMGB proteins: *ABF2*, *HMO1*, *NHP6A*, *NHP6B*, *NHP10*, *IXR1* and *ROX1* (Bustin, 2001). The structural characteristics and functions of these yeast proteins are shown in Table 1. Only one HMG-box domain is present in five of them, but Abf2 and Ixr1 have two in tandem "HMG-box" motifs.

16
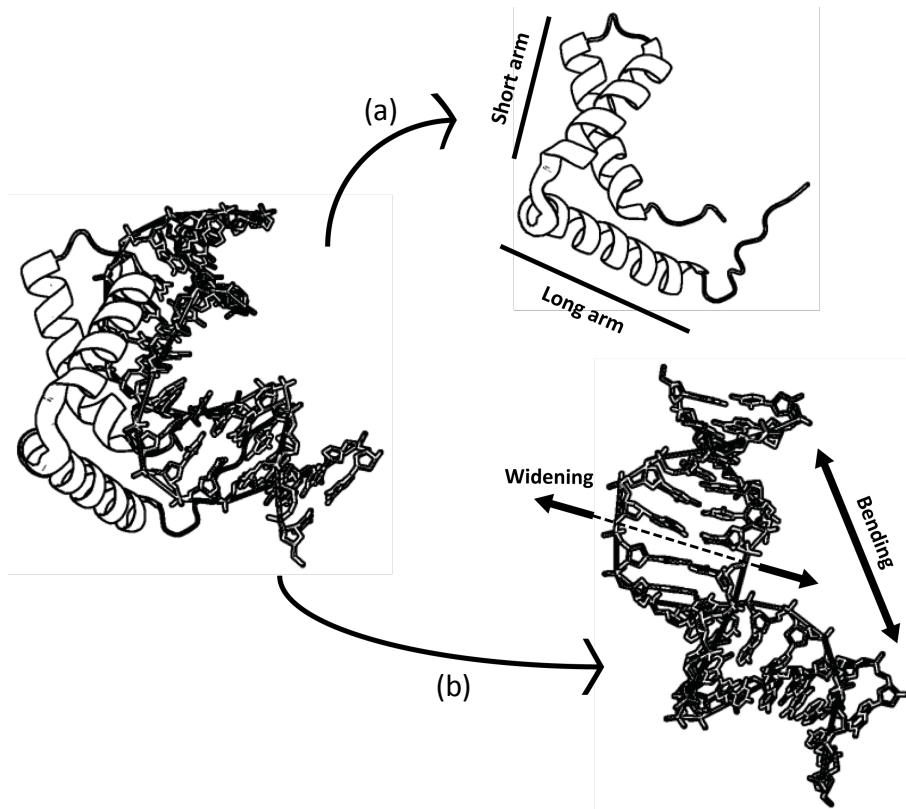
Figure 1. (a) Characteristic L-shaped fold based on Sox17 protein structure (PDB ID: 3F27). (b) Bending and widening produced in the double strand of DNA.

2.- The hypoxic response in yeast and the role of HMGB proteins

2.1.- Wide Genome Duplication (WGD) and adaptation to low oxygen levels

Due to the increased levels of atmospheric oxygen over the past 3.5 billion years, prokaryotes and eukaryotes have developed multiple optimization strategies related to the use of oxygen. Classical mitochondrial respiratory chain is part of a catabolic pathway involved in oxidative phosphorylation. Thus, oxidation of substrates, the respiratory electron transport complexes (I, II, III) and the final reduction of oxygen in water for cytochrome *c* oxidase (complex IV) results in the

synthesis of ATP through a chemo-osmotic coupling (Beauvoit & Rosenfeld, 2003) (Figure 2).

Assimilative sugar sources vary greatly among species (Barnett, 1976), but glucose remains as the universal carbon source in yeast. In *S. cerevisiae*, facultative aerobic yeast, glucose is primarily directed to the glycolytic fermentation, in detriment of respiration pathways, strongly repressed by glucose (positive *Crabtree effect*). This phenomenon can play a crucial role in the balance between fermentation and respiration when cells grow in oxygen-limited environments. However, this glucose repression does not occur in other yeast species such as *Kluyveromyces lactis* or *Pichia stipitis*, capable of catabolize glucose through bot respiration and fermentation simultaneously.
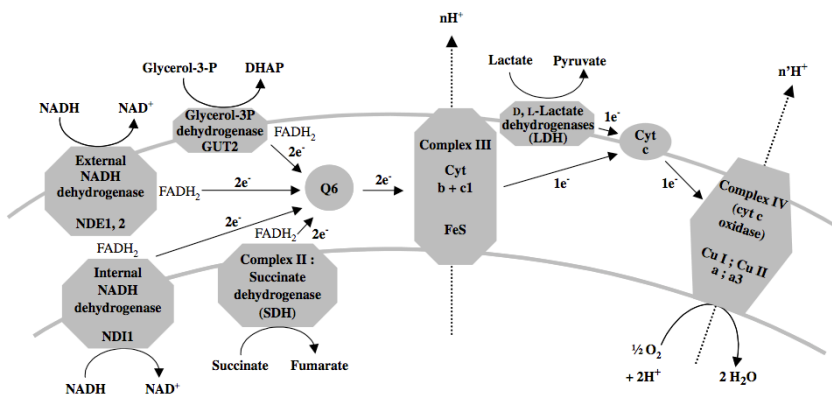
Figure 2. Scheme of the respiratory chain in *Saccharomyces cerevisiae* (Rosenfeld & Beauvoit, 2003). Only inner mitochondrial membrane is represented. Cytochrome *b*, subunits I, II and III of the cytochrome *c* oxidase are respiratory components encoded by mitochondrial genes (as subunits VI, VIII and IX of the Fo-ATPase, not represented). The two types D- and L- of lactate dehydrogenase are not detailed. Q6: ubiquinone 6; Cyt: cytochrome.

Pyruvate is metabolized in part by fermentation through pyruvate decarboxylase and in part by oxidation through pyruvate dehydrogenase (Breuning *et al.*, 2000). Other yeast species such as *Yarrowia lipolytica*, are exclusively

respiratory, unable to ferment (Kurtzman & Fell, 1998). Although most yeast species identified are able to grow under limiting oxygen conditions, by fermenting sugars to ethanol and carbon dioxide, only a few can grow in complete absence of oxygen (Sunnerhagen & Piskur, 2006). This is the case of *S. cerevisiae*, which can grow rapidly in both environments.

Adaptation of yeast to an anaerobic environment has an evolutionary component. Several data indicate that the ability of the *Saccharomyces* genus to grow in anaerobic environments is due to a large genome duplication (*Whole Genome Duplication*, WGD) which took place about 100 million years ago (Piskur, 2001). Species such as *K. lactis*, which comes from a common ancestor with *S. cerevisiae* before the genomic duplication, cannot grow in the absence of oxygen. As a consequence of WGD, this genus acquired the ability to develop new capabilities, either by increased gene dosage, where the presence of additional copies can confer selective advantage without function diverging in the *loci* (for example the chaperones *SSB1/SSB2* and *HSP82/HSC82*); or by a neofunctionalization process, wherein the second gene copy acquires a new different function from the original one. This latter process is well known in the aerobic/ anaerobic response. For example, the gene couples *CYC1/CYC7* or *COX5a/COX5b* of the mitochondrial electron transport chain, the *HMG1/HMG2* iso-enzymes of the sterol biosynthesis pathway or the *AAC2/AAC3* isoforms of ADP/ATP translocation. Transcriptomic experiments using DNA arrays showed that up to 25% of the genes (called ohnologs) present duplication due to the WGD (Sunnerhagen & Piskur, 2006), having at least one member showing differential expression as a function of oxygen levels (Kwast *et al*., 2002).

To achieve a metabolic adaptation to these two alternative states, i.e. the presence and absence of oxygen, there are genes that are differentially expressed (Zitomer & Lowry, 1992). Nuclear and mitochondrial genes coding for proteins involved in oxidative phosphorylation, as cytochromes and subunits of the

respiratory chain, and enzymes for protection against oxidative damage, as catalase or superoxide dismutases, are only expressed under aerobic conditions and turn repressed when oxygen concentrations are low (Rosenfeld & Beauvoit, 2003). Conversely, there are genes that are slightly expressed in normoxia and are induced when oxygen availability decreases. These are called hypoxic genes and they are classified into four groups according to their cellular functions (Kwast *et al*., 2002), which are related to cell wall, cellular response to stress, carbohydrate metabolism and genes related to the metabolism of lipids, fatty acids and isoprenoids.

Frequently, cellular heme levels mediate the signaling of oxygen availability. Heme biosynthesis requires molecular oxygen in two consecutive reactions, first as an electron acceptor for oxidative decarboxylation and after for the oxidation of two methylene groups to methenyl (Zagorec & Labbe-Bois, 1986). The first of these two steps is limited under hypoxic conditions and is catalyzed by the enzyme coproporphyrinogen III oxidase, encoded by the hypoxic gene *HEM13* (Klinkenberg *et al*., 2005). Heme is a prosthetic group of several proteins, some of which are transcriptional factors, whose activity and expression are dependent of intracellular heme levels. Five genes, *HAP1*- 2, 3, 4, *5*, participate in the activation of gene expression in response to heme. *HAP1* encodes a transcription factor that plays a central role in the response to oxygen levels through heme.

2.2.- Rox1, an HMG box-protein, as a key regulator in the aerobic repression of hypoxic genes

Up to one third of hypoxic genes are transcriptionally repressed during aerobic growth by Rox1 through the recruitment of the general repression complex Ssn6/Tup1. Rox1 is a DNA binding protein with an HMG-box domain that binds to the consensus sequence **YYYATTGTTCTC** present in the promoter regions of genes related to hypoxia, causing a DNA bending of 90° in the double strand (Deckert *et*

al., 1999; Xin *et al*, 2000). *ROX1* expression is dependent of intracellular oxygen and heme levels, and since its expression is under the control of Hap1 (Keng, 1992), ROX1 is therefore transcriptionally induced aerobically (Deckert *et al*., 1998). In addition to up-regulation produced by Hap1, *ROX1* is closely regulated by self-repression under aerobic conditions to avoid cellular toxic effects produced by an eventual over-expression. At low oxygen levels, the Rox1 protein is rapidly degraded, since it is labile, and besides, the *ROX1* gene is no longer transcribed. Under normoxic (aerobic) conditions, the heme-activated Hap1 complex increases *ROX1* expression, allowing in turn that Rox1 represses hypoxic genes. In hypoxia the situation is reversed, as the low levels of Rox1 allow derepression of these genes.

The genes that are under the control of Rox1, either directly by the protein binding to their promoter regions, or indirectly through signal transduction pathways, are those related to efficient metabolism under low oxygen levels, ergosterol and heme synthesis, cell wall maintenance or electron chain transport (Ter Linde & Steensma, 2002).

A cross regulation between Rox1 and Ixr1 has been reported (Castro-Prego *et al*., 2010b) and Ixr1 also takes part in the transcriptional regulation in response to hypoxia (Lambert *et al*. 1994; Bordineaud *et al*., 2000; Castro-Prego *et al*., 2010a;)

3.- Cisplatin and the connection to HMGB proteins

3.1.- Mechanisms of action

Cisplatin, also called *cis*-diamminedichloroplatinum (II) (CDDP), is a metallic (platinum) coordination compound with a square planar geometry, slightly soluble in water and soluble in dimethylprimanide and N,N-dimethylformamide. The *cis* configuration is required for its antitumor activity. It has two labile chloride

groups and two relatively inert amine ligands. It was first synthesized by Michele Peyrone in 1844 and its chemical structure was first elucidated by Alfred Werner in 1893. It was the first FDA-approved platinum compound for cancer treatment in 1978 (Kelland, 2007). Since then, cisplatin is one of the most used chemotherapy drugs for the treatment of multiple other solid neoplasms, including head and neck, lung, colorectal and ovarian cancers (Rosenberg *et al.*, 1965; Kelland, 2007).

Nowadays, nine platinum analogs are currently in clinical trials around the world termed ormaplatin (tetraplatin), oxaliplatin, DWA2114R, enloplatin, lobaplatin, CI-973 (NK-121), 254-S, JM-216, and liposome-entrapped cis-bis-neodecanoato-trans-R,R-1,2-diaminocyclohexane platinum (II) (LNDDP) (Weiss and Christian, 1993; Frezza *et al.*, 2010). Figure 3 presents the chemical structures of cisplatin and four of its analogs including carboplatin, oxaliplatin and ormaplatin.
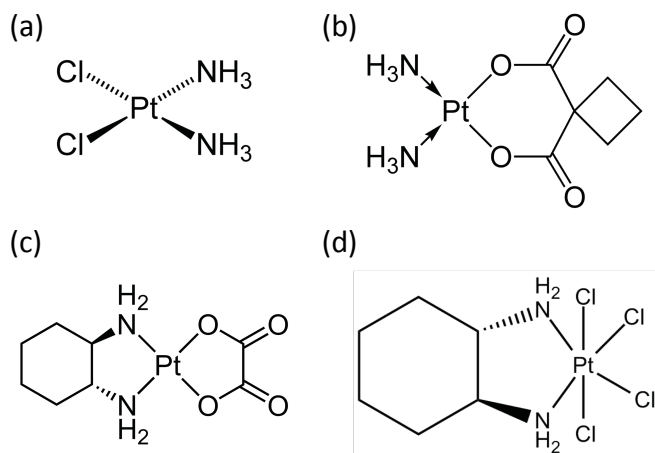
Figure 3. Chemical structures of platinum drugs cisplatin (a), carboplatin (b), oxaliplatin (c) and ormaplatin (d).

During the past, it was believed that cisplatin get access to the cytoplasmic compartment by passively diffusing across the plasma membrane. Nowadays, it is known that majority of cisplatin is actively moved in and out of the cell by copper transporters. In mammals, there are several membrane transporters

of platinum compounds, some dependent on ATP such as Mdr1 and other MRPs (Multidrug Resistance-associated Proteins) or analogous (Atp7A/B). There are also solute carrier importers, independent of ATP, that transport cisplatin by passive or facilitated diffusion, such as Ctr1 and other SLCs (SoLute Carriers), Aqp2 or Aqp9. Depending on the cellular context, multiple transporters may be involved in cisplatin uptake. Therefore, it is difficult to correlate cisplatin sensitivity/resistance with a particular transporter.

Cisplatin becomes activated once it enters the cell. In the cytoplasm, water molecules displace the chloride atoms on cisplatin. Once hydrolyzed, activated cisplatin is a strong electrophile that can react with any nucleophile, including the sulfhydryl groups on proteins and nitrogen donor atoms on nucleic acids. Cisplatin binds to the N7 reactive center on purine residues to form cisplatin adducts, a critical lesion that disrupts the structure of the DNA and interferes with DNA replication and transcription. The 1,2-intrastrand cross-links of purine bases with cisplatin are the most frequent changes in DNA, that includes 1,2- intrastrand d(GpG) and 1,2-intrastrand d(ApG) adducts, and represent about 90% and 10% of adducts, respectively. The 1,3-intrastrand d(GpXpG) adducts and other adducts such as inter-strand cross-links and nonfunctional adducts have been reported to contribute to cisplatin toxicity (Dasari & Tchounwou, 2014). This alteration in the structure is recognized by cellular proteins implied in the repair of cisplatin-induced DNA damage. Cisplatin lesions are primarily repaired via NER (Nucleotide Excision Repair) system, a multilayered process that includes epigenetic, transcriptional and posttranslational regulation. In addition to NER, cisplatin can also induce TCR (Transcription-Coupled Repair). The intra-strand crosslink stalls RNA polymerase II to trigger TCR (Damsma *et al*., 2007). It has been reported that p53 protects against cisplatin-induced apoptosis in a TCR-dependent manner (McKay *et al*., 2001). In addition, the homology-directed DNA repair (HR) that allows error-free repair of the double-strand breaks caused by the excision of cisplatin-DNA adducts has been implicated in the repair of cisplatin-induced DNA

damage (Borst *et al.*, 2008). Finally, MMR (MisMatch Repair) system also recognizes cisplatin-induced DNA damage, but instead of increasing cell viability, MMR system was shown to be important for cisplatin-mediated cytotoxicity (Sedletska *et al.*, 2007).

Although the major target of cisplatin is the nuclear DNA, other cell biomolecules are important targets of cisplatin. Proteins contain several potential reactive sites for platinum (Histidine, Methionine, and Cysteine). In this sense, the reaction of activated cisplatin with sulfur containing groups is a kinetically favored process, forming stable Pt-S bonds (Pinato *et al.*, 2014). Reactions on proteins can occur at different stages along the long pathway that these drugs need to complete to exert their therapeutic DNA platination. One of the first potential targets is the human serum ambumin (HAS), the most abundant protein in plasma, since cisplatin is generally administered intravenously. Indeed, nephrotoxicity effects that accompany cisplatin treatment are related to the HAS oligomerization caused by the platinum compound (Pinato *et al.*, 2013).

Once in the cytoplasm, the platinum compound become even more reactive because of chloride concentration drops further, enhancing cisplatin hydration. In this sense, synthesis and accumulation of glutathione (GSH) and metallothioneins are raised, and other important components actively involved in controlling cellular redox homeostasis have also been found to be directly targeted by the platinum complex (i.e. the thioredoxin reductase system whose functions are impaired). The platinum atom in cisplatin is chelated by glutathione (GSH) and the glutathione-Pt complex effluxes out the cell in an ATP-dependent process by members of the glutathione transporter family, termed the GS-X pumps (Ishikawa & Ali-Osman, 1993). Metallothioneins are cysteine-rich proteins (around 20 cysteines in peptides of 61-68 amino acid length) involved in zinc and copper homeostasis, heavy metal detoxification, protection from apoptosis and that are able to interact with cisplatin (Basu & Lazo, 1990). The cisplatin-caused redox

imbalance is actually a beneficial promotion of the cytotoxic activity of the drug. However, the same unbalanced redox processes have been associated with side effects such as nephrotoxicity and hepatotoxicity (Jungwirth *et al.*, 2011).

Cellular sensitivity to cisplatin is not only regulated by its uptake, efflux or interaction with its targets DNA or proteins, but also cellular responses to cisplatin-induced DNA damage play a major role in deciding the ultimate cell fate. Following DNA damage, cell cycle checkpoints are activated to delay cell-cycle progression to provide time for DNA repair or eliminate genetically unstable cells by inducing cell death. Furthermore, it is now recognized that inhibition of DNA replication is not sufficient to explain cisplatin cytotoxicity. The tumor suppressor protein p53, considered the "guardian of genome", play a critical role in the cisplatin response, trans-activating genes involved in cell cycle arrest (e.g., p21), DNA repair (e.g., growth arrest and DNA damage-inducible 45, Gadd45), and apoptosis (e.g., Bax) (De Laurenzi & Melino, 2000). Also, the tyrosine kinase c-Abl plays an important role in stress response to DNA damaging agents, and it is activated upon the recognition of cisplatin- induced DNA damage by the MMR causing activation of JNK/SAPK (c-Jun-N-terminal kinase/stress- activated protein kinase) (Kharbanda *et al.*, 1995). c-Abl also acts in cooperation with p53, p73 (Gong *et al.*, 1999) and p38 MAPK pathway (Galan-Moya *et al.*, 2008) to trigger cisplatin-induced apoptosis.

Although the molecular mechanisms that underlie the cytotoxic potential of cytoplasmic cisplatin are poorly understood, it is known that includes the accumulation of reactive oxygen species (ROS), nitric oxide (NO), and cell apoptosis. Apoptosis is a controlled type of cell death, which is energy-dependent leading to cell shrinkage, chromatin condensation, membrane budding, phosphatidylserine externalization, and activation of a family of cysteine proteases called caspases (Salvesen & Dixit, 1997; Cummings *et al.*, 2000). There are two major pathways of cell death mediated by caspases; the first is mediated by TNF-α (Tumor Necrosis Factor α) receptor superfamily and recruitment of procaspase-8

to form the death-inducing signaling complex (DISC); the second is mediated by the Bcl-2 family proteins, which regulate cytochrome-*c* release from mitochondria, thus activating procaspase-9 through the interaction with Apaf-1 (Apoptosis Promoting Activating Factor-1).

3.2.- Medical importance of Cisplatin resistance

In spite of the general effectiveness of cisplatin, one of the major problems of cisplatin treatment is the intrinsic or acquired resistance that many cancer cell types show. Cisplatin is highly efficient only against testicular germ cell cancer, leading to a durable complete remission in >80% of the patients (Winter & Albers, 2011). On the contrary, the clinical responses showed by cisplatin-based chemo-radio therapeutic regimens in patients affected by other solid tumors (e.g., ovarian carcinoma) are temporary and malignant cells become chemo-resistant. Others, like lung, prostate or colorectal cancer, entail neoplastic lesions that are intrinsically resistant to the cytostatic/cytotoxic activity of cisplatin.

The proposed mechanisms of cisplatin resistance include changes in cellular uptake and efflux of cisplatin, increased biotransformation and detoxification in the liver, and increase in DNA repair and anti-apoptotic mechanisms (Gottesman *et al.*, 2002). In this sense, cisplatin resistance has been classified in four phases: first, those alterations in processes that precede the binding of cisplatin to its actual targets, including DNA and cytoplasmic structures, and called "pre-target resistance"; second, those alterations directly related to the molecular damage produced by cisplatin and called "on-target resistance"; third, those alterations in the lethal signalling pathways triggered by cisplatin lesions and called "post-target resistance"; and fourth, those changes that alter some molecular circuit that are not intimately associated with main cisplatin- elicited signals and called "off-target resistance" (Gottesman *et al.*, 2002; Galluzzi *et al.*, 2014).

Among "pre-target resistance" reasons are the reduced amount of "reactive" cisplatin forms in the cytoplasm, the extrusion of the drug, mainly by the multidrug resistance-associated protein 2 (*MRP2*), and the increased levels of metallothioneins and glutathione, or enzymes related to glutathion synthesis, such as g-glutamylcysteine synthetase or glutathione S-transferase (Kelley *et al*., 1988; Lewis *et al*., 1988; Kasahara *et al*., 1991; Koike *et al*., 1997; Cui *et al*., 1999; Korita *et al*., 2010; Chen *et al*., 2010).

On-target resistance refers to a DNA repair system that becomes specially-efficient. In particular, the nucleotide excision repair (NER) system is believed to fix the majority of cisplatin-DNA adducts, although components of the mismatch repair (MMR) machinery have also been implicated in this process (Chaney & Sancar, 1996; Furuta *et al*., 2002). Also, double-strand break lesions that sometimes produce cisplatin are normally repaired by homologous recombination (Smith *et al*., 2006). Accordingly, homologous recombination-proficient neoplasms are generally more resistant to cisplatin treatment than homologous recombination-deficient cells, such as those bearing loss-of-function mutations in the genes encoding Brca1 and Brca2, (Venkitaraman, 2002). On the other hand, only a few extra-nuclear cisplatin-binding partners have been identified to contribute to cisplatin resistance, including cytosolic (e.g., myosin IIa, Hsp90), ribosomal (e.g., ribosomal protein L5), reticular (e.g., calreticulin) and mitochondrial components (e.g., mitochondrial DNA, Vdac1), as well as others ubiquitously spread like Vcp (valosin-containing protein) a type II member of AAA+-ATPase family (Sancho-Martinez *et al*., 2012; Karasawa *et al*., 2013).

Post-target cisplatin resistance refers to the set of alterations in the cell checkpoints mechanisms that detect the cisplatin damage and convert it into a lethal signal, as well as in the components of the regulation and signalling pathways that promotes the activation of the cell death via apoptotic or necrotic systems (Gottesman *et al*., 2002). The cell reaction to cisplatin treatment is

promoting a rapid activation of an integrated adaptive response aimed at the re-establishment of cellular homeostasis, avoiding the cell death path. Only when homeostasis cannot be restored, because stress conditions are excessive in intensity or duration, lethal signals are transmitted, a survival mechanism for the preservation of organism homeostasis (Gottesman *et al.*, 2002). These signals consist in the switch of the DNA damage response from a cytoprotective to a cytotoxic mode, followed by the activation of Bax and Bak1*,* and also the accumulation of ROS and consequent PTPC (Permeability Transition Pore) opening (Mandic *et al.*, 2001). Both these processes eventually promote increased permeability of the mitochondrial outer membrane, in turn resulting in the functional and physical breakdown of mitochondria, followed by the activation of caspase-dependent and independent mechanisms of cell death (Galluzi *et al.*, 2012). Thus, post-target cisplatin resistance has been associated not only with genetic and epigenetic alterations that impair p53 signalling, but also with defects in several other pro-apoptotic signal transducers, including p38 MAPK pathway and JNK1/SAPK (Mansouri *et al.*, 2003).

Finally, the susceptibility of cancer cells to cisplatin can also be limited by off-target mechanisms, that is, molecular pathways that deliver compensatory survival signals even though they are not directly activated by cisplatin, some mediated by the AKT1 signalling pathway, such as Dyrk1b or Tmem205 proteins (Gottesman *et al.*, 2002).

To overcome resistance, cisplatin is commonly used in combination with some other drugs in treating ovarian cancer, biliary tract cancer, lung cancer (diffuse malignant pleural mesothelioma), gastric cancer, carcinoma of salivary gland origin, breast, colon, lung, prostate, melanoma and pancreatic cancer cell lines, squamous cell carcinoma of male genitial tract, urothelial bladder cancer, and cervical cancer (Dasari & Tchounwou, 2014).

3.3.- HMGB proteins and cisplatin

The cascade of cellular and molecular events that produces cisplatin treatment, such as impair transcription and replication, cell-cycle arrest to provide a time frame for DNA repair, or apoptosis when the damage extent exceeds repair capacities, involve a wide range of proteins that interact with DNA lesions or with subsequent DNA distortions and constitute the ''platinated DNA interactome'' (Bounaix Morand du Puch *et al*., 2011). Platinum modification distorts the structure of duplex DNA in a distinctive manner. A variety of cellular proteins specifically recognize these uniquely altered structural forms of DNA. These proteins include those involved in repair processes like XPC-hHR23B, XPA, RPA, and TFIIH, that recognize platinum adducts cooperatively during the early stage of NER (Riedl *et al*., 2003; Dip *et al*., 2004); hMutSR heterodimer (MSH2- MSH6) and bacterial MutS damage recognition proteins in MMR (Mello *et al*., 1996; Yamada *et al*., 1998); the DNA-dependent protein kinase (DNA-PK), that participates in cellular DNA repair processes such as double-strand break (DSB) restoration (Turchi & Henkels, 1996); or the human 3-methyladenine DNA glycosylase (AAG), a damage recognition protein involved in base excision repair, selectively binds to various cisplatin adducts (Kartalou *et al*., 2000). Other proteins not related to repair systems that preferentially recognize cisplatin-modified DNA are the TATA-binding protein (TBP) (Vichi *et al*., 1997), the tumor suppressor protein p53 (Wetzel & Berberich, 2001), the Poly(ADP-ribose) polymerase 1 (PARP-1) (Zhang *et al*., 2004), Y-box binding protein-1 (YB-1) (Ise *et al*., 1999), Structural maintenance of chromosome protein 3 (SMC3), Chromatin-specific transcription elongation factor (SPT16), among others (Jung & Lippard, 2007; Bounaix Morand du Puch *et al*., 2011).

Nevertheless, the group of proteins that have long been known to interact with cisplatin DNA is the HMGB protein family, particularly hHMGB1 (Scovell *et al*., 1987; Bruhn *et al*., 1992). High-mobility group protein 1 (HMGB1) is one of the

early proteins discovered to bind cisplatin-modified DNA (Hughes *et al*., 1992; Pil & Lippard, 1992). HMGB1 is an abundant and highly conserved non-histone chromosomal protein of 30-kDa protein and 215 amino acids, comprises two HMG box domains A and B and an acidic C-terminal tail and that preferentially binds to DNA with bent or distorted structures. Each HMG domain, as well as the full-length HMGB1 protein, binds selectively to cisplatin-modified DNA (Pil & Lippard, 1992; Dunham & Lippard, 1997). Although the two HMG box domains of HMGB1 are structurally similar and positioned in tandem, domain A interacts more strongly with cisplatin 1,2-intrastrand DNA crosslinks than domain B (Dunham & Lippard, 1997). As a NSS DNA binding protein, it regulates numerous nuclear functions including transcription, replication, recombination, and general chromatin remodeling, serving as an architectural facilitator by assisting the assembly of nucleoprotein complexes (Thomas & Travers, 2001). For the past several years, HMGB1 has also been investigated as an extracellular mediator, performing significant roles in inflammation, differentiation, migration, tumor metastasis, and the immune response (Dumitriu *et al*., 2005). This variety in the properties of HMGB1, suggesting a likely involvement of the protein in the cisplatin mechanism of action, also make the biological repercussions of HMGB1 binding still controversial, as some published works demonstrated a subsequent facilitation of DNA repair whereas others suggest a shielding effect towards NER factors. Firstly, *in vitro* studies demonstrated that HMGB1 could inhibit reparation of 1,2-intrastrand DNA crosslinks by the NER system, presumably by binding to and shielding the damage site from recognition by the repair apparatus (Huang *et al*., 1994; Zamble *et al*., 1996). Supporting these results, an increased protein level of HMGB1 following hormone treatment sensitizes breast cancer cells to cisplatin by a factor of 2 (He *et al*., 2000), and also the additional expression of HMGB2 (85% identity with HMGB1) in human lung cancer cells enhanced cisplatin sensitivity more than 3-fold (Arioka *et al*., 1999). On the other hand, HMGB1 is overexpressed in various cisplatin-resistant cell lines and has been identified as a proapoptotic

signaling protein (Brezniceanu *et al.*, 2005). Even mouse embryonic native and HMGB1 knockout cell lines show no significant differences in their sensitivity to cisplatin (Wei *et al.*, 2003). Then, it seems that the ability of HMGB1 to impact the cytotoxicity of cisplatin can depend upon the cell type, the experimental method used to change the protein level, and possibly even the growth conditions and number of passages of the cells.

Many other proteins containing one or more HMG domains bind to cisplatin-modified DNA, including the structure-specific recognition protein Ssrp1 (Bruhn *et al.*, 1992), the ribosomal RNA transcription factor hUBF (Treiber *et al.*, 1994), the mitochondrial transcription factor A mtTFA (Jung & Lippard, 2007), the lymphoid enhancer binding protein Lef-1 (Chow *et al.*, 1994), the sex-determining factor Sry (Trimmer *et al.*, 1998), testis-specific HMG protein tsHMG (Ohndorf *et al.*, 1997), the drosophila homologue of HMGB1 HMG-D (Churchill *et al.*, 1995), the TOX high mobility group box family member 4 Tox4 (Bounaix Morand du Puch *et al.*, 2011), and the yeast proteins Cmb1 (Fleck *et al.*, 1998), Nhp6 (Wong *et al.*, 2002) or Ixr1 (McA'Nulty *et al.*, 1996). Indead, Ixr1 participates in DNA repair after DNA damage induced by intra-strand crossing agents like cisplatin (Brown *et al.*, 1993; Chow *et al.*, 1994; McA'Nulty *et al.*, 1996; McA´Nulty & Lippard, 1996; Huang *et al.*, 2005).

4.- Prions and gene regulation in yeast

4.1.- General description of prions

Prions are self-replicating protein bodies that are involved in the spreading of several mammalian neurodegenerative diseases, variously known as Kuru, scrapie, and bovine spongiform encephalopathy, in humans, sheep and cows, respectively (Aguzzi *et al.*, 2008). However, most prions and their properties have been discovered and studied in lower organisms, particularly in the yeast *Saccharomyces cerevisiae* (table 2).

Table 2. Known and candidate prions

Prion determinant	Prion state	Organism	Protein function	Consequences of prion state
Native prions PrP	PrP^{Sc}	Mammals	Neuronal growth and maintance	Neurodegeneration and death
Ure2	[URE3]	*S. cerevisiae* and related yeasts	Represses transcription of nitrogen catabolic genes	Indiscriminate utilization of nitrogen sources
Sup35	[*PSI*$^+$]	*S. cerevisiae* and related yeasts	Translation termination	Increased nonsense suppression, translation frameshifting, changes in mRNA stability
Rnq1	[*PIN*$^+$]	*S. cerevisiae*	Unknown	Increased appearance of other prions
HET-s	[*Het-s*]	*P. anserina*	Heterokaryon incompatibility	Inhibit phusion between [Het-s] and het-S mycelia
Swi1	[*SWI*$^+$]	*S. cerevisiae*	Transcription regulation	Altered carbon source utilization
Mca1	[MCA]	*S. cerevisiae*	Regulation of apoptosis, cell cycle progression	Unknown
Cyc8	[*OCT*$^+$]	*S. cerevisiae*	Transcription repression	Altered carbon source utilization, flocculation
Mot3	[*MOT3*$^+$]	*S. cerevisiae*	Transcription regulation	Altered cell wall composition
Pma1/Std1	[*GAR*$^+$]	*S. cerevisiae*	Plasma membrane proton pump (Pma1) and glucose signaling (Std1)	Indiscriminate utilization of carbon sources
Candidate prions CPEB, neuronal isoforms	-	*A. californica*	Translation regulation of synapse-specific mRNAs	Localized protein synthesis at activated synapses; maintains long-term facilitation
Sfp1	[*ISP*$^+$]	*S. cerevisiae*	Transcriptional regulation of ribosomal protein and biogenesis genes	Neutralisation of nonsense suppressors
19 other proteins	-	*S. cerevisiae*	Diverse	Undetermined

It is in general accepted that prions have a benign or even beneficial role in the yeast cell, acting as epigenetic elements that increase phenotypic diversity in a heritable way, increasing survival in front of changing and diverse environmental conditions (True *et al*., 2004; Alberti *et al*., 2009). Thus, prions show reversible switching capabilities among diverse phenotypes and promote the evolution of such phenotypic novelty.

Amyloids represent the most frequent, and in consequence, the most biochemically characterized state of prions (Glover *et al*., 1997; Alberti *et al*., 2009), but exist other types of self-propagating protein conformations that may also result in the prion phenomena (Wickner *et al*., 2007; Brown & Lindquist, 2009). Amyloid is a highly ordered, fibrillar protein aggregate with a unique set of biophysical characteristics that facilitate prion propagation: extreme stability, assembly by nucleated polymerization, and a high degree of templating specificity. Prion spread cycle begins with a single nucleating event that occurs from and within a stable intracellular population of non-prion conformers of the same protein. This prion nucleus is then elongated into a fibrillary conformation by addition of new non-prion conformers (Serio *et al*., 2000; Tessier & Lindquist, 2009). In final steps, the growing protein fibers fragment into smaller propagating units, which are spread among daughter cells (Shorter & Lindquist, 2005). Because the change in protein conformation results in alterations of function, these self-perpetuating conformational changes create heritable phenotypes associated to the determinant protein and its genetic background. Prion phenotypes are dominant in genetic crosses and exhibit non-Mendelian inheritance patterns. In this sense, prion-based genetic elements are denoted with capital letters and brackets – "[*PRION*]".

4.2.- Concept of prion as "bet-hedging device"

Bet-hedging device refers to its ability to allow simple organisms to switch spontaneously between distinct phenotypic states (True & Lindquist, 2000). Prions

raise the reproductive fitness through the creation of several subpopulations with distinct phenotypic states in organisms exposed to fluctuating environments (Seger & Brockmann, 1987). Until recently, the number of known proteins capable to form prions has been short, and important aspects of their crucial roles in adaptation and evolution were ignored. However, numerous discoveries in yeast last years have enormously expanded the prion knowledge. Proteins that were recently discovered to form prions include several chromatin remodeling and transcription factors (Du *et al*., 2008; Alberti *et al*., 2009; Patel *et al*., 2009), and others prionogenic proteins whose prion states and their roles into the cell need yet to be examined (Alberti *et al*., 2009).

4.3.- Concept of prion as "evolutionary capacitor"

An evolutionary capacitor is any entity that normally hides the effects of genetic polymorphisms, allowing for their storage in a silent form, and releases them in a sudden stepwise fashion (Masel & Siegal, 2009). In this sense, some of these phenotypes produced by the expression of accumulated genetic variation on occasion will be beneficial to the organism, and occasionally further genetic and epigenetic variations would accumulate and stabilize the beneficial phenotype as a consequence of cell proliferation. In accordance with this, it has been proposed that prions have the ability to act as evolutionary capacitors (Shorter & Lindquist, 2005). For example, the lower translation fidelity as a consequence of [PSI^+] (prion state of Sup35), results in the translation of previously silent genetic information through a variety of mechanisms including stop-codon read-through and ribosome frameshifting (Liebman & Sherman, 1979; True *et al*., 2004; Wilson *et al*., 2005; Namy *et al*., 2008). Furthermore, stop-codon read-through allows also changes in mRNA stabilities because of a relaxed selection in untranslated regions and cryptic RNA transcripts under normal [psi^-] conditions. As a consequence, transcripts are free to accumulate genetic variations, to finally affect their genetic expression. Upon the appearance of [PSI^+], these polymorphisms become phenotypically

expressed.

The epigenetic role of prions in its evolutionary capacitance is essential and necessary, providing a mechanism for the persistence, and if it is appropriate, genetic assimilation of the revealed phenotypes (Masel & Siegal, 2009). Then, phenotypes that appear as a consequence of prion formation could be more advantageous by other capacitors like regulatory network responses, where spontaneity and persistence for multiple generations could premium over transience and punctual stress response.

It is interesting to note that among prionogenic proteins there are overrepresented important components of gene expression control, cell signalling and the response to stimuli such as stress (Table 2) (Alberti *et al*., 2009). Many of them represent highly connected nodes in the genetic network of yeast. The Swi1 chromatin remodeler, for instance, regulates the expression of 6% of the yeast genome (Du *et al*., 2008). Likewise, Cyc8 represses 7% of the yeast gene complement (Green & Johnson, 2004). The prion candidates Pub1, Ptr69 and Puf2 are members of a family of RNA-binding proteins that regulate the stability of hundreds of mRNAs encoding functionally related proteins (Hogan *et al*., 2008). This large enrichment of putative prions among proteins with key regulatory and signaling functions suggests that prion-based switches evolve preferentially among proteins with a strong impact on multiple downstream biological processes. Moreover, the existence of a genetic polymorphisms background whose expression is altered by these prions would create different and complex phenotypes in front of different environmental situations upon which natural selection can act.

4.4.- Prion switching regulated by stress conditions

Prion formation is a special type of the protein misfolding process and requires of environmental stresses to perturb protein stability. Indeed, the frequency of prion switching is affected by environmental factors (Avery, 2006).

Protein quality control machinery is essential for both normal protein folding and for cellular stress responses. Prion formation is strongly influenced by several components of the ubiquitin-proteasome system (Chernoff, 2007). Furthermore, prion propagation requires the actions of members of the Hsp40, Hsp70, and Hsp110 chaperone families as well as the AAA+ protein dis-aggregase Hsp104 (Chernoff, 2007; Sweeny & Shorter, 2008). Hsp104 is a member of the ClpA/ClpB family of chaperones, whose members are found throughout bacteria, fungi, plants and eukaryotic mitochondria. Hsp104 provides thermotolerance by re-solubilizing stress-induced protein aggregates, and also has the unique ability to sever amyloid fibers into new prion propagons. This property has been conserved for hundreds of millions of years of fungal evolution (Zenthon *et al*., 2006).

The distribution of proteins between soluble and aggregated states is highly sensitive to the status of the protein homeostasis network, which includes protein synthesis, folding, sorting, and degradation machinery (Morimoto, 2008). In this sense, chaperones take part in protein homeostasis creating interconnected networks to act as important transducers of the stress response (Morimoto, 2008). Thus, alterations in the abundance, availability, and connectivity of chaperones like Hsp104 and Hsp70s in response to stress conditions may therefore be influencing the prion conformational switching and propagation capacities, reflecting the ancestral chaperone involvement in the relationship between environment and phenotype.

4.5.- Intrinsically conformational disorder of HMGB proteins as possible source of Prion candidates

Until now, there are no HMGB proteins among the prion-prone proteins described. Proteins with prion propensities are, in general, associated with the presence of intrinsically disordered regions (IDR) that can aggregate orderly.

Transcriptional factors are essential protein "hubs" for controlling many

aspects of biological activity and is responsible for the binding diversity of the broad cascade of protein-protein interactions, since one of the major functional advantages for intrinsically disordered proteins is the ability to bind to multiple different targets without sacrificing specificity to form the flexible nets [Dunker *et al*., 2005]. Intrinsically Disordered Proteins (IDPs) can bind to multiple targets due to its structural plasticity, which allows for IDPs to adopt several conformations (Wright and Dyson, 2009). Liu and coworkers have shown that 82% of eukaryotic transcription factors contain intrinsically disordered regions (Liu *et al*., 2006). The intrinsic disorder may have evolved to overcome the thermodynamic and kinetic challenges in sequence specific binding of DNA (Dyson and Wright, 2005). The flexibility found in transcriptional factors is utilized during transcription activation where complexes of many different proteins need to be formed and changes in the chromatin structure occurs.

HMG-box domains are characterized to possess high flexibility. Using calorimetric measurements, the HMG-box of Sox-5 from mouse have shown that significant levels of protein refolding occur on association, in addition to the DNA bending (Crane-Robinson *et al* 1998; Privalov *et al*. 1999). The undergoing DNA-dependent order-disorder transition of HMG-box domains appears to play an important role for the adaptability of the angular surface of the domain (figure 1) to enable a HMGB protein to induce different architectures in different functional contexts. Thus, context-dependent changes in overall architecture may differentially affect transcription, and a single transcription factor may exert fine control over relative levels of expression within a set of target genes.

Besides DNA-binding domains, remote IDRs in the transcription factors can also affect DNA binding. Fuzzy interactions of IDRs with the DNA-binding domain can, in addition to protein-protein interactions and PTMs of IDRs, change the affinity by modulating the flexibility or the conformation preferences of residues in the DNA-protein interface. That is the case of human HMGB1, in which

the acidic tail of its carboxylic side organizes the HMG boxes and linkers into an "auto-inhibited" complex with the DNA-binding faces of the HMG-box domains are hidden and not available. When functional partners interact with hHMGB1 produce a conformational change to promote an open "binding competent" form of the protein, releasing and leaving exposed and available their HMG-box domains to bind to the DNA (Watson el al., 2014).

Studies about the priogenic nature of HMGB proteins in yeast have not been reported and, in this sense, our study about the existence of IDRs in Ixr1 and their function in amyloid formation is pioneer.

5.- REFERENCES

Aguzzi, A., Baumann, F., and Bremer, J. (2008). "The prion's elusive reason for being." Annual review of neuroscience **31**:439-477.

Alberti, S., R. Halfmann, O. King, A. Kapila, and S. Lindquist (2009). "A systematic survey identifies prions and illuminates sequence features of prionogenic proteins." Cell **137**(1): 146-158.

Arioka H., K. Nishio, T. Ishida, H. Fukumoto, K. Fukuoka, T. Nomoto, H. Kurokawa, H. Yokote, S. Abe, N. Saijo (1999). "Enhancement of cisplatin sensitivity in high mobility group 2 cDNA-transfected human lung cancer cells." Japanese Journal of Cancer Research **90**(1):108-15.

Avery S.V. (2006). "Microbial cell individuality and the underlying sources of heterogeneity." Nature Reviews Microbiology **4**(8):577-87. Review.

Basu A., J.S. Lazo (1990). "A hypothesis regarding the protective role of metallothioneins against the toxicity of DNA interactive anticancer drugs." Toxicology Letters **50**(2-3):123-35

Bianchi M.E., A. Agresti (2005). "HMG proteins: dynamic players in gene regulation and differentiation." Current Opinion in Genetics & Development **15**(5):496-506.

Bonaldi T., G. Längst, R. Strohner, P.B. Becker, M.E. Bianchi (2002). "The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding." EMBO Journal **21**(24):6865-73.

Bounaix Morand du Puch C., E. Barbier, A. Kraut, Y. Couté, J. Fuchs, A. Buhot, T. Livache, M. Sève, A. Favier, T. Douki, D. Gasparutto, S. Sauvaigo, J. Breton (2011). "TOX4 and its binding partners recognize DNA adducts generated by platinum anticancer drugs." Archives of Biochemistry and Biophysics **507**(2):296-303."

Bourdineaud J.P., G. De Sampaïo, G.P. Lauquin (2000). "A Rox1-independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the SRP1/TIR1 gene." Molecular Microbiol

38(4):879-90.

Borst P., S. Rottenberg, J. Jonkers (2008). "How do real tumors become resistant to cisplatin?." Cell Cycle **7**(10):1353-9.

Brezniceanu M.L., K. Völp, S. Bösser, C. Solbach, P. Lichter, S. Joos, M. Zörnig (2003). "HMGB1 inhibits cell death in yeast and mammalian cells and is abundantly expressed in human breast carcinoma." FASEB Journal **17**(10):1295-7.

Brown S.J., P.J. Kellett, S.J. Lippard (1993) "Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin" Science **261**(5121):603-605

Brown, J.C., and S. Lindquist (2009). "A heritable switch in carbon source utilization driven by an unusual yeast prion." Genes & Development **23**(19): 2320-2332.

Bruhn S.L., P.M. Pil, J.M. Essigmann, D.E. Housman, S.J. Lippard (1992). "Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin." Proceedings of the National Academy of Sciences of the United States of America **89**(6):2307-11.

Bustin, M. and R. Reeves (1996). "High-mobility-group chromosomal proteins. Architectural components that facilitate chromatin function." Progress in Nucleic Acid Research and Molecular Biology **54**: 35-100.

Bustin M. (1999). "Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins." Molecular and Cellular Biology **19**(8):5237-46.

Bustin M. (2001) "Revised nomenclature for high mobility group (HMG) chromosomalproteins" Trends in Biochemical Sciences **26** (3) 152-153.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso, M.E. Cerdán (2010a) "Ixr1p regulates oxygen-dependent *HEM13* transcription" FEMS Yeast Research **10**(3):309-321

Castro-Prego M., M. Lamas-Maceiras, P. Soengas, I. Carneiro, M.I. González Siso, M.E. Cerdán (2010b) "Regulatory factors controlling transcription of *Saccharomyces cerevisiae IXR1* (*ORD1*) by oxygen levels. A model of transcriptional adaptation from aerobiosis to hypoxia implicating *ROX1* and *IXR1* cross-regulation" Biochemical Journal **425**(1):235-243

Catez F., R. Hock (2010). "Binding and interplay of HMG proteins on chromatin: lessons from live cell imaging." Biochimie et Biophysic Acta **1799**(1-2):15-27.

Chaney S.G., A. Sancar (1996). "DNA repair: enzymatic mechanisms and relevance to drug response." Journal of the National Cancer Institute **88**(19):1346-60. Review.

Chen H.H., M.T. Kuo (2010). "Role of glutathione in the regulation of cisplatin resistance in cancer chemotherapy." Metal-Based Drugs **2010**: Article ID: 430939.

Chernoff Y.O. (2007). "Stress and prions: lessons from the yeast model." FEBS Letters **581**(19):3695-701.

Chow C.S., J.P. Whitehead, S.J. Lippard (1994). "HMG domain proteins induce sharp bends in cisplatin-modified DNA." Biochemistry **33**(50):15124-30.

Churchill M.E., D.N. Jones, T. Glaser, H. Hefner, M.A. Searles, A.A. Travers (1995). "HMG-D is an architecture-specific protein that preferentially binds to DNA

containing the dinucleotide TG." EMBO Journal **14**(6):1264-75.

Crane-Robinson C., C.M. Read, P.D. Cary, P.C. Driscoll, A.I. Dragan, P.L. Privalov (1998). "The energetics of HMG box interactions with DNA. Thermodynamic description of the box from mouse Sox-5." Journal of Molecular Biology **281**(4):705-17.

Cui Y., J. Konig, J.K. Buchholz, H. Spring, I. Leier, D. Keppler (1999). "Drug resistance and ATP- dependent conjugate transport mediated by the apical multidrug resistance protein, MRP2, permanently expressed in human and canine cells." Molecular Pharmacology **55**(5): 929-937.

Cummings B.S., J.M. Lasker, L.H. Lash (2000). "Expression of glutathione-dependent enzymes and cytochrome P450s in freshly isolated and primary cultures of proximal tubular cells from human kidney." Journal of Pharmacology and Experimental Therapeutics **293**(2):677-85.

Damsma G.E., A. Alt, F. Brueckner, T. Carell, P. Cramer (2007). "Mechanism of transcriptional stalling at cisplatin-damaged DNA." Nature Structural & Molecular Biology **14**(12):1127-33.

Dasari S., P.B. Tchounwou (2014). "Cisplatin in cancer therapy: molecular mechanisms of action." European Journal of Pharmacology **740**:364-78.

De Laurenzi V., G. Melino (2000). "Apoptosis. The little devil of death." Nature **406**(6792):135-6.

Deckert J., R.A. Khalaf, S.M. Hwang, R.S. Zitomer (1999). "Characterization of the DNA binding and bending HMG domain of the yeast hypoxic repressor Rox1." Nucleic Acids Research **27**(17):3518-26.

Dip R., U. Camenisch, H. Naegeli (2004). Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. DNA Repair **3**(11):1409-23.

Du, Z., K. W. Park, H. Yu, Q. Fan, and L. Li (2008). "Newly identified prion linked to the chromatin-remodeling factor Swi1 in *Saccharomyces cerevisiae*." Nature Genetics **40**(4): 460-465.

Dumitriu I.E., P. Baruah, A.A. Manfredi, M.E. Bianchi, P. Rovere-Querini (2005). "HMGB1: guiding immunity from within." Trends in Immunology **26**(7):381-7.

Dunham S.U., S.J. Lippard (1997). "DNA sequence context and protein composition modulate HMG-domain protein recognition of cisplatin-modified DNA." Biochemistry **36**(38):11428-36.

Dunker A.K., M.S. Cortese, P. Romero, L.M. Iakoucheva, V.N. Uversky (2005). "Flexible nets. The roles of intrinsic disorder in protein interaction networks." FEBS Journal **272**(20):5129-48.

Dyson, H.J., P.E. Wright (2005). "Intrinsically unstructured proteins and their function." Nature Reviews Molecular Cell Biology **6**(3):197-208.

Fleck O., C. Kunz, C. Rudolph, J. Kohli (1998). "The high mobility group domain protein Cmb1 of *Schizosaccharomyces pombe* binds to cytosines in base mismatches and opposite chemically altered guanines." Journal of Biological Chemistry **273**(46):30398-405.

Frezza, M., S. Hindo, D. Chen, A. Davenport, S. Schmitt, D, Tomco, Q.P. Dou (2010). "Novel metals and metal complexes as platforms for cancer therapy." Current Pharmaceutical Design. **16**(16):1813-1825.

Furuta T., T. Ueda, G. Aune, A. Sarasin, K.H. Kraemer, Y. Pommier (2002). "Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells." Cancer Research **62**(17):4899-902.

Galan-Moya E.M., J. Hernandez-Losa, C.I. Aceves Luquero, M.A. de la Cruz-Morcillo, C. Ramírez-Castillejo, J.L. Callejas-Valera, A. Arriaga, A.F. Aranburo, S. Ramón y Cajal, J. Silvio Gutkind, R. Sánchez-Prieto (2008). "c-Abl activates p38 MAPK independently of its tyrosine kinase activity: Implications in cisplatin-based therapy." International Journal of Cancer **122**(2):289-97.

Galluzzi L., O. Kepp, C. Trojel-Hansen, G. Kroemer (2012). "Mitochondrial control of cellular life, stress, and death." Circulation Research **111**(9):1198-207.

Gerlitz G., R. Hock, T. Ueda, M. Bustin (2009). "The dynamics of HMG protein-chromatin interactions in living cells." Biochemistry and Cellular Biology **87**(1):127-37.

Glover, J. R., A. S. Kowal, E. C. Schirmer, M. M. Patino, J. J. Liu and S. Lindquist (1997). "Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*." Cell **89**(5): 811-819.

Goodwin GH, C. Sanders & E.W. Johns (1973). "A new group of chromatin-associated proteins with a high content of acidic and basic amino acids" European Journal of Biochemistry **38**(1):14-9.

Gottesman M.M., T. Fojo, S.E. Bates (2002). "Multidrug resistance in cancer: role of ATP-dependent transporters." Nature Reviews Cancer **2**(1): 48-58.

Green, S.R. and A.D. Johnson (2004). "Promoter-dependent roles for the Srb10 cyclin-dependent kinase and the Hda1 deacetylase in Tup1- mediated repression in *Saccharomyces cerevisiae*." Molecular Biology of the Cell **15**(9), 4191-4202.

Grosschedl R., K. Giese, J. Pagel (1994). "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures" Trends in Genetics **10**(3):94-100. Review.

He Q., C.H. Liang, S.J. Lippard (2000). "Steroid hormones induce HMG1 overexpression and sensitize breast cancer cells to cisplatin and carboplatin." Proceedings of the National Academy of Sciences of the United States of America **97**(11):5768-72.

Hock R, T. Furusawa, T. Ueda, M. Bustin (2007). "HMG chromosomal proteins in development and disease" Trends in Cell Biology **17**(2):72-9.

Hogan, D.J., D.P. Riordan, A.P. Gerber, D. Herschlag and P.O Brown (2008). "Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system." PLoS Biology **6**(10), e255

Huang J.C., D.B. Zamble, J.T. Reardon, S.J. Lippard, A. Sancar (1994). "HMG-domain proteins specifically inhibit the repair of the major DNA adduct of the anticancer drug cisplatin by human excision nuclease." Proceedings of the National Academy of Sciences of the United States of America **91**(22):10394-8.

Huang R.Y., M. Eddy, M. Vujcic, D. Kowalski (2005). "Genome-wide screen

identifies genes whose inactivation confer resistance to cisplatin in *Saccharomyces cerevisiae*." Cancer Research **65**(13):5890-7.

Hughes E.N., B.N. Engelsberg, P.C. Billings (1992). "Purification of nuclear proteins that bind to cisplatin-damaged DNA. Identity with high mobility group proteins 1 and 2." Journal of Biological Chemistry **267**(19):13520-7.

Ise T., G. Nagatani, T. Imamura, K. Kato, H. Takano, M. Nomoto, H. Izumi, H. Ohmori, T. Okamoto, T. Ohga, T. Uchiumi, M. Kuwano, K. Kohno (1999). "Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen." Cancer Research **59**(2):342-6.

Ishikawa T., F. Ali-Osman (1993). "Glutathione-associated cis-diammine dichloroplatinum(II) metabolism and ATP-dependent efflux from leukemia cells. Molecular characterization of glutathione-platinum complex and its biological significance." Journal of Biological Chemistry **268**(27):20116-25.

Jung Y., S.J. Lippard (2007). "Direct cellular responses to platinum-induced DNA damage." Chemical Reviews **107**(5):1387-407.

Jungwirth U., C.R. Kowol, B.K. Keppler, C.G. Hartinger, W. Berger, P. Heffeter (2011). "Anticancer activity of metal complexes: involvement of redox processes." Antioxidants & Redox Signals **15**(4):1085-127.

Karasawa T., M. Sibrian-Vazquez, R.M. Strongin, P.S. Steyger (2013). "Identification of cisplatin-binding proteins using agarose conjugates of platinum compounds." PLoS One **8**(6):e66220.

Kartalou M., L.D. Samson, J.M. Essigmann (2000). "Cisplatin adducts inhibit 1,N(6)-ethenoadenine repair by interacting with the human 3-methyladenine DNA glycosylase." Biochemistry **39**(27):8032-8.

Kasahara K., Y. Fujiwara, K. Nishio, T. Ohmori, Y. Sugimoto, K. Komiya, T. Matsuda, N. Saijo (1991). "Metallothionein content correlates with the sensitivity of human small cell lung cancer cell lines to cisplatin." Cancer Research **51**(12):3237-42.

Kasinsky H.E., J.D. Lewis, J.B. Dacks, J. Ausio (2001). "Origin of H1 linker histones, FASEB Journal." **15**(1):34-42.

Kelland L. (2007). "The resurgence of platinum-based cancer chemotherapy." Nature Reviews Cancer **7**(8):573-584.

Kelley S.L., A. Basu, B.A. Teicher, M.P. Hacker, D.H. Hamer, J.S Lazo (1988). "Overexpression of metallothionein confers resistance to anticancer drugs." Science **241**(4874):1813-5.

Keng T. (1992). "*HAP1* and *ROX1* form a regulatory pathway in the repression of HEM13 transcription in *Saccharomyces cerevisiae.*" Molecular Cell Biology **12**(6):2616-23.

Kharbanda S., P. Pandey, R. Ren, B. Mayer, L. Zon, D. Kufe (1995). "c-Abl activation regulates induction of the SEK1/stress-activated protein kinase pathway in the cellular response to 1-beta-D arabinofuranosylcytosine." Journal of Biological Chemistry **270**(51):30278-81.

Kim Y.C., G. Gerlitz, T. Furusawa, F. Catez, A. Nussenzweig, K.S. Oh, K.H. Kraemer,

Y. Shiloh, M. Bustin (2009). "Activation of ATM depends on chromatin interactions occurring before induction of DNA damage." Nature Cell Biology **11**(1):92-6.

Klinkenberg L.G., T.A. Mennella, K. Luetkenhaus, R.S. Zitomer (2005). "Combinatorial repression of the hypoxic genes of *Saccharomyces cerevisiae* by DNA binding proteins Rox1 and Mot3." Eukaryotic Cell **4**(4):649-60.

Koike K., T. Kawabe, T. Tanaka, S. Toh, T. Uchiumi, M. Wada, S. Akiyama, M. Ono, M. Kuwano (1997). "A canalicular multispecific organic anion transporter (cMOAT) antisense cDNA enhances drug sensitivity in human hepatic cancer cells." Cancer Research **57**(24): 5475-5479.

Korita P.V., T. Wakai, Y. Shirai, Y. Matsuda, J. Sakata, M. Takamura, M. Yano, A. Sanpei, Y. Aoyagi, K. Hatakeyama, Y. Ajioka (2010). "Multidrug resistance-associated protein 2 determines the efficacy of cisplatin in patients with hepatocellular carcinoma." Oncology Reports **23**(4):965-972.

Kwast K.E., L.C. Lai, N. Menda, D.T. 3rd James, S. Aref, P.V. Burke (2002). "Genomic analyses of anaerobically induced genes in *Saccharomyces cerevisiae*: functional roles of Rox1 and other factors in mediating the anoxic response." Journal of Bacteriology **184**(1):250-65.

Lambert J.R., V.W. Bilanchone, M.G. Cumsky (1994) "The *ORD1* gene encodes a transcription factor involved in oxygen regulation and is identical to *IXR1*, a gene that confers cisplatin sensitivity to *Saccharomyces cerevisiae*" Proceedings of the National Academy of Sciences United States of America **91**(15):7345-7349.

Lange S.S., D.L. Mitchell, K.M. Vasquez (2008). "High mobility group protein B1 enhances DNA repair and chromatin modification after DNA damage." Proceedings of the National Academy of Sciences United States of America **105**(30):10320-5.

Lewis A.D., J.D. Hayes, C.R. Wolf (1988). "Glutathione and glutathione-dependent enzymes in ovarian adenocarcinoma cell lines derived from a patient before and after the onset of drug resistance: intrinsic differences and cell cycle effects." Carcinogenesis **9**(7):1283-1287.

Liebman S.W., F. Sherman (1979). "Extrachromosomal psi+ determinant suppresses nonsense mutations in yeast." Journal of Bacteriology **139**(3):1068-71.

Lim J.H., F. Catez, Y. Birger, K.L. West, M. Prymakowska-Bosak, Y.V. Postnikov, M. Bustin (2004). "Chromosomal protein HMGN1 modulates histone H3 phosphorylation." Molecular Cell **15**(4):573-84.

Liu J., N.B. Perumal, C.J. Oldfield, E.W. Su, V.N. Uversky, A.K. Dunker (2006). "Intrinsic disorder in transcription factors." Biochemistry **45**(22):6873-6888.

Mansouri A., L.D. Ridgway, A.L. Korapati, Q. Zhang, L. Tian, Y. Wang, Z.H. Siddik, G.B. Mills, F.X. Claret (2003). "Sustained activation of JNK/p38 MAPK pathways in response to cisplatin leads to Fas ligand induction and cell death in ovarian carcinoma cells." Journal of Biological Chemistry **278**(21):19245-56.

Masel, J. and M.L. Siegal (2009) "Robustness: mechanisms and consequences." Trends in Genetics. 25(9): 395-403.

McA'Nulty M.M., J.P. Whitehead, S.J. Lippard (1996). "Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts in vitro and *in vivo*." Biochemistry **35**(19):6089-99.

McA'Nulty M.M. and S.J. Lippard (1996). "The HMG-domain protein Ixr1 blocks excision repair of cisplatin-DNA adducts in yeast." Mutat Research **362**(1):75-86.

McKay B.C., C. Becerril, M. Ljungman (2001). "p53 plays a protective role against UV- and cisplatin-induced apoptosis in transcription-coupled repair proficient fibroblasts." Oncogene **20**(46):6805-8.

Mello J.A., S. Acharya, R. Fishel, J.M. Essigmann (1996). "The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin." Chemistry & Biology **3**(7):579-89.

Misteli T., A. Gunjan, R. Hock, M. Bustin, D.T. Brown (2000). "Dynamic binding of histone H1 to chromatin in living cells." Nature **408**(6814):877-81.

Morimoto R.I. (2008). "Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging." Genes Development **22**(11):1427-38.

Müller S., L. Ronfani, M.E. Bianchi (2004). "Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function." Journal of International Medicine **255**(3):332-43.

Namy O., A. Galopier, C. Martini, S. Matsufuji, C. Fabret, J.P. Rousset (2008). "Epigenetic control of polyamines by the prion [PSI+]." Nature Cell Biology **10**(9):1069-75.

Ohndorf U.M., J.P. Whitehead, N.L. Raju, S.J. Lippard (1997). "Binding of tsHMG, a mouse testis-specific HMG-domain protein, to cisplatin-DNA adducts." Biochemistry **36**(48):14807-15.

Pil P.M., S.J. Lippard (1992). "Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin." Science **256**(5054):234-7.

Pinato O., C. Musetti, N.P. Farrell, C. Sissi (2013). "Platinum-based drugs and proteins: reactivity and relevance to DNA adduct formation." Journal of Inorganic Biochemistry **122**:27-37.

Pinato O., C. Musetti, C. Sissi (2014). "Pt-based drugs: the spotlight will be on proteins." Metallomics **6**(3):380-95.

Piskur J. (2001). "Origin of the duplicated regions in the yeast genomes." Trends in Genetics **17**(6):302-3.

Privalov P.L., I. Jelesarov, C.M. Read, A.I. Dragan, C. Crane-Robinson (2003). "The energetics of HMG box interactions with DNA: thermodynamics of the DNA binding of the HMG box from mouse sox-5." Journal of Molecular Biology **294**(4):997-1013.

Rattner B.P., T. Yusufzai, J.T. Kadonaga (2009). "HMGN proteins act in opposition to ATP-dependent chromatin remodeling factors to restrict nucleosome mobility." Molecular Cell **34**(5):620-6.

Riedl T., F. Hanaoka, J.M. Egly (2003). "The comings and goings of nucleotide excision repair factors on damaged DNA." EMBO Journal **22**(19):5293-303.

Rodriguez Lombardero S., A. Vizoso Vazquez, E. Rodriguez Belmonte, M.I. Gonza lez Siso and M.E. Cerdan (2012). "SKY1 and IXR1 interactions, their effects on cisplatin and spermine resistance in Saccharomyces cerevisiae." Canadian Journal of Microbiology **58**(2): 184-188.

Rosenberg B., L. Vancamp, T. Krigas (1965). "Inhibition of cell division in escherichia coli by electrolysis products from a platinum electrode." Nature **205**:698-699.

Rosenfeld E., B. Beauvoit (2003). "Role of the non-respiratory pathways in the utilization of molecular oxygen by *Saccharomyces cerevisiae*." *Yeast*. **20**(13):1115-44.

Salvesen G.S., V.M. Dixit (1997). "Caspases: intracellular signaling by proteolysis." Cell **91**(4):443-6. Review.

Sancho-Martínez S.M., L. Prieto-García, M. Prieto, J.M. López-Novoa, F.J. López-Hernández (2012). "Subcellular targets of cisplatin cytotoxicity: an integrated view." Pharmacology & Therapeutics **136**(1):35-55.

Scovell W.M., N. Muirhead, L.R. Kroos (1987). "*cis*-Diamminedichloroplatinum(II) selectively cross-links high mobility group proteins 1 and 2 to DNA in micrococcal nuclease accessible regions of chromatin." Biochemical and Biophysical Research Communications **142**(3):826-35.

Sedletska Y., L. Fourrier, J.M. Malinge (2007). "Modulation of MutS ATP-dependent functional activities by DNA containing a cisplatin compound lesion (base damage and mismatch)." Journal of Molecular Biology **369**(1):27-40.

Seger, J. and H.J. Brockmann (1987). "What is bet-hedging?" Oxford Surveys in Evolutionary Biology **4**:182—211.

Serio T.R., A.G. Cashikar, A.S. Kowal, G.J. Sawicki, J.J. Moslehi, L. Serpell M.F Arnsdorf and S.L. Lindquist (2000). "Nucleated conformational conversion and the replication of conformational information by a prion determinant." Science **289**(5483): 1317-1321.

Sezen B., D. Sames (2004). "Oxidative C-arylation of free (NH)-heterocycles via direct sp3 C-H bond functionalization." Journal of the American Chemical Society. 2004, **126**(41):13244-13246.

Shorter J. and S. Lindquist (2005). "Prions as adaptive conduits of memory and inheritance." Nature Review Genetics **6**(6): 435-450.

Smith J., L.M. Tho, N. Xu, D.A. Gillespie (2010). "The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer." Advances in Cancer Research **108**:73-112.

Sunnerhagen P., J. Piskur (2006). "Comparative Genomics. Using Fungi as Models." Springer-Verlag Berlin Heidelberg. Germany.

Sweeny E.A., J. Shorter (2008). "Prion proteostasis: Hsp104 meets its supporting cast." Prion **2**(4):135-40.

Ter Linde J.J., H.Y. Steensma (2002). "A microarray-assisted screen for potential Hap1 and Rox1 target genes in *Saccharomyces cerevisiae*." Yeast **19**(10):825-40.

Tessier P.M. and S. Lindquist (2009). "Unraveling infectious structures, strain variants and species barriers for the yeast prion [PSI+]." Nature Structural Molecullar Biology **16**(6): 598-605.

Treiber D.K., X. Zhai, H.M. Jantzen, J.M. Essigmann (1994). "Cisplatin-DNA adducts are molecular decoys for the ribosomal RNA transcription factor hUBF (human upstream binding factor)." Proceedings of the National Academy of Sciences of the

United States of America **91**(12):5672-6.

Trimmer E.E., D.B. Zamble, S.J. Lippard, J.M. Essigmann (1998). "Human testis-determining factor SRY binds to the major DNA adduct of cisplatin and a putative target sequence with comparable affinities." Biochemistry **37**(1):352-62.

True, H.L. and S.L. Lindquist (2000). "A yeast prion provides a mechanism for genetic variation and phenotypic diversity." Nature **407**(6803): 477-483.

Turchi J.J., K. Henkels (1996). "Human Ku autoantigen binds cisplatin-damaged DNA but fails to stimulate human DNA-activated protein kinase." Journal of Biological Chemistry **271**(23):13861-7.

Venkitaraman A.R. (2002). "Cancer susceptibility and the functions of BRCA1 and BRCA2." Cell **108**(2):171-82. Review.

Vichi P., F. Coin, J.P. Renaud, W. Vermeulen, J.H. Hoeijmakers, D. Moras, J.M. Egly (1997). "Cisplatin- and UV-damaged DNA lure the basal transcription factor TFIID/TBP." EMBO Journal **16**(24):7444-56.

Watson M., K. Stott, H. Fischl, L. Cato, J.O. Thomas (2014). "Characterization of the interaction between HMGB1 and H3-a possible means of positioning HMGB1 in chromatin." Nucleic Acids Research **42**(2):848-59.

Wei M., O. Burenkova, S.J. Lippard (2003). "Cisplatin sensitivity in Hmbg1-/- and Hmbg1+/+ mouse cells." Journal of Biological Chemistry **278**(3):1769-73.

Weiss R.B., M.C. Christian (1993). "New cisplatin analogues in development. A review." Drugs **46**(3):360-377.

Wetzel C.C., S.J. Berberich (2001). "p53 binds to cisplatin-damaged DNA." Biochimica et Biophysica Acta **1517**(3):392-7.

Wickner R.B., H.K. Edskes, F. Shewmaker, T. Nakayashiki, A. Engel, L. McCann and D. Kryndushkin (2007). "Yeast prions: evolution of the prion concept." Prion **1**(2): 94-100.

Winter C., P. Albers (2011). "Testicular germ cell tumors: pathogenesis, diagnosis and treatment." Nature Reviews Endocrinology **7**: 43-53

Wong B., J.E. Masse, Y.M. Yen, P. Giannikopoulos, J. Feigon, R.C. Johnson (2002). "Binding to cisplatin-modified DNA by the *Saccharomyces cerevisiae* HMGB protein Nhp6A." Biochemistry **41**(17):5404-14.

Wright P.E., H.J. Dyson (2009). "Linking folding and binding." Current Opinion in Structural Biology **19**(1):31-8.

Xin H., S. Taudte, N.R. Kallenbach, M.P. Limbach, R.S. Zitomer (2000). "DNA binding by single HMG box model proteins." Nucleic Acids Research **28**(20):4044-50.

Yamada M., E. O'Regan, R. Brown, P. Karran (1997). "Selective recognition of a cisplatin-DNA adduct by human mismatch repair proteins." Nucleic Acids Research **25**(3):491-6.

Zagorec M., R. Labbe-Bois (1986). "Negative control of yeast coproporphyrinogen oxidase synthesis by heme and oxygen." Journal of Biological Chemistry **261**(6):2506-9.

Zamble D.B., D. Mu, J.T. Reardon, A. Sancar, S.J. Lippard (1996). "Repair of cisplatin-DNA adducts by the mammalian excision nuclease." Biochemistry

35(31):10004-13.

Zenthon J.F., F. Ness, B. Cox, M.F. Tuite (2006). "The [PSI+] prion of *Saccharomyces cerevisiae* can be propagated by an Hsp104 orthologue from *Candida albicans*." Eukaryotic Cell **5**(2):217-25.

Zitomer R.S., C.V. Lowry (1992). "Regulation of gene expression by oxygen in *Saccharomyces cerevisiae*." Microbiological Reviews **56**(1):1-11.

Objectives

The main objective of this Thesis work is the characterization of the structure and functions of the HMGB protein Ixr1 from *Saccharomyces cerevisiae.* To reach this main objective the following particular objectives have been undertaken:

1.- To analyse the role of Ixr1 in the transcriptional response of *S. cerevisiae* to hypoxia by a transcriptome approach.

2.- To characterize DNA regulatory sequences important for *in vivo* Ixr1 binding and to find whether there is a relationship between the binding of DNA regulatory sequences of the two HMGB proteins of *S. cerevisiae* that are related to the hypoxic response: Rox1 and Ixr1. This study will be based on a ChIP-on-chip approach.

3.- To analyse the role of Ixr1 in the transcriptional response of *S. cerevisiae* to the anti-cancer drug cisplatin by a transcriptome approach.

4.- To differentiate the *in vivo* binding of Ixr1 to DNA in cells non treated and treated with cisplatin and correlate these results with the transcriptome response of *S. cerevisiae* cells to this drug by a ChIP-on-chip approach.

5.- To quantify *in vitro* by diverse biochemical and thermodynamic methods the particular interactions of each HMG-box domains (A and B) present in Ixr1 with several DNA regulatory sequences previously characterized in the *ROX1* and *HEM13* promoters and selected as a model of study; as well as the "four-way junctions" form and modifications produced by DNA-platination in the studied interactions.

6.- To characterize from a structural point of view the Ixr1 protein and their principal domains.

Outline of this thesis

The high-mobility group (HMG) proteins were discovered as nuclear factors (NFs) more than 40 years ago and they were named for their high electrophoretic mobility in polyacrylamide gels (Goodwin *et al.*, 1973). They are present in almost all metazoans and plants. Although HMG motifs are present in many nuclear proteins, the classification and nomenclature of the considered "canonical" HMG proteins is organized in 3 families named HMGA, HMGB, and HMGN, each one having a specific functional motif; the "AT-hook" in HMGA, the "HMG-box" in HMGB and the "nucleosomal binding domain" in HMGN (Bustin, 2001). In the nucleus, they act as non-histone architectural chromatin-proteins, but they also have other regulatory functions upon replication, transcription and DNA repair. Some HMG proteins have been related to extra-nuclear and extracellular functions during inflammation, cell differentiation, cell migration, and tumor metastasis (Hock, 2006).

The HMG-box that characterizes the HMGB family is composed of three α-helices folded into an L-shaped configuration whose concave surface binds into the minor groove of DNA (Thomas and Travers 2001). The binding causes widening of the groove and bending of the DNA axis (Figure 1). HMGB proteins also bind with high affinity to already distorted DNA structures such as four-way junctions, bulges, kinks and modified DNA containing cisplatin adducts (Pil & Lippard, 1992).

Saccharomyces cerevisiae has seven genes expressing HMGB proteins: *ABF2*, *HMO1*, *NHP6A*, *NHP6B*, *NHP10*, *IXR1* and *ROX1* (Bustin, 2001). Only one HMG-box domain is present in five of them, but Abf2 and Ixr1 have two in tandem "HMG-box" motifs. The main objective of this PhD Thesis is to achieve a substantial progress in the molecular characterization of the structure and functions of Ixr1.

Functionally, Ixr1 (formerly named Ord1, for <u>o</u>xygen/<u>o</u>xidase <u>r</u>egulation <u>d</u>efective) has been previously related to the transcriptional control of the

response to changes normoxia-hypoxia of several yeast genes such as *COX5B*, which encodes the hypoxic isoform of subunit Vb of cytochrome *c* oxidase (Lambert *et al*., 1994); *TIR1*, which encodes a cell wall mannoprotein; induced by cold shock and anaerobiosis (Bourdineaud *et al*., 2000); *HEM13*, which encodes the enzyme coproporphyrinogen III oxidase that catalyzes the rate-limiting step in heme biosynthesis (Castro-Prego *et al*., 2010a); and *ROX1*, which encodes a transcriptional repressor of hypoxic genes in normoxia (Castro-Prego *et al*., 2010b). In order to better understand the role of Ixr1 in the hypoxic response and to increase the repertory of Ixr1 target genes, we first characterized the transcriptome of two yeast isogenic strains, Ixr1-wild type and *Δixr1*, grown in aerobic and hypoxic conditions. Two different experimental approaches were designed to analyze the response during long-term hypoxic growth and the response after a shift from aerobic to hypoxic conditions. Chapter 1 describes the main results obtained using DNA arrays and confirmed by qPCR.

Two HMGB proteins, Rox1 and Ixr1, participate in the transcriptional regulation caused by changes in oxygen availability in *S. cerevisiae* (Bourdineaud *et al*., 2000) and *in vitro* and *in vivo* experiments indicate that both regulators might co-regulate the gene *HEM13* by direct binding to target sequences in its promoter region (Castro-Prego *et al*., 2010a). Although genome-wide analysis of changes in mRNA levels have been carried out in *S. cerevisaie* strains with the Rox1-wild type allele and *Δrox1* in normoxia and hypoxia (Becerra *et al*., 2002), the binding of Ixr1 and Rox1 to promoter regions of target genes had not been studied at a genomic scale. Considering that mRNA levels of Rox1 (Deckert *et al*., 1995; Zitomer *et al*., 1997) and Ixr1 (Castro-Prego *et al*., 2010b) are both subject to changes fired by adaptation to normoxia-hypoxia, we have carried out ChIP-on-chip analyses of Ixr1 and Rox1 binding to the *S. cerevisiae* genome. The results obtained are described and discussed in Chapter 2.

Although Ixr1 was first identified by its ability to bind DNA previously modified by the anticancer drug cisplatin and deletion of *IXR1* increases the resistance of yeasts cells to this drug (Brown *et al*., 1993), very little is known about their molecular function in relation to DNA damage and DNA repair. Yeast strains lacking Ixr1 have decreased amounts of dNTPs (Tsaponina *et al*., 2011). It has been demonstrated that Ixr1 is required for the expression of Rnr1, one of the four subunits of the enzyme ribonucleotide reductase, which catalyzes the rate-limiting step in the production of all four dNTPs (Tsaponina *et al*., 2011). These are the building blocks necessary for DNA synthesis both during cell cycle progression and after DNA damage. The relationship between this molecular function of Ixr1 and the increased resistance caused by *IXR1* deletion was not evident. More recently, it has been observed that inactivation of *IXR1* renders cells resistant not only to cisplatin but also to other DNA-damaging drugs with different mechanisms of action (Tsaponina and Chaves, 2013). A hypothesis has been formulated in the sense that dNTPs depletion observed in the *Δixr1* mutant strain is the cause of constitutive genome integrity checkpoint activation and consequently of the broad DNA damage tolerance observed in the mutant. In support of this hypothesis it has been demonstrated that the wild-type yeast cells exhibit increased DNA damage tolerance when the genome integrity checkpoint is pre-activated by low concentrations of hydroxyurea, a drug that slows down DNA replication by depleting dNTPs (Tsaponina and Chaves, 2013). In the course of this PhD Thesis the effects of cisplatin treatment on the transcriptome of a wild-type and an isogenic *Δixr1* strain have been analyzed and discussed in relation to Ixr1 binding to promoter and non-promoter regions of the *S. cerevisiae* genome. The results of these transcriptome and ChIP-on-chip analyses are explained in Chapter 3.

The determination of Ixr1 structure is attractive and challenging since the in tandem organization of its two HMG-boxes is not similar to those present in other HMGB proteins, whose structures have been experimentally determined. Besides, it contains three Q-rich regions that are not present in the other HMGB

proteins from *S. cerevisiae*, with the exception of a short Q-stretch present in Rox1. Modelling of the two HMG-boxes present in Ixr1 and comparison with experimentally determined structures of the unique HMG-box of mammalian SRY (Werner, 1995) and the two boxes present in human HMGB1 (Read, 1993; Weir, 1993; Hardman, 1995; Stott, 2006) predicts that binding to DNA might be non-sequence-specific (or structure-specific) through the first HMG-box and sequence-specific through the second (Castro-Prego *et al.*, 2010a). The binding of Ixr1 to a sequence present in the *ROX1* promoter (−428 to −254) and previously shown as necessary for Rox1 binding for aerobic auto-repression of *ROX1* (Decker *et al.*, 1995) was demonstrated *in vivo* and *in vitro* (Castro-Prego *et al.*, 2010b). It was observed that binding of Ixr1 to the *ROX1* promoter increased during hypoxia in contrast with Rox1, which binds the *ROX1* promoter during aerobiosis (Castro-Prego *et al.*, 2010b). Previously reported cross-regulation between *ROX1* and *IXR1* (Castro- Prego *et al.*, 2010b) is functionally relevant since both are related to the *S. cerevisiae* hypoxic response. To get a better understanding of regulation of *ROX1* by Ixr1, in terms of structural commitments, the binding of the HMG-boxes of Ixr1 to the consensus present in the *ROX1* promoter has been analyzed in this work. Structural characteristics of both HMG-boxes have been determined by CD and NMR. The binding of each HMG-box to the consensus in the *ROX1* promoter have been analysed by EMSA experiments; the thermo-physical constants of the binding have been calculated by diverse experiments including FP and calorimetric studies. The results obtained and their structural and functional implications are explained in Chapter 4.

Several attempts were done for crystallization of the Ixr1 protein purified from yeast and bacteria. The production in yeast had been considered as advantageous since post-translational modifications could be integrated in the model. However, this strategy rendered a low yield, not enough for crystallization due to toxic effects of over-production. In bacteria the yield was higher, but the protein did not produce good diffracting crystals due to its peculiar sequence and

amino acidic composition. We therefore focused our study in other affordable structural aspects like the oligomeric state, the identification of ordered and disordered stretches, the significance of the poly-glutamine regions and their tendency to form amyloids. All these aspects are exposed in Chapter 5.

REFERENCES

Becerra M, L.J. Lombardía-Ferreira, N.C. Hauser, J.D. Hoheisel, B. Tizón, M.E. Cerdán (2002) "The yeast transcriptome in aerobic and hypoxic conditions: effects of hap1, rox1, rox3 and srb10 deletions" Molecular Microbiology **43**(3):545-55.

Bourdineaud, J. P., G. De Sampaio & G.J. Lauquin (2000) "A Rox1-independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the SRP1/TIR1 gene" Molecular Microbiology **38**(4):879-90.

Brown SJ, P.J. Kellett, S.J. Lippard (1993) "Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin" Science **261**(5121):603-5.

Bustin M. (2001) "Revised nomenclature for high mobility group (HMG) chromosomalproteins" Trends in Biochemical Sciences **26** (3) 152–153.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, I. Carneiro, I. González-Siso and M.E. Cerdán (2010a) "Regulatory factors controlling transcription of Saccharomyces cerevisiae IXR1 by oxygen levels: a model of transcriptional adaptation from aerobiosis to hypoxia implicating ROX1 and IXR1 cross-regulation." Biochemical Journal **425**(1):235-43.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso and M.E. Cerdán (2010b) "Ixr1p regulates oxygen-dependent HEM13 transcription." FEMS Yeast Research **10**(3):309-21.

Deckert J, R. Perini, B. Balasubramanian, R.S. Zitomer (1995) "Multiple elements and auto-repression regulate Rox1, a repressor of hypoxic genes in *Saccharomyces cerevisiae*" Genetics **139**(3):1149-58.

Goodwin GH, C. Sanders & E.W. Johns (1973) "A new group of chromatin-associated proteins with a high content of acidic and basic amino acids" European Journal of Biochemistry **38**(1):14-9.

Hock R, T. Furusawa, T. Ueda, M. Bustin (2007) "HMG chromosomal proteins in development and disease" Trends in Cell Biology **17**(2):72-9.

Lambert JR, V.W. Bilanchone & M.G. Cumsky (1994) "The ORD1 gene encodes a transcription factor involved in oxygen regulation and is identical to IXR1, a gene that confers cisplatin sensitivity to Saccharomyces cerevisiae" Proceedings of the National Academy of Sciences U S A **91**(15):7345-9

Pil P. M. and S.J. Lippard (1992). "Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin." Science **256**(5054):234-7.

Thomas J.O. and A.A. Travers (2001). "HMG1 and 2, and related 'architectural' DNA-binding proteins." Trends in Biochemical Science **26**(3):167–174.

Tsaponina O., E. Barsoum, S.U. Aström and A. Chabes (2011). "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools." PLoS Genetics **7**(5):e1002061.

Tsaponina O. and A. Chabes (2013). "Pre-activation of the genome integrity checkpoint increases DNA damage tolerance." Nucleic Acids Research **41**(22):10371-8.

Zitomer RS, P. Carrico, J. Deckert (1997) "Regulation of hypoxic gene expression in yeast" Kidney International **51**(2):507-13.

Chapter 1

Ixr1p and the control of the *Saccharomyces cerevisiae* hypoxic response

Ángel Vizoso-Vázquez, Mónica Lamas-Maceiras, Manuel Becerra, M. Isabel González-Siso, Esther Rodríguez-Belmonte and M. Esperanza Cerdán.

Dpto. Biología Celular y Molecular. Universidad de A Coruña

F. Ciencias, Campus de A Zapateira s/n 15071. A Coruña. SPAIN.

ABSTRACT

In *Saccharomyces cerevisiae*, adaptation to hypoxia/anaerobiosis requires the transcriptional induction or de-repression of multiple genes organized in regulons controlled by specific transcriptional regulators. Ixr1 is a transcriptional regulatory factor that causes aerobic repression of several hypoxic genes (*COX5B, TIR1* and *HEM13*) and also activation of *HEM13* during hypoxic growth. Analysis of the transcriptome of the wild-type strain BY4741 and its isogenic derivative *Δixr1,* grown in aerobic and hypoxic conditions, reveals differential regulation of genes related to the hypoxic and oxidative stress responses, but also to the re-adaptation of catabolic and anabolic fluxes in response to oxygen limitation. The function of Ixr1 in transcriptional regulation of genes from the sulphate assimilation pathway and other pathways producing α-keto-acids is of biotechnological importance for the industries based on yeast-derived fermentation products.

1.- INTRODUCTION

The yeast *Saccharomyces cerevisiae* uses both respiration and fermentation to obtain energy through metabolic pathways. *S. cerevisiae* grows under aerobic (normoxic) and also under limited (hypoxic) or even depleted (anaerobic/anoxic) oxygen conditions, the latter excluding oxygen-dependent respiratory metabolism. During hypoxia, it is advantageous for the cell to adapt the pattern of gene expression in order to improve oxygen utilization. Among the genes that are induced during hypoxia are those whose products fit principally into four categories: cell wall composition, lipid and carbohydrate metabolism and cellular stress response (ter Linde *et al*. 2002; Becerra *et al*. 2002; Kwast *et al*. 2002).

In *S. cerevisiae*, adaptation to hypoxia/anaerobiosis requires the transcriptional induction or de-repression of multiple genes organized in regulons controlled by specific transcriptional regulators. The regulon controlled by the aerobic repressor Rox1 was the first and best characterized (Zitomer & Lowry 1992; Zitomer *et al*. 1997; Kastaniotis & Zitomer 2000; Klinkenberg *et al*. 2005) but it is not unique. Upc2, Sut1 and Sut2 are also transcriptional regulators related to the induction of hypoxic/anaerobic genes (Abramova *et al*. 2001a; Abramova *et al*. 2001b; Vik & Rine 2001; Regnacq *et al*. 2001). Ixr1 is a transcriptional regulatory factor that causes aerobic repression of the *COX5B* gene, which encodes the hypoxic isoform of the subunit Vb of the mitochondrial complex cytochrome *c* oxidase (Lambert *et al*. 1994). Ixr1 has also been related to aerobic and anaerobic regulation of *TIR1,* a cell wall mannoprotein of the Srp1/Tip1 family of serine-alanine-rich proteins (Bourdineaud *et al*. 2000) and *HEM13,* which encodes the enzyme coproporphyrinogen III oxidase necessary in the heme biosynthetic pathway (Castro *et al*. 2010a).

Rox1 and Ixr1 both contain HMG (high-mobility group) domains, which bind to and bend DNA (Deckert *et al*. 1995; Deckert *et al*. 1999; McA'Nulty *et al*. 1996). Although initially the regulation caused by these two proteins was considered mainly independent, we have recently found that *HEM13* transcription is controlled by both factors (Castro *et al*. 2010a). Rox1 and Mot3 repress aerobic expression of *HEM13* (Klinkenberg *et al*. 2005), while Ixr1 is a positive regulator of *HEM13* transcription in hypoxia (Castro *et al*. 2010a). Besides, a transcriptional cross-regulation between the genes *ROX1* and *IXR1* exists (Castro *et al*. 2010b). In aerobiosis, low levels of *IXR1* expression are maintained by Rox1 repression through the general co-repressor complex Tup1-Ssn6. Ixr1 is also required for hypoxic repression of *ROX1* and binds to its promoter (Castro *et al*. 2010b).

The transcriptomes of *S. cerevisiae* during aerobic, hypoxic and anaerobic growth have been previously analyzed (ter Linde *et al*. 2002; Becerra *et al*. 2002;

Kwast *et al.* 2002) and the effects of *ROX1* deletion in these conditions were also studied by genomic approaches (Becerra *et al.* 2002; Kwast *et al.* 2002). Here we present a genome-wide transcriptional analysis in the *S. cerevisiae* wild-type strain BY4741 and in its isogenic derivative *Δixr1* strain grown in aerobic and hypoxic conditions. Two different approaches to the hypoxic-elicited response have been driven; the response during hypoxic growth and the response after a shift from aerobic to hypoxic conditions.

2.- MATERIALS AND METHODS

2.- Strains and growth conditions

The *S. cerevisiae* strains, BY4741 *(MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0)* and BY4741-*Δixr1* (*MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 YKL032c::kanMX4*) were obtained from EUROSCARF (European *S. cerevisiae* archive for functional analysis; http://web.uni-frankfurt.de/fb15/mikro/euroscarf/). Growth and handling of yeasts were carried out according to standard procedures. For RNA extractions, the yeast cells were cultured in Erlenmeyer flasks at 250 rpm. The flasks were filled with 40% volume of culture medium. Yeast cells were grown at 30 ^{o}C in synthetic complete medium (CM) containing per liter: 6.7 g of bacto-yeast nitrogen base without amino acids from Difco (Franklin Lakes, New Jersey, USA); 40 mg each of histidine, leucine, adenine, uracil, lysine and tyrosine, 10 mg each of arginine, methionine and threonine, 30 mg tryptophan; 60 mg each of phenylalanine and isoleucine; 2% glucose (w/v). During hypoxic growth, cells were cultured over night in anaerobic jars with the GasPack EZAnaerobe system from Becton, Dickinson and Company (Franklin Lakes, New Jersey, USA) and under these conditions (oxygen concentration <1%) the medium was supplemented with 20 mg/L ergosterol and 0.5% tween 80. In experiments analyzing the consequences of a shift from aerobic to hypoxic conditions, cells were grown first in aerobiosis until OD_{600} = 0.8 and then

hypoxia was generated by bubbling nitrogen in the flasks at constant pressure during 3 hours.

Differences of growth between the BY4741 and BY4741-*Δixr1* strains were monitored in YP (1% w/v yeast extract 2% w/v peptone) with 2% w/v of glucose, galactose or glycerol as specified. The cells were pre-adapted over night to each carbon source by growing the pre-inocula in the same liquid media. Yeast cells were cultured in Erlenmeyer flasks filled to 20% volume and with continuous agitation at 250 rpm. The starting OD_{600} was 0.05 and OD was measured at different time points during 40 hours.

2.2.- RNA isolation

For RNA isolations, cells were harvested, immediately frozen in liquid nitrogen, disrupted using a Micro-Dismembrator (B. Braun Biotech International, Allentown, Pennsylvania, USA). The resulting powder was mixed with TRIZOL Reagent (GIBCO BRL- Life Technologies, Gaithersburg, Maryland, USA) and total RNA was extracted by the method previously described (Chomczynski and Sacchi 1987).

2.3.- cDNA Probe generation

The cDNA probes were prepared as described (Hauser *et al*. 1998). Briefly, 60 μg of total RNA was annealed to oligonucleotide dT18 and used as a template to synthesize and radiolabel the corresponding first strand cDNA with 50 μCi of [α-^{32}P]-dATP (Amersham Biosciences, Little Chalfont, UK) and SuperScript II (GIBCO BRL Life Technologies, Gaithersburg, Maryland, USA). The reactions were carried out at 43 °C for 1 h, after which the RNA was hydrolyzed with NaOH at 65 °C for 30 minutes. The probe was purified by isopropanol precipitation and the incorporation of ^{32}P was measured to check the efficiency of the reaction.

2.4.- Filter hybridization, washing and stripping

The use of arrays was as previously described (Becerra *et al*. 2002). Arrays (yeast DNA chips version 5) were supplied by the Valencia University (DNA-Chips Services from SGAI UVA, Valencia. Spain). These arrays contain all of the yeast ORFs and have been previously described (Alverola *et al*. 2004). Filters were pre-hybridized for 1 hour at 65 ^{o}C in the hybridization mix (750 mM NaCl, 75 mM sodium acetate, 0.1% w/v bovine serum albumine, 0.1% w/v ficoll, 0.1% w/v polyvinylpyrrolidone and 0.5% w/v sodium dodecylsulphate, pH 7). The probe was then denatured for 5 minutes at 100 ^{o}C, cooled quickly on ice, and hybridized with the arrays over night at 65 ^{o}C. The following day, two washes were performed at 65 ^{o}C, for 5 and 20 minutes respectively, in 300 mM NaCl, 30 mM sodium acetate, 0.1% (w/v) sodium dodecylsulphate, pH 7. Filter regeneration was done by pouring a boiling solution of 5 mM sodium phosphate (pH 7.5) and 0.1% (w/v) sodium dodecylsulphate over the filters prior to their reuse. We re-used each DNA array three times, being the number of re-utilizations described without significant loss of signal over nine, as previously described for these arrays (Alverola *et al*. 2004).

2.5.- Experimental design

Biological replicates were run in duplicate. Two independent RNA extractions were obtained from biological replicates. The same array was hybridized in successive uses with cDNA obtained from cells grown in normoxia, hypoxia generated in anaerobic jars or shifted to hypoxia by bubbling nitrogen in the flasks. Four independent DNA arrays were used.

2.6.- Signal quantification and data analysis

The filters were exposed for 24 hours to a phosphor screen and the data were collected using a Phosphor Imager Scanning Instrument 425 (Molecular

Dynamics-Amersham Biosciences, Little Chalfont, UK). Signal quantification was performed with Array Vision software (Molecular Dynamics-Amersham Biosciences, Little Chalfont, UK).

Normalization of data from each membrane was carried out by dividing the measurement for each gene by the median of all spots in the filter. Median and standard deviations were calculated for the expression of each single gene in the total number of replica. To find those genes which were expressed at higher levels in the wild type than in the mutant strain the ratio wt/*Δixr1* was calculated (activation ratio). To find those genes which were expressed at higher levels in the mutant than in the wt strain the ratio *Δixr1*/wt was calculated (repression ratio). A *t-test* was applied to evaluate the differences between means. Genes were considered to be differentially expressed in the two strains (wild type and *Δixr1* mutant) or in response to different culture conditions (normoxia, hypoxia or shift from normoxia to hypoxia) if they satisfied both a *p-value* lower than 0.05 in the t-test statistical evaluation of differences between means and at least a two-fold change in expression levels. The original data from this study are available from GEO (http://www.ncbi.nlm.nih.gov/geo/info/linking.html); the accession number is GSE30046. The data were processed, clustered and analyzed using the bio-informatics tools of Genespring (Agilent Technologies, Palo Alto, California, USA). Processed data from differentially expressed genes (DEGs) are available through supplementary material online resources accompanying this paper.

Functional distribution of genes in the differentially regulated clusters was analyzed using FunSpec (http://funspec.ccbr. utoronto.ca/) developed by Robinsons and coworkers (Robinson *et al*. 2002). The MIPS Functional Catalogue Database (FunCatDB) was used in the analyses (http://mips.helmholtz-muenchen.de/proj/ funcatDB/). For these analyses a *p-value* lower than 0.01 was selected. These *p-value*s represent the probability that the intersection of one given list of genes with any given functional category occurs by chance. In the

68

report of the analyses carried out with FunSpec (Tables 2 to 7)*,* k is the number of genes from the input cluster in given category and f is the total number of genes in given category.

2.7.- Analysis of expression by qRT-PCR

Total RNA isolated with the NucleoSpin RNAII kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany) was converted into cDNA and labeled with the KAPA SYBR FAST universal one-step qRT-PCR kit (Kapa Biosystems, Inc, Woburn, Massachusetts, USA). PCR primers for individual genes selected after the microarray analysis were designed with the Universal ProbeLibrary Assay Design Center developed by Roche Diagnostics Corp., Indianapolis, Indiana, USA (https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id= uplct_030000) to generate 60-85 base pairs amplicons with a Tm of 59 or 60°C. The ECO Real-Time PCR System was used for the experiments and calculations (Illumina, Inc., San Diego, California, USA). Three independent RNA extractions were assayed for each strain or condition. The mRNA levels of the selected genes were corrected by the mRNA levels of *TAF10*, a gene previously verified to be constitutive in the assayed conditions and not affected by the *Δixr1* deletion. A t-test was applied to evaluate the differences between means as previously described.

3.- RESULTS

3.1.- Effects of *IXR1* deletion on the *S. cerevisiae* transcriptome during normoxic growth

The expression of the whole set of genes in wild type and *Δixr1 S. cerevisiae* haploid strains were first compared under normoxic growth. Analysis of the DNA-array data by t-test statistical analysis as described in Material and

methods revealed a total of 322 differentially expressed genes (DEGs). They were clustered with Genespring using the gene tree clustering method, measuring similarity by standard correlation, with a separation ratio of 0.5 and a minimum distance of 0.001 (Figure 1a). The whole data sheet of these DEGs and relevant functional annotations as reported by Genespring is available through Table S1.

A total of 248 genes are expressed at higher levels in the mutant than in the wild type strain. Functional distribution (according to MIPS Functional Classification) of the genes from these clusters was analyzed with FunSpec and functional groups that are over represented in these clusters, according to the cut *p-value* selected, are summarized in Table 1 with the statistical parameters for each group. Representative groups include genes of the sulphate assimilation pathway, stress response, oxidative stress response, vacuolar protein degradation or pyrimidine metabolism.

Only 74 genes are expressed at higher levels in the wild type than in the mutant. Over represented functional groups (Table 2) include genes of the biosynthesis of the branched chain (BC) amino acids valine, isoleucine and leucine (VIL), sugar and purine metabolism and ion transporters. Regarding Ixr1-mediated regulation of genes related to carbohydrate metabolism we have found that several genes, which are necessary for producing enzymes from the glycolytic pathway, are positively regulated by Ixr1 during normoxic growth (Table 2). We have verified (figure 2) that differences of growth-rate in glucose or galactose are not detected between BY4741 wild type and *Δixr1* strains. However, growth rate in glycerol is minor in the *Δixr1* strain than in the wild type.
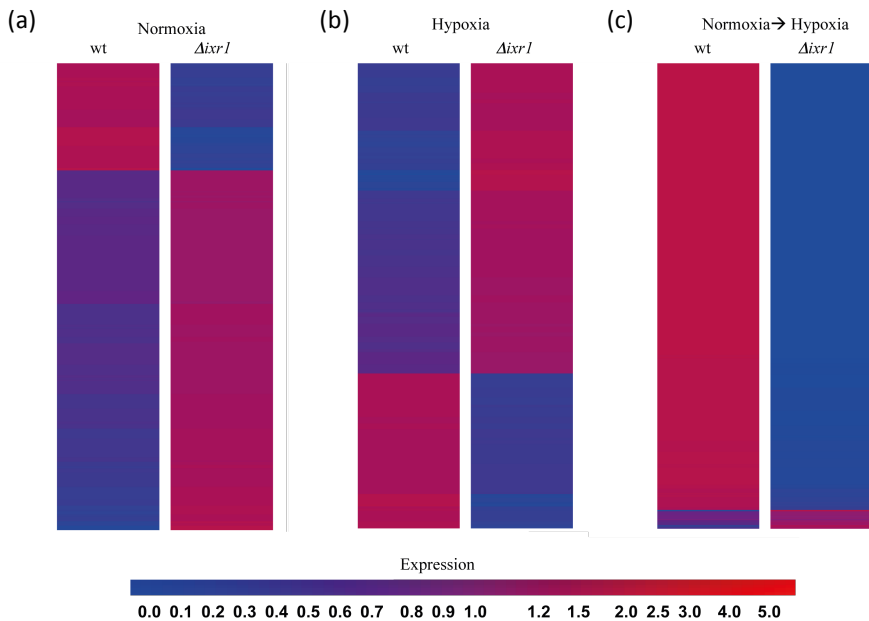
Figure 1. Cluster analysis of DEGs included in Table S1 (a), Table S2 (b) and Table S3 (c) with Genespring. The values analyzed are the means of significant changes in expression obtained from independent cultures as described in the experimental design in Materials and methods. The genes that are above the median of the normalized values between the different samples are shown in red and those below are shown in blue. The color intensity is proportional to the normalized values, as shown on the scale at the bottom of the figure. Cells were grown under normoxia (a) hypoxia (b) or after a hypoxic sift (c).

3.2.- Effects of *IXR1* deletion on the *S. cerevisiae* transcriptome during hypoxic growth

For hypoxic growth, cells well inoculated in media supplemented with 20 mg/L ergosterol and 0.5% Tween 80 to OD$_{600}$ = 0.05 and cultured overnight in anaerobic jars. In these conditions, statistical analysis of the DNA-array data as previously described revealed a total of 536 DEGs in wild type and *Δixr1 S. cerevisiae* strains. The whole data sheet of these DEGs is available through Table S2 and their clustering, using Genespring in the conditions previously specified, is shown (figure 1b).

71

Among DEGs detected in hypoxia, 358 genes are expressed at higher levels in the mutant than in the wild type strain. Data obtained by functional clustering of these DEGs is included in Table 3.
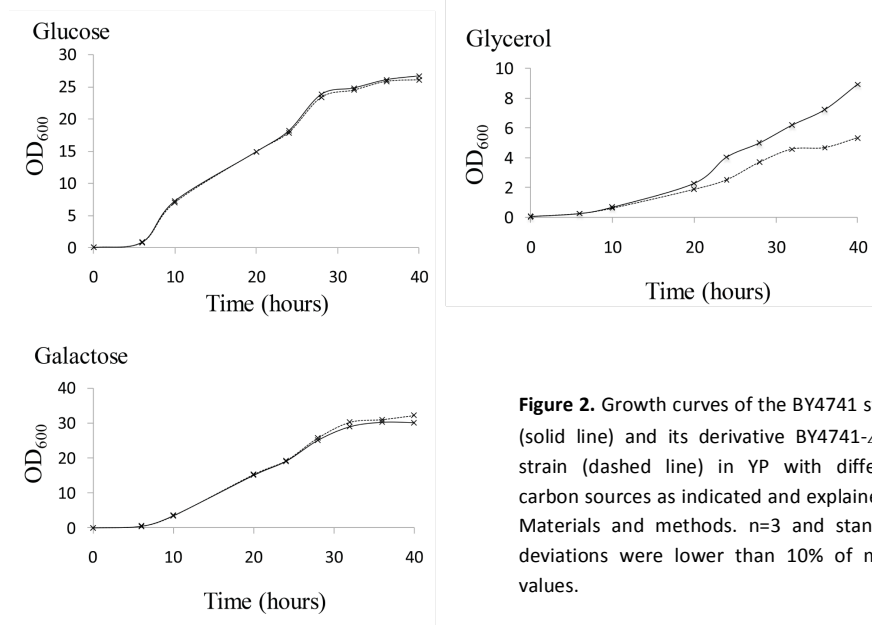
Figure 2. Growth curves of the BY4741 strain (solid line) and its derivative BY4741-*Δixr1* strain (dashed line) in YP with different carbon sources as indicated and explained in Materials and methods. n=3 and standard deviations were lower than 10% of mean values.

Table 1. Enriched functional gene groups whose expression in the *Δixr1* strain is higher than the wild type (BY4741) during normoxia

MIPS functional group	*p*-value	Genes in category from cluster	*κ*
Sulfate assimilation [01.02.03.01]	3.67×10^{-3}	MET10 MET5 MET14 MET1 MET16	5
Stress response [32.01]	0.000355	SSE2 PAU2 STF2 TIR3 UBA1 DAK1 TSL1 RIM11 ALD3 PAI3 HOR7 SIW14 YGP1 DDR2 TIR4 ASR1	16
Protease inhibitor [18.02.01.02.03]	0.000495	TFS1 PAI3 PBI2	3
Vacuolar protein degradation [14.13.04.02]	0.000520	APE3 VID30 PAI3 TRE2	4
Biosynthesis of homocysteine [01.01.06.05.01.01]	0.001638	MET10 STR3 MET5	3
Oxidative stress response [32.01.01]	0.004212	GRX1 TSA2 GRX2 MXR1 HSP12 CTT1 GRE1	7
Pyrimidine metabolism [01.03.04]	0.005477	HNT2 GNA1 THI5 PUS5 URA10	5

Table 2. Enriched functional gene groups whose expression in the *Δixr1* strain is lower than in the wild type (BY4741) during normoxia

MIPS functional group	*p*-value	Genes in category from cluster	κ	*f*
Biosynthesis of valine [01.01.11.03.01]	2.14×10^{-4}	*ILV6 BAT1 ILV3 ILV5*	4	6
Biosynthesis of isoleucine [01.01.11.02.01]	1.33×10^{-5}	*BAT1 ILV3 ILV5*	3	5
Sugar, glucoside, polyol and carboxylate catabolism [01.05.02.07]	2.99×10^{-2}	*CDC19 TDH3 XKS1 TDH1 TDH2 GPM1 HPF1*	7	81
Biosynthesis of leucine [01.01.11.04.01]	7.27×10^{-2}	*BAT1 ILV5 LEU9*	3	8
Glycolysis and gluconeogenesis [02.01]	8.44×10^{-5}	*CDC19 TDH3 TDH1 TDH2 GPM1*	5	41
C4-dicarboxylate transport [20.01.03.03]	0.000733	*OAC1 DIC1*	2	4
Ion transport [20.01.01]	0.002509	*FIT1 FIT3*	2	7
Purin metabolism [01.03.01]	0.004415	*YBR284W ADO1 IMD4*	3	30
Phosphate transport [20.01.01.07.07]	0.006384	*MIR1 DIC1*	2	11
Anion transport [20.01.01.07]	0.006384	*FZF1 MEP2*	2	11
Drug/toxin transport [20.01.27]	0.009249	*QDR3 ARN2 TPO4*	3	39

Representative groups include genes of the biosynthesis of BC amino acids, stress response, sugar transport and metabolism. Genes related to nitrogen, sulphur, selenium and phosphate metabolism as well as purine anabolism are also included. Besides, this group of DEGs is also enriched in different genes related to metabolic flux regulation (MIPS codes [01.02.07], [01.01.13] and [01.05.25]). During hypoxic growth 178 genes are expressed at higher levels in the wild type than in the *Δixr1* mutant strain. Enriched functional groups, as revealed by analysis with FunSpec (Table 4), include genes from the metabolism of methionine and sulphate assimilation, genes from the oxidative stress response or genes for DNA synthesis and replication among others.

Table 3. Enriched functional gene groups whose expression in the *Δixr1* strain is higher than in the wild type (BY4741) during hypoxia

MIPS functional group	*p*-value	Genes in category from cluster	κ	f
Biosynthesis of valine [01.01.11.03.01]	0.000117	*ILV6 ILV3 ILV5 ILV2*	4	6
Stress response [32.01]	0.000160	*YBR016W TIP1 TPS2 SED1 PAM1 SSD1 ECM10 TIR1 RSP5 NSR1 HSP150 PIR3 PIR1 UTH1 PSR1 UBI4 PSR2 MCM1 YCK2 YNL234W TIR4*	21	162
Sugar transport [20.01.03.01]	0.000181	*MAL31 MPH2 HXT7 HXT4 HXT1 HXT5 RGT1 HXT2*	8	31
Detoxification [32.07]	0.000323	*ADH5 ROG3 FZF1 TPO2 SNG1 AZR1 ERC1 TPO1 MCM1 AQR1 ALR1 TPO4 SVS1*	13	80
C-compound and carbohydrate transport [20.01.03]	0.000361	*YBR241C GLK1 YAT2 DUR3 AQR1 ITR2 ODC1 SAM3*	8	34
Alcohol fermentation [02.16.01]	0.000408	*ADH5 PDC5 ADH3 ADH2 ALD4*	5	13
Biosynthesis of leucine [01.01.11.04.01]	0.000500	*LEU1 ILV5 LEU4 LEU9*	4	8
Amino acid/amino acid derivatives transport [20.01.07]	0.000570	*AGP2 BAP3 HIP1 AVT1 LYP1 NRT1 ODC1 DIP5 SAM3*	9	45
Drug/toxin transport [20.01.27]	0.000965	*TPO2 SNG1 AZR1 ERC1 QDR2 TPO1 AQR1 TPO4*	8	39
Cellular import [20.09.18]	0.001032	*MAL31 GLK1 BAP3 HXT7 HIP1 HXT4 HXT1 HXT5 HXT2 FET4 LYP1 ITR2 ALR1*	13	90
Nitrogen, sulfur and selenium metabolism [01.02]	0.002222	*GDH3 AMD2 IRC7 TRR2 NIT1 DAL81 NIT2 ARG1 GDH1*	9	54
Phosphate metabolism [01.04]	0.003188	*CDC19 HIS4 GLK1 PTC6 TPS2 VHS1 SAP1 VTC2 PKP2 NPY1 MUQ1 RTS3 BUB1 PRS3 PKP1 YAK1 ADO1 PRR1 RPT1 KKQ8 PSR1 PSR2 HOG1 CNA1 NAM7 PFK2 HEF3 YCK2 CBK1 NPR1 SSB2 CLA4 THI80 LCB4 VPH1*	35	401
Purine anabolism [01.03.01.03]	0.003996	*HIS4 ADE5,7 ADE3 PRS3 IMD2 IMD3*	6	29
Amine / polyamine transport [20.01.11]	0.005510	*TPO2 DUR3 TPO1 TPO4*	4	14
Regulation of nitrogen, sulfur and selenium metabolism [01.02.07]	0.005578	*FZF1 DAL81 ARG81 ARG80 MCM1*	5	22
C-compound and carbohydrate metabolism [01.05]	0.006478	*ACH1 ADH5 KTR4 ILV6 GLK1 AAD3 TPS2 ANP1 YEL047C YAT2 LPD1 MIG1 DSE2 RGT1 PDC5 ADH3 ILV2 GPD2 ALG6 GLO4 ATF1 ERR1*	22	230
Regulation of amino acid metabolism [01.01.13]	0.007761	*CCR4 ARO8 DAL81 ARG81 ARG80 MCM1*	6	33
Regulation of C-compound and carbohydrate metabolism [01.05.25]	0.007993	*CCR4 SNF5 ACK1 YFL052W MIG1 HAP2 RGT1 HOG1 GSF2 MCM1 PFK2 MKS1 POP2 GAL11*	14	126

Table 4 Enriched functional gene groups whose expression in the *Δixr1* strain is lower than in the wild type (BY4741) during hypoxia

MIPS functional group	*p*-value	Genes in category from cluster	*κ*	*f*
Metabolism of methionine [01.01.06.05]	1.41E-02	SAM2 MET3 MET1 SAM1 ADI1 MET16	6	21
Sulfate assimilation [01.02.03.01]	3.36E-02	MET3 MET5 MET1 MET16	4	8
Oxidative stress response [32.01.01]	0.000102	GPX2 GRX1 MXR1 HSP12 TRX2 HYR1 GTT1 FMP46	8	55
Protease inhibitor [18.02.01.02.03]	0.000188	YHR138C PAI3 PBI2	3	5
Peroxidase reaction [32.07.07.05]	0.000633	GPX2 GRX1 HYR1	3	7
Detoxification by modification [32.07.03]	0.000992	GRX1 GTT1 AYT1	3	8
Oxygen and radical detoxification [32.07.07]	0.003598	MXR1 TRX2 SLN1	3	12
Secondary metabolism [01.20]	0.003598	YJR096W AYT1 GCY1	3	12
Mitotic cell cycle [10.03.01.01]	0.004585	AME1 SPO12 IPL1	3	13
Cytoplasmic and nuclear protein degradation [14.13.01]	0.005430	YHR138C YUH1 PAI3 PBI2 UBP2 AXL1	6	60
DNA synthesis and replication [10.01.03]	0.006572	CDC28 TRX2 EST3 CDC8 ABF2 RRG9 CLB5	7	82
Glutathione conjugation reaction [32.07.07.03]	0.006925	GRX1 GTT1	2	5

3.3.- Effects of *IXR1* deletion on the *S. cerevisiae* transcriptome after a shift from aerobic to hypoxic growth

In these experiments cells were grown under normoxia until OD_{600} = 0.8 and then hypoxia was generated by bubbling nitrogen in the flasks at constant pressure during 3 hours. In these conditions, statistical analysis of the DNA-array data as previously described revealed a total of 502 DEGs in wild type and *Δixr1 S. cerevisiae* strains. The whole data sheet of these DEGs is available through Table S3

and their clustering, using Genespring in the conditions previously specified, is shown (Figure 1c).

In this experimental condition only 20 genes are expressed at higher levels in the mutant than in the wild type strain, while the majority of genes, 482, are expressed at higher levels in the wild type than in the mutant.

Several genes encoding for different subunits of cytochrome *c* oxidase or related to petite phenotype caused by mitochondrial dis-function are also found among the genes expressed at higher levels in the wild type than the mutant. However, analysis with FunSpec (Tables 5 and 6) reveals that few functional groups are over-represented among these DEGs. Some genes related with cell aging or dead are included as significant groups.

A comparison of the DEGs that are affected by the *Δixr1* mutation during long-term hypoxia (anaerobic jars) and during the shift from aerobic to hypoxic conditions reveals that there is little overlap. Only eight genes are co-regulated in both conditions. Four genes are expressed at lower levels in the mutant than in the wild type strain (*PSA1*, *SED1*, *DSE2* and YJL016) and the others (*GPX2*, YFR049W, YGL177W, YOR097C) are expressed at higher levels in the mutant than in the wild type strain.

Table 5. Enriched functional gene groups whose expression in the *Δixr1* strain is higher than in the wild type (BY4741) after the hypoxic shift

MIPS functional group	*p*-value	Genes in category from cluster	κ	f
Fatty acid binding (e.g. acyl-carrier protein) [16.13.03]	0.002877	ACB1	1	1
Apoptosis (type I programmed cell death) [40.10.02]	0.005747	CDC48	1	2
Toxins [32.05.05.01]	0.005747	FLR1	1	2
Lipid/fatty acid transport [20.01.13]	0.006906	PDI1 ACB1	2	44

Table 6. Enriched functional gene groups whose expression in the *Δixr1* strain is lower than in the wild type (BY4741) after the hypoxic shift

MIPS functional group	*p*-value	Genes in category from cluster	*κ*	*f*
Cell aging [40.20]	0.003286	*SNF1 MPT5 PHB2 LAG1 ZDS1 RAS2 LAG2*	7	28
Regulation by modification [18.01.01]	0.003327	*SNF1 PTP3 CNA1 SIW14 SSU72*	5	15
Cell death [40.10]	0.006164	*LAG1 RAS2 LAG2 MAM3*	4	11

3.4.- Effect caused by *IXR1* deletion upon expression of hypoxic genes or genes essential to anaerobic growth.

The first transcriptional function of Ixr1 described in the literature was as aerobic repressor of the hypoxic gene *COX5B* (Lambert *et al*. 1994), however experimental data supporting that Ixr1 regulates mRNA levels of other hypoxic genes are limited to a few reports (Bourdineaud *et al*. 2000; Castro *et al*. 2010a; Castro *et al*. 2010b).

Since hypoxic genes are distributed over different functional groups in FunCatDB, the analysis of our data with the FunSpec tool, as above commented, was not expected to reveal the existence of a hypoxic response mediated by Ixr1. We have therefore compared the expression of a set of genes in the wt and *Δixr1* mutant strains during aerobic or hypoxic growth. The genes for this analysis (the complete set of analyzed genes is listed in Table S4) were selected among those previously defined to be induced during the anoxic/hypoxic response in *S. cerevisiae* (Kwast *et al*. 2002; Becerra *et al*. 2002) as well as among those which are essential during anaerobic or hypoxic growth according to our search in SGD (*Saccharomyces* Genome Data Bank; http://www.yeastgenome.org/). The results obtained revealed that significant changes affected by *IXR1* deletion were not observed during the shift from normoxia to hypoxia. Therefore, only the data corresponding to normoxic or long-term hypoxic growth were included in the data

shown in Table 7. Only eight genes from this selection (Table S4) are activated by Ixr1 in normoxia (*GPM1, EGD1*, *CWP1*, *GTT1*, *TDH3*, *HEM3, CDC19* and *SKN7*) and they have diverse functions. However, the number of genes repressed during normoxia is higher (14 genes) and includes several members of the seripauperin multigene family, negatively regulated by oxygen and heme. Among the genes repressed by Ixr1 during hypoxic conditions (22 genes) there are also genes from the seripauperin multigene family, genes related to fatty acid (*OLE1*) or ergosterol synthesis (*ERG4*, *ERG11, HES1*) and transcriptional repressors related to the hypoxic response (*MOT3*, *RIM101*).

Table 7. Genes related to the aerobic/hypoxic response, which change their expression in wt and *Δixr1* strains during normoxic (O_2) or hypoxic (Hy) growth

O_2	Hy	GENE	ORF	O_2 FC*	Hy FC*	O_2	Hy	GENE	ORF	O_2 FC*	Hy FC*
green	green	*GPM1*	YKL152C	+3.5	+1.8	black	red	*OLE1*	YGL055W	nc	-10
black	green	*EGD1*	YPL037C	nc	+2.6	black	red	*TIR1*	YER011W	nc	-5
black	green	*CWP1*	YKL096W	nc	+2.2	black	red	*ADH2*	YMR303C	nc	-5
black	green	*GTT1*	YIR038C	nc	+2.1	black	red	*DAN1*	YJR150C	nc	-5
green	black	*TDH3*	YGR192C	+7.4	nc	black	red	*ADH5*	YBR145W	nc	-5
green	black	*HEM3*	YDL205C	+4.9	nc	black	red	*GLK1*	YCL040W	nc	-5
green	black	*CDC19*	YAL038W	+2.3	nc	black	red	*MOT3*	YMR070W	nc	-5
green	black	*SKN7*	YHR206W	+2.1	nc	black	red	*ERG4*	YGL012W	nc	-3.3
red	black	*HEM4*	YOR278W	-10	nc	black	red	*ERG11*	YHR007C	nc	-3.3
red	black	*MSC1*	YML128C	-5	nc	black	red	*HES1*	YOR237W	nc	-3.3
red	black	*PAU2*	YEL049W	-5	nc	black	red	*YNL234W*	YNL234W	nc	-3.3
red	black	*HXK1*	YFR053C	-5	nc	black	red	*TPS2*	YDR074W	nc	-3.3
red	black	*SPI1*	YER150W	-3.3	nc	black	red	*PAU21*	YOR394W	nc	-3.3
red	black	*TIR3*	YIL011W	-3.3	nc	black	red	*YEL047C*	YEL047C	nc	-3.3
red	black	*GSY2*	YLR258W	-2.5	nc	black	red	*GPD2*	YOL059W	nc	-3.3
red	black	*DBP7*	YKR024C	-2.5	nc	black	red	*ADH3*	YMR083W	nc	-2.5
red	black	*UGP1*	YKL035W	-2.5	nc	black	red	*RIM101*	YHL027W	nc	-2.5
red	black	*KGD1*	YIL125W	-2	nc	black	red	*GLC3*	YEL011W	nc	-2.5
red	black	*SSE2*	YBR169C	-2	nc	black	red	*LPD1*	YFL018C	nc	-2
red	black	*HSP78*	YDR258C	-2	nc	black	red	*PFK2*	YMR205C	nc	-2
red	black	*AAC1*	YMR056C	-2	nc	black	red	*GSY1*	YFR015C	nc	-2
red	red	*TIR4*	YOR009W	-3.3	-10						

Ixr1 activation is indicated in code green and repression in red; black, no significant change

*FC Fold change, activation (wt/Δixr1, +) or repression (Δixr1/wt, -); nc, no significant change.

Table 8. Gene expression changes in wild type and *Δixr1 S. cerevisiae* strains observed in the qRT-PCR approach

Growth condition	Gene	Fold change*	*p*-value
Normoxia	MET1	+2.0	0.000035
Normoxia	MET5	+1.5	0.002383
Normoxia	MET16	+1.5	0.001295
Normoxia	MET28	nc	-
Normoxia	MET32	+2.2	0.004081
Normoxia	GRX1	+5.8	0.025885
Normoxia	MXR1	-1.4	0.034799
Normoxia	HSP12	+1.7	0.023660
Normoxia	BAT1	-1.8	0.011491
Normoxia	BAT2	+2.4	0.015641
Normoxia	GPM1	-2.6	0.003157
Normoxia	HEM4	+1.6	0.013508
Normoxia	PAU2	+1.9	0.012722
Normoxia	TIR3	+1.6	0.014961
Normoxia	TIR4	+1.6	0.017020
Normoxia	SPI1	+2.1	0.000447
Hypoxia	MET1	-1.6	0.017822
Hypoxia	MET5	-1.6	0.026961
Hypoxia	MET16	-1.7	0.006591
Hypoxia	MET28	nc	-
Hypoxia	MET32	-1.8	0.000015
Hypoxia	GRX1	-1.5	0.035506
Hypoxia	MXR1	-1.9	0.022183
Hypoxia	HSP12	+3.2^{ns}	0.111730
Hypoxia	BAT1	+1.6	0.002467
Hypoxia	BAT2	+2.1	0.025333
Hypoxic shift	COX7	nc	-
Hypoxic shift	COX9	-1.3	0.042918
Hypoxic shift	MSS18	-1.3	0.028839

* Fold change, activation (wt/Δixr1, +) or repression (Δixr1/wt, -); nc, no change; ^{ns} no significant change

3.5.- qRT-PCR validation of changes in expression of selected genes

The 20 genes selected for this validation as well the sequences of the designed primers are listed in Table S5. The results obtained are summarized in Table 8 and they confirmed the data obtained in the DNA-array approach with the following exceptions. Among the genes selected from the sulphate assimilation pathway, *MET28* was not repressed by Ixr1 during normoxia using this approach. Among the genes selected from the oxidative stress response, *MXR1* was not activated in normoxia and *HSP12* was not repressed during hypoxic growth. Finally, *COX7* was not repressed during hypoxic growth.

4.- DISCUSSION

Ixr1 is a high mobility group (HMG) transcription factor first identified by its ability to bind DNA modified by the anticancer drug cisplatin (cis-diaminedichloriplatinum-II) (Brown *et al*., 1993). Ixr1 cellular function has been related to the regulation of diverse genes related to the hypoxic response (Lambert *et al*. 1994; Bourdineaud *et al*., 2000; Castro *et al*., 2010a; Castro *et al*., 2010b) or to the maintenance of an adequate supply and balance of dNTPs for DNA synthesis and repair (Tsaponina *et al*., 2011). We have studied the effects of *IXR1* deletion upon the transcriptome of *S. cerevisiae* cells from the BY4741 strain grown in aerobic and hypoxic conditions or after a shift from aerobic to hypoxic conditions.

4.1.- Ixr1 and the control of hypoxic genes

It has been previously reported that Ixr1 causes aerobic repression of three hypoxic genes *COX5B* (Lambert *et al*., 1994), *TIR1* (Bourdineaud *et al*., 2000) and *HEM13* (Castro *et al*., 2010a). In the normalized data sheet obtained from the DNA-array expression analysis, the expression of *COX5B* during aerobic growth increases 2.3 folds in the *Δixr1* strain in reference to the wild type. In a parallel comparison, *TIR1* and *HEM13* expression also increases 2.1 and 1.5 folds

respectively. During hypoxic growth, it has been reported that Ixr1 is an activator of *HEM13* (Castro *et al*., 2010a). In the normalized data sheet obtained from the DNA-array expression analysis, the hypoxic expression of *HEM13* decreases 1.6 fold in the *Δixr1* strain with reference to the wild type strain. It can be concluded that previously reported changes in gene expression between the wild type and *Δixr1* strains obtained by conventional northern analysis or by fusion of their promoters to reporter genes are also observed with the DNA-array, although relative changes (fold ratio of repression or activation) are minor in this approach than in previous ones (Table 9). This comparison of results obtained by different approaches suggests that the methods used in this work allow the identification of genes that are differentially expressed in wild type and *Δixr1* strains in the growth conditions selected although, certainly, not all of them because those whose expression-change is minor than two fold were not included for further analysis.

Table 9. Gene expression changes in wild type and *Δixr1 S. cerevisiae* strains observed in the DNA array approach (values in arbitrary units after normalization) or previously published data

Growth condition	Gene	DNA array Fold change*	Reported change*	Method	Reference
Normoxia	COX5B	-2.3	Undetectable in wt; it increases in *Δixr1*	Northern	Lambert *et al*. (1994)
Normoxia	TIR1	-2.1	Fold change -8.5	Reporter gene	Bourdineaud *et al*. (2000)
Normoxia	HEM13	-1.5	Undetectable in wt; it increases in *Δixr1*	Northern	Castro *et al*. (2010b)
Hypoxia	HEM13	+1.6	Fold change +1.9	Northern	Castro *et al*. (2010b)

*Ixr1 Activation (+) Ixr1 Repression (-)

We have also analyzed the regulation of a set of genes previously related to the aerobic/hypoxic response in *S. cerevisiae*. Heme is a necessary cofactor for the cytochromes, catalases, and P450 enzymes of sterol synthesis, and its synthesis

requires oxygen (Labbe-Bois & Labbe, 1990). Some of the heme-biosynthetic genes including *HEM4, HEM13* and *HEM15*, are up-regulated under anoxia, although only *HEM13* and *HEM15* are repressed by Rox1 (Kwast *et al*., 2002). Data in Tables 7-8 indicate that the aerobic repression of *HEM4* could be mediated by Ixr1. Several genes from Table 7 repressed by Ixr1 during normoxia belong to the seripauperin family (*PAU2, TIR3* and *TIR4*). Up-regulation of these three genes, as well as *SPI1*, in the *Δixr1* strain has been verified by qRT-PCR (Table 8). The hypothesis that the anaerobic induction of this *PAU* gene family influence cell wall porosity and may be necessitated in part by the changes in membrane fluidity that occur under anaerobiosis has been raised (reviewed in Smits *et al*., 1999). However the genes belonging to this family are apparently regulated by a Rox1-independent mechanism (Kwast *et al*., 2002) and its regulation has been associated to other transcriptional factors (Abramanova *et al*., 2001a; Abramanova *et al*., 2001b). Our data suggest that at least part of the members of this *PAU* gene family as well as genes related to cell wall, like *SPI1* (Hamada *et al*., 1998), are repressed by Ixr1 during aerobic growth.

According to data presented in this work, Ixr1 has a negative effect upon mRNA levels of the majority of the differentially expressed genes (DEGs) observed both during normoxic or hypoxic growth (Figure 1a and 1b). However the opposite effect is observed in the transition from normoxic to hypoxic conditions (Figure 1c). The molecular mechanisms of the positive or negative effects of Ixr1 upon expression of different genes, or even if these effects are direct or indirect in the assayed conditions, are not known. Nevertheless, this dual behavior has been previously found studying the regulatory effect of Ixr1 upon other promoters (Bourdineaud *et al*., 2000; Castro *et al*., 2010a) and is actually under study in our laboratory. Besides, the Ixr1-mediated response to the hypoxic shift notoriously differs from the response to long-term hypoxia as shown by the absence of overlap between DEGs in both conditions (Tables S2 and S3). We have observed (Table S3)

that several genes encoding for subunits of the enzyme cytochrome *c* oxidase (*COX7, COX8* and *COX9*) or related to the process of RNA splicing of *COX1* transcripts (*MSS18*) are expressed at lower levels in the *Δixr1* strain than in the wild-type strain during the hypoxic shift. The qRT-PCR validation (Table 8) confirms that *COX9* and *MSS18* are at least regulated by Ixr1. This might indicate that Ixr1 transcriptional regulation takes part in the mechanism of sensing oxygen levels that has been attributed to cytochrome *c* oxidase in eukaryotic systems (Kwast *et al*., 1999). It has been described that the turnover rate of cytochrome *c* oxidase increases during the anaerobic shift and that this increase cannot be explained merely by replacing the aerobic isoform Va of cytochrome *c* oxidase subunit V with the more active hypoxic isoform Vb (David and Poyton, 2005). Since Ixr1 activates transcription of several subunits of the cytochrome c oxidase complex during the hypoxic shift (this work), a correlation with increased turnover might exist.

4.2.- Ixr1 and the control of the glycolytic pathway

Several genes related to the glycolytic pathway are positively regulated by Ixr1 during normoxic growth (Table 2) and we have verified (Figure 2) that differences of growth-rate between BY4741 wild type and *Δixr1* strains are not detected in glucose or galactose, but are observed in glycerol. Probably growth in hexoses is possible through the pentose phosphate pathway, even when the glycolytic pathway is down-regulated in the *Δixr1* strain. Oppositely, growth in glycerol requires *GPM1* which is down-regulated in the *Δixr1* strain (Tables 2 and 8).

4.3.- Ixr1 and the control of the sulphate assimilation pathway

After functional clustering of DEGs we have found that Ixr1 regulates a number of genes related to oxidative stress response, cell-wall composition, cell-

energetic, the metabolism of sulphur and branched-chain (BC) amino acids and DNA synthesis in a way that is highly dependent on oxygen availability. Interestingly, the sign of Ixr1-mediated regulation, positive or negative, change when we compare the expression of functional groups of genes in cells grown under normoxia or hypoxia (compare Tables 1 and 4). Even, the same genes are oppositely regulated by Ixr1 depending of these growth conditions. The genes from the sulphate assimilation pathway illustrate this mechanism of regulation mediated by Ixr1 (Figure 3). Genes from this functional group are up-regulated in the mutant during normoxic growth (repressed by Ixr1) and down-regulated in the mutant during hypoxic growth (activated by Ixr1). Besides the genes shown in Tables 1 and 4, a direct inspection of Table S1 allows the identification of two other genes, encoding transcriptional factors related to the sulphate assimilation pathway, *MET28* and *MET32,* which were not present in the list of genes obtained using FunSpec. Three genes from this pathway are oppositely regulated by Ixr1 accordingly to oxygen levels. In the *Δixr1* mutant strain, the genes *MET5*, *MET1* and *MET16* are up-regulated during normoxia but are down-regulated during hypoxia (Tables 1, 4 and 8). *MET5* encodes for the beta-subunit of the enzyme sulphite reductase (Thomas *et al.*, 1992). *MET1* encodes for the enzyme S-adenosyl-L-methionine uroporphyrinogen III transmethylase, involved in the biosynthesis of siroheme (Hansen *et al*. 1997), a prosthetic group used by sulphite reductase; required for sulphate assimilation and methionine biosynthesis.

MET16 encodes for the 3'-phosphoadenylsulphate reductase that reduces 3'-phosphoadenylyl sulphate to adenosine-3',5'-bisphosphate and free sulphite using reduced thioredoxin as co-substrate (Thomas *et al.*, 1990). An essential function of the sulphur-pathway is its involvement, through S-adenosylmethionine, in the biosynthesis of polyamines (Thomas & Surdin-Kerjan, 1997). Other essential function of this pathway is glutathione biosynthesis and therefore the defense against oxidative stress. Indeed, other group of genes that is co-regulated with the

genes from the sulphate assimilation pathway is related to the oxidative stress response and three genes of glutaredoxins are up-regulated in the *Δixr1* mutant during normoxic growth (Tables 1 and 8), but down-regulated during hypoxic growth (Tables 4 and 8).
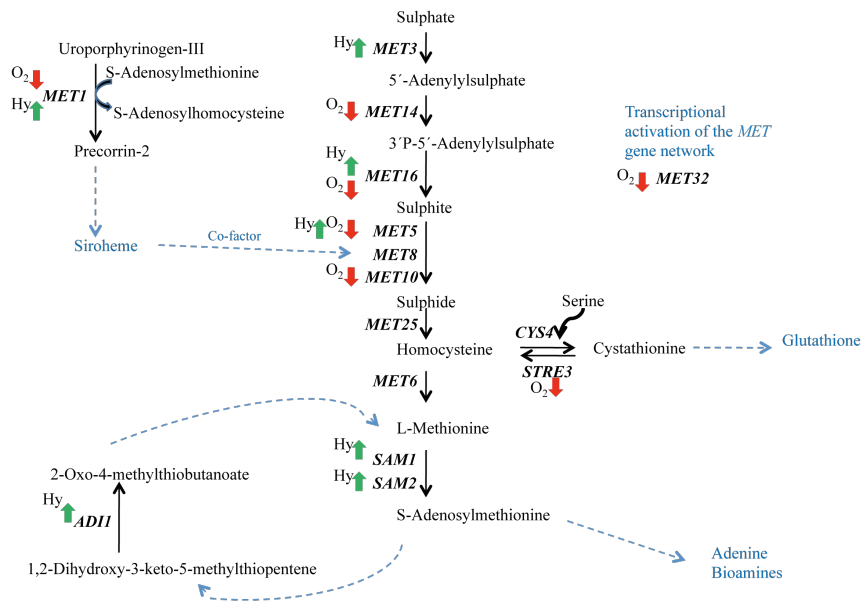
Figure 3. Regulatory role of Ixr1 on the sulphate assimilation pathway and interconnected routes as deduced from the transcriptome analysis. Dashed lines indicate that several individual steps have been omitted or, when signaled, that the compound is a cofactor of other catalyzed reaction. Down-regulation (down directed arrow) by Ixr1 during normoxia (O_2) or up-regulation (up directed arrow) by Ixr1 during Hypoxia (Hy) is indicated for each regulated gene.

Grx1p is a glutathione-dependent disulphide oxidoreductase induced by hydroperoxide and superoxide-radicals that protects cells from oxidative damage (Luikenhuis *et al.*, 1998; Collinson *et al.*, 2002). It has been previously reported that low-oxygen levels induce an oxidative stress response in *S. cerevisiae* (Dirmeier *et*

al., 2002) and the data above commented reveals that Ixr1-mediated regulation during hypoxia might enhance a protective defense against oxidative stress.

4.4.- Ixr1 and the control of the biosynthesis of the branched chain amino acids

Among the functional gene-groups that are positively regulated by Ixr1 during normoxic growth and negatively regulated during hypoxic growth are those related to the biosynthesis of the BC amino acids L-valine, L-isoleucine and L-leucine (VIL). The metabolic pathways and the observed effects of Ixr1 upon transcript levels of the corresponding genes are summarized in Figure 4. The regulatory effect caused by Ixr1 during normoxic growth would favor the BC amino acid anabolism, while during hypoxia their biosynthesis would be repressed. *BAT1* is one of the genes down-regulated in the *Δixr1* strain during normoxic growth (Tables 2 and 8) and it encodes for the BC amino acid aminotransferase that catalyzes the transfer of amino groups between the amino acids valine, isoleucine, and leucine and their corresponding α-keto-acids. α-keto-acids are biosynthetic precursors of fusel alcohols, which influence the aroma and flavor of yeast-derived fermentation products. It has been recently proposed that in *S. cerevisiae BAT1* and *BAT2* are paralogous genes, 77% identical to each other, coming from a multifunctional common ancestor (Colón *et al*., 2011). The aminotransferases encoded by these genes are required for both BC amino acid biosynthesis and the Ehrlich pathway amino acid catabolism, but along evolution there has been a specialization of their function and cellular localization. *BAT1* encodes the mitochondrial and *BAT2* the cytosolic isoforms respectively. *BAT1* is highly expressed under biosynthetic conditions, while *BAT2* expression is highest under catabolic conditions (Colón *et al*., 2011). It is possible that Ixr1 mediated regulation might also contribute to this change in the expression of these paralogous genes and, in general, to other genes in order to accommodate to the different catabolic or anabolic fluxes predominant during normoxic or hypoxic growth. During normoxic growth pyruvate oxidation for energy production does not limit its use in

biosynthetic pathways, however during hypoxia, due to the minor efficiency of fermentation than respiration to produce energy the available levels of pyruvate for biosynthesis might become a limiting factor for the anabolic pathways.

4.5.- Interest of Ixr1 transcriptional control for biotechnological purposes

In conclusion, the transcriptome analysis here presented suggests that Ixr1 contributes to the changes necessary to adapt yeast cells to hypoxia, but also to oxidative stress conditions. Besides, Ixr1 regulates the levels of transcripts encoding enzymes and regulators of metabolic pathways that are profoundly remodeled to adapt the use and production of cellular energy to oxygen availability. Remodeling of the sulphate assimilation pathway is of interest in wine and beer biotechnology.

Metabolic pathways producing α-keto-acids are also of biotechnological importance, because they are biosynthetic precursors of fusel alcohols, which influence the aroma and flavor of yeast-derived fermentation products. As above discussed, the transcriptome analysis here presented indicates that Ixr1 is important for the regulation of these metabolic pathways. Therefore, Ixr1 mutations and their effects upon regulation of target genes are of biotechnological relevance in the design and evaluation of industrial yeast strains.

5.- ACKNOWLEDGEMENTS
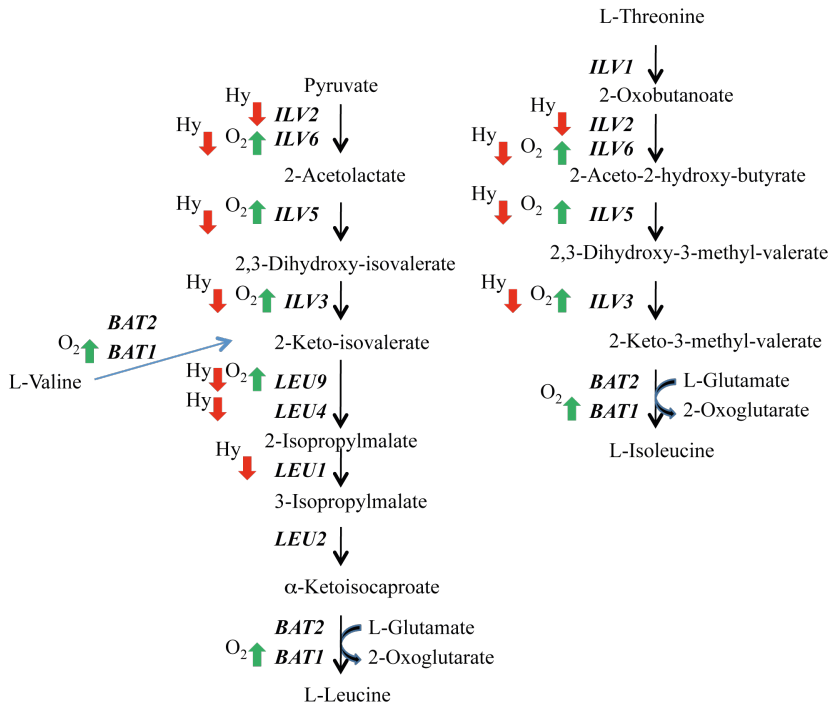
Figure 4. Regulatory role of Ixr1 on the branched chain amino acids metabolism as deduced from the transcriptome analysis. Up-regulation (up directed arrow) by Ixr1 during normoxia (O_2) or down-regulation (down directed arrow) by Ixr1 during Hypoxia (Hy) is indicated for each regulated gene.

6.- REFERENCES

Abramova N.E., B.D. Cohen, O. Sertil, R. Kapoor, K.J. Davies, C.V. Lowry (2001a) "Regulatory mechanisms controlling expression of the *DAN/TIR* mannoprotein genes during anaerobic remodeling of the cell wall in *Saccharomyces cerevisiae*" Genetics **157**(3):1169-1177

Abramova N.E., O. Sertil, S. Mehta, C.V. Lowry (2001b) "Reciprocal regulation of anaerobic and aerobic cell wall mannoprotein gene expression in *Saccharomyces cerevisiae*" Journal of Bacteriology **183**(9):2881-2887

Alberola T.M., J. García-Martínez, O. Antúnez, L. Viladevall, A. Barceló, J. Ariño, J.E. Pérez-Ortín (2004) "A new set of DNA macrochips for the yeast *Saccharomyces cerevisiae*: features and uses" International Microbiology **7**(3):199-206

Becerra M., L.J. Lombardía-Ferreira, N.C. Hauser, J.D. Hoheisel, B. Tizon, M.E. Cerdán (2002) "The yeast transcriptome in aerobic and hypoxic conditions: effects

of *hap1*, *rox1*, *rox3* and *srb10* deletions" Molecular Microbiology **43**(3):545-555 (*Erratum in*: Mol Microbiol 45:265)

Bourdineaud, J.P., G. De Sampaio, G.J. Lauquin (2000) "A Rox1-independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the *SRP1/TIR1* gene" Molecular Microbioly **38**(4):879-890

Brown S.J., P.J. Kellett, S.J. Lippard (1993) "Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin" Science **261**(5121):603-605

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso, M.E. Cerdán (2010a) "Ixr1p regulates oxygen-dependent *HEM13* transcription" FEMS Yeast Research **10**(3):309-321

Castro-Prego M., M. Lamas-Maceiras, P. Soengas, I. Carneiro, M.I. González Siso, M.E. Cerdán (2010b) "Regulatory factors controlling transcription of *Saccharomyces cerevisiae IXR1* (*ORD1*) by oxygen levels. A model of transcriptional adaptation from aerobiosis to hypoxia implicating *ROX1* and *IXR1* cross-regulation" Biochemical Journal **425**(1):235-243

Collinson E.J., G.L. Wheeler, E.O. Garrido, A.M. Avery, S.V. Avery, C.M. Grant (2002) "The yeast glutaredoxins are active as glutathione peroxidases" The Journal of Biological Chemistry **277**(19):16712-16717

Colón M., F. Hernández, K. López, H. Quezada, J. González, G. López, C. Aranda, A. González (2011) "*Saccharomyces cerevisiae* Bat1 and Bat2 aminotransferases have functionally diverged from the ancestral-like *Kluyveromyces la*ctis orthologous enzyme" PLoS One **6**(1):e16099

Chomczynski P. & N. Sacchi (1987) "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction" Analytical Biochemistry **162**(1):156-159

David P.S. & R.O. Poyton (2005) "Effects of a transition from normoxia to anoxia on yeast cytochrome c oxidase and the mitochondrial respiratory chain: implications for hypoxic gene induction" Biochimica et Biophysica Acta **1709**(2):169-180

Deckert J., R.A. Khalaf, S.M. Hwang, R.S. Zitomer (1999) "Characterization of the DNA binding and bending HMG domain of the yeast hypoxic repressor Rox1" Nucleic Acids Research **27**(17):3518-3526

Deckert J., A.M. Rodríguez Torres, J.T. Simon, R.S. Zitomer (1995) "Mutational analysis of Rox1, a DNA-bending repressor of hypoxic genes in *Saccharomyces cerevisiae"* Molecular and Cellular Biology **15**(11):6109-6117

Dirmeier R., K.M. O'Brien, M. Engle, A. Dodd, E. Spears, R.O. Poyton (2002) "Exposure of yeast cells to anoxia induces transient oxidative stress. Implications for the induction of hypoxic genes" The Journal of Biological Chemistry **277**(38):34773-34784

Hamada K., S. Fukuchi, M. Arisawa, M. Baba, K. Kitada (1998) "Screening for glycosylphosphatidylinositol (GPI)-dependent cell wall proteins in *Saccharomyces cerevisiae*" Molecular Genetics and Genomics **258**(1-2):53-59

Hansen J., M. Muldbjerg, H. Chérest, Y. Surdin-Kerjan (1997) "Siroheme biosynthesis in *Saccharomyces cerevisiae* requires the products of both the *MET1* and *MET8* genes" FEBS Letters **401**(1):20-24

Hauser N.C., M. Vingron, M. Scheideler, B. Krems, K. Hellmuth, K.D. Entian, J.D. Hoheisel (1998) "Transcriptional profiling on all open reading frames of *Saccharomyces cerevisiae*" Yeast **14**(13):1209-1221

Kastaniotis A.J. & R.S. Zitomer (2000) "Rox1 mediated repression. Oxygen dependent repression in yeast" Advances in Experimental Medicine and Biology 475:185-195

Klinkenberg L.G., T.A. Mennella, K. Luetkenhaus, R.S. Zitomer (2005) "Combinatorial repression of the hypoxic genes of *Saccharomyces cerevisiae* by DNA binding proteins Rox1 and Mot3" Eukaryotic Cell **4**(4):649-660

Kwast K.E., P.V. Burke, B.T. Staahl, R.O. Poyton (1999) "Oxygen sensing in yeast: evidence for the involvement of the respiratory chain in regulating the transcription of a subset of hypoxic genes" Proceedings of the National Academy of Sciences USA **96**(10):5446-5451

Kwast K.E., L.C. Lai, N. Menda, D.T. James, S. Aref, P.V. Burke (2002) "Genomic analyses of anaerobically induced genes in *Saccharomyces cerevisiae*: functional roles of Rox1 and other factors in mediating the anoxic response" Journal of Bacteriology **184**(1):250-265

Labbe-Bois R. & P. Labbe (1990) "Tetrapyrrole and heme biosynthesis in the yeast *Saccharomyces cerevisiae*. In: Dailey HA (ed) Biosynthesis of heme and chlorophylls" McGraw-Hill, New York, pp 235-285

Lambert J.R., V.W. Bilanchone, M.G. Cumsky (1994) "The *ORD1* gene encodes a transcription factor involved in oxygen regulation and is identical to *IXR1*, a gene that confers cisplatin sensitivity to *Saccharomyces cerevisiae*" Proceedings of the National Academy of Sciences USA **91**(15):7345-7349

Luikenhuis S., G. Perrone, I.W. Dawes, C.M. Grant (1998) "The yeast *Saccharomyces cerevisiae* contains two glutaredoxin genes that are required for protection against reactive oxygen species" Molecular Biology of the Cell **9**(5):1081-1091

McA'Nulty M.M., J.P. Whitehead, S.J. Lippard (1996) "Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts *in vitro* and *in vivo*" Biochemistry **35**(19):6089-6099

Regnacq M., P. Alimardani, B. El Moudni, T. Berges (2001) "Sut1p interaction with Cyc8p (Ssn6p) relieves hypoxic genes from Cyc8p-Tup1p repression in *Saccharomyces cerevisiae*" Molecular Microbiology **40**(5):1085-1096

Robinson MD, J. Grigull, N. Mohammad, T.R. Hughes (2002) "FunSpec: a web-based cluster interpreter for yeast" BMC Bioinformatics 3:e35

Smits G.J., J.C. Kapteyn, E.H. van Den, F.M. Klis (1999) "Cell wall dynamics in yeast" Current Opinion in Microbiology **2**(4):348-352

Tsaponina O., E. Barsoum, S.U. Astrom, A. Chabes (2011) "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools" PLoS Genetics 7(5):e1002061

Ter Linde J.J. & H.Y. Steensma (2002) "A microarray-assisted screen for potential Hap1 and Rox1 target genes in *Saccharomyces cerevisiae*" Yeast **19**(10):825-840

Thomas D., R. Barbey, Y. Surdin-Kerjan (1990) "Gene-enzyme relationship in the sulphate assimilation pathway of *Saccharomyces cerevisiae*. Study of the 3'-phosphoadenylylsulphate reductase structural gene" The Journal Biology Chemistry 265:15518-15524

Thomas D., R. Barbey, D. Henry, Y. Surdin-Kerjan (1992) "Physiological analysis of mutants of *Saccharomyces cerevisiae* impaired in sulphate assimilation" Journal of General Microbiology **138**(10):2021-2028

Thomas D. & Y. Surdin-Kerjan (1997) "Metabolism of sulfur amino acids in *Saccharomyces cerevisiae*" Microbiology and Molecular Biology **61**(4):503-532

Vik A. & J. Rine (2001) "Upc2p and Ecm22p, dual regulators of sterol biosynthesis in *Saccharomyces cerevisiae*" Molecular Cell Biology **21**(19):6395-6405

Zitomer R.S., C.V. Lowry (1992) "Regulation of gene expression by oxygen in *Saccharomyces cerevisiae*" Microbiology Reviews **56**(1):1-11

Zitomer R.S., P. Carrico, J. Deckert (1997) "Regulation of hypoxic gene expression in yeast" Kidney International **51**(2):507-513

Chapter 2

Chromatin immunoprecipitation and *in silico* studies looking for a consensus of Ixr1 binding

SUMMARY

In *Saccharomyces cerevisiae*, adaptation to low oxygen levels leads to deep changes in cell metabolism. The implications of Ixr1 in the regulation of genes related to the hypoxic and oxidative stress responses, and also to the re-adaptation of catabolic and anabolic fluxes in response to oxygen limitation, have been previously described. Although the promoters of several genes (*HEM13, COX5B, TIR1, IXR1* or *ROX1*) have been experimentally studied *in vivo* and/or *in vitro*, a general consensus for Ixr1 binding in a DNA-sequence specific way has not been well established yet. In the present study, ChIP-on-chip analysis of Ixr1 DNA-binding sites both during normoxia and hypoxia revealed low enrichment in promoter regions, ranging from 50% to 60% of the total significant binding sites. Regulation of transcription by Ixr1 is mainly attributable to indirect or transient mechanisms, instead of stable binding to regulated promoters, as deduced from the low overlap obtained between transcriptome and ChIP-on-chip data. Besides, ChIP-on-chip analyses of Rox1 DNA-binding under normoxia were done, and its participation in a cross-regulation between Rox1 and Ixr1 is also discussed.

1.- INTRODUCTION

The first report about the participation of Ixr1 in the yeast hypoxic response was done more than twenty years ago, when Lambert and col. published that Ixr1 causes aerobic repression of the *COX5B* gene, which encodes the hypoxic isoform of the subunit Vb of the mitochondrial complex cytochrome *c* oxidase (Lambert *et al*., 1994). Along these years this transcriptional factor has also been related to other hypoxic genes like *TIR1,* a cell wall mannoprotein of the serine-alanine-rich protein familly (Bourdineaud *et al*., 2000) and *HEM13,* which encodes the enzyme coproporphyrinogen III oxidase in the heme biosynthetic pathway (Castro-Prego *et al*., 2010b). More recently, a genome wide study has allowed the

identification of the whole set of genes that are regulated by changes in mRNA levels during the hypoxic response by an Ixr1-dependent mechanism (Vizoso-Vázquez *et al*., 2012). However, very little is known about the interaction of Ixr1 with the promoters of the target genes.

The analysis of the amino acid sequence of Ixr1 reveals two HMG-box domains, which have been related to interaction with DNA in other transcriptional regulators (Jantzen *et al*., 1990; Oosterwegel *et al*., 1991; Waterman *et al*., 1991). Although the folding of HMG-boxes is well conserved in the HMGB proteins (Thomas & Travers 2001), these domains are functionally heterogeneous; that is, depending on the amino acids present at specific positions they could bind DNA in a way that requires the recognition of a specific sequence (SS) in the nucleic acid, a tractable consensus; otherwise they could bind DNA without recognizing a specific nucleotide sequence (NSS) but a specific DNA structure (Bruhn *et al*., 1992; Landsman & Bustin, 1993; Grosschedl *et al*., 1994). Modelling of the two HMG-boxes from Ixr1 and comparison with experimentally determined structures of other HMG-boxes, previously classified as DNA binding domains with DNA-sequence specificity or with DNA-structure specificity, predicts that binding of Ixr1 to DNA might be structure-specific through the first HMG-box and sequence-specific through the second (Castro-Prego *et al*., 2010a).

The existence in *S. cerevisiae* of a general consensus for Ixr1 binding in a DNA-sequence specific way has not been well established yet. Although the promoters of five genes have been experimentally studied *in vivo* and/or *in vitro,* it is still a matter of controversy. Binding of Ixr1 to *COX5B* promoter was studied (Lambert *et al*., 1994) by electrophoretic mobility-shift assays (EMSA) and using a DNA probe of 44-bp (-239 to -195), which mediated the aerobic repression of *COX5B* transcription (Hodge *et al*., 1990). This DNA probe contained in the direct strand the 8 bp hypoxic operator **TATTGTTC**, which is found upstream of many hypoxic genes and includes the core for Rox1 binding (Lowry *et al*., 1990; Zitomer &

Lowry 1992); in the reverse strand also carried a 13-bp sequence **TCGTTCGTTGCCT**, which is found upstream of several hypoxic genes (Hodge *et al*., 1990; Lowry *et al*., 1990). Two sequences in the *TIR1* promoter (-299 to -251 and -218 to -156) allowed constitutive normoxic expression of the gene and bound Ixr1 in EMSA (Bourdineaud *et al*., 2000); it was shown that Ixr1 could bind to probes deleted for either of the two boxes, but could not bind when both boxes were removed (Bourdineaud *et al*., 2000). Ixr1 is necessary for the high expression of *HEM13* under hypoxic conditions and its function is exerted *in vivo* through the *HEM13* promoter region extending from -577 to -419 (Castro *et al*., 2010a). Chromatin immune-precipitation (ChIP) analyses showed that Ixr1 binds *in vivo* to the *HEM13* promoter both under aerobic and under hypoxic conditions (Castro-Prego *et al*., 2010a). *In vitro* studies by EMSA demonstrated that Ixr1 binds to two sequences extending from -534 to -509 and from -497 to -450, which competed between them and from which the consensus **KTTSA<u>AYKGTT</u>YASA** was deduced (Castro *et al*., 2010a). The sequence underlined in the consensus represents a degenerate form of the core in the consensus for Rox1 binding (**ATTGTT**). Indeed the actual sequence for Rox1 binding to the *HEM13* promoter in aerobic conditions (**TTTCAATTGTTTAGA**) extends from positions -476 to -462 and was therefore included in the alignment to deduce the consensus (Castro-Prego *et al*., 2010a). Ixr1 also binds to the region -300 to -102 of the YOL104w promoter, which contains a sequence (from -175 to -161) matching the deduced consensus **KTTSA<u>AYKGTT</u>YASA** (Castro-Prego *et al*., 2010a). Using a general wide-genome approach to identify specific sequences for binding of all *S. cerevisiae* transcription factors the consensus **AArcmrgRAGCGGkG** was deduced for Ixr1 binding (MacIsaac *et al*., 2006). Curiously Ixr1 auto-activates the expression of the *IXR1* gene during hypoxia by direct binding to the promoter region extending from −557 to −376, which contains two sequences that resemble the consensus proposed by MacIsaac for Ixr1 binding (Castro-Prego *et al*., 2010b).

As reported, some data argue in favour of a possible competition or alternative use of Rox1 and Ixr1 for binding to the same consensus in promoters co-regulated by these two transcriptional factors that contain HMG-boxes for DNA binding. However the main problem to sustain or refute this hypothesis is that the data available do not allow the identification of one unique and highly specific DNA consensus for Ixr1 binding. The scarce number of promoters studied so far and the intrinsic difficulty of characterizing the nature of the binding of Ixr1 to the DNA, since it has two HMG-boxes with structure-specificity and sequence-specificity respectively, make it necessary a whole-genome approach. In this work we present the data of ChIP-on-chip analyses carried out with *S. cerevisiae* strains in which the Ixr1 and Rox1 proteins had been tagged for immunoprecipitation. The binding of the two regulators was monitored in normoxia. It was also studied the binding of Ixr1 in normoxia in a strain in which the gene *ROX1* was deleted. The binding of Ixr1 was also studied in hypoxia, but Rox1 binding was not assayed in hypoxia because the protein is quickly degraded when oxygen levels decay. Results are discussed in relation to previously published transcriptome data.

2.- MATERIAL AND METHODS

2.1.- Cell culture and strains

The *S. cerevisiae* strains Z1465 (*MAT**a** ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ psi+ ROX1:myc9::TRP1*)   and Z1580 (*MAT**a** ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ psi+ IXR1:myc9::TRP1*) where obtained from Young's lab (Lee *et al*., 2002). The knockout strain Z1580-*Δrox1* was obtained by one-step replacement with the *URA3* marker.

The plasmid Yeplac195 (Gietz & Sugino, 1998) was used as a template to amplify a linear fragment containing the *URA3* gene and two flanking regions of homology to the ORF ends of *ROX1* by PCR using the primers ECV688 and ECV689

(Table 1). After transformation of the Z1580 strain with amplified fragments, cells were selected in complete media without uracil. The correct replacement in the *S. cerevisiae* genome was verified by PCR as described previously (Tizón *et al*., 1999) using primers designed inside the URA3 ORF and the flanking regions of *ROX1* (see Table 1 oligos ECV698AV, ECV699AV, ECV700AV, ECV701AV). The handling of yeast cells was carried out according to standard procedures.

Table 1. Oligos used to obtain and check Z1580-*Δrox1 strain.*

Oligo name	Sequence	Gene	Strand[a]	Position[b]
ECV688	atgaatcctaaatcctctacacctaagattCCTTTAGCTGTTCTATATGCTGC	*ROX1*	C	+1
ECV689	tcatttcggagaaactaggctagttttagcCCACCTGACGTCTAAGAAACC	*ROX1*	W	+1107
ECV698AV	GTGATCTTCGGCTCGGC	*ROX1*	W	-554
ECV699AV	AAGAGATGAAGGTTACGATTGGT	*URA3*	C	+592
ECV700AV	TTGTACTTGGCGGATAATGC	*URA3*	W	+234
ECV701AV	ATATCTTGCAGTCCATCCTCG	*ROX1*	C	+1537

[a]W: Watson strand; C: Crick strand.
[b]Numbering is considering +1 for the adenine in the first start codon.

Three biological replicates of cultures and treatments were run. Yeast cells were pre-cultured overnight in 10 mL of complete synthetic medium (SD) prepared as previously described (Zitomer & Hall, 1976). For ChIP-on-chip experiments the cultures were inoculated at initial OD_{600} of 0.1 in 200 mL SD (6.7 g of bacto-yeast nitrogen base without amino acids from Difco (Franklin Lakes, New Jersey, USA); 40 mg each of histidine, leucine, adenine, uracil and tyrosine, 10 mg of methionine and 30 mg tryptophan; 2% glucose (w/v)) and grown in 1 L Erlenmeyer flasks at 30 ^{o}C and with agitation at 250 rpm in the normoxic or hypoxic conditions. During hypoxic growth, cells were cultured overnight in anaerobic jars with the GasPack EZ-Anaerobe system from Becton, Dickinson and Company (Franklin Lakes, New Jersey, USA) and under these conditions (oxygen concentration <1%) the medium was supplemented with 20 mg/L ergosterol and 0.5% tween 80.

2.2.- Chromatin immunoprecipitation and ChIP-on-chip analysis

Chromatin immunoprecipitation experiments were carried out as described previously (Lee *et al.*, 2002) with minor modifications. Shortly, 200 mL of yeast culture were collected at OD_{600} of ≈ 0.9-1. Crosslinking was performed by adding 1% formaldehyde to the culture and incubating at room temperature for 20 min. 125 mM of glycine was then added and culture was incubated 5 min. Cells were then harvested and washed four times with 50 ml Tris-HCl buffer saline (20 mM Tris-HCl, pH 7,5, 150 mM NaCl) at 4 ^{o}C. The cell breakage was performed in 800 µL of lysis buffer (50 mM HEPES-KOH, pH 7,5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0,1% sodium deoxycholate, 2X complete protease inhibitor cocktail, Roche, and 2X complete phosphatase inhibitor cocktail, Roche) with glass beads; cell extracts were sonicated for 5 min in 10 sec on/59 sec off cycles (chromatin was sheared into an average size of 400 bp). Immunoprecipitations were performed with magnetic Dynabeads (Invitrogen) following the manufacturer's instructions and using anti-(c-Myc) antibodes (sc47694; Santa Cruz Biotechnology) for specific Ixr1–(c-Myc) immunoprecipitation. Negative controls, with rabbit IgG immunoprecipitation, were also performed. Samples were washed three times with 1 mL of Lysis Buffer, three times with 1 mL of lysis buffer high salt (50 mM HEPES-KOH, pH 7,5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0,1% sodium deoxycholate), three times with 1 mL of wash buffer (10 mM Tris-HCl, pH 8, 250 mM LiCl, 0,5% NP-40, 0,5% sodium deoxycholate, 1 mM EDTA), and once with 1 mL of TE Buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA). Immunoprecipitations were then eluted in 250 µL of elution buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS) and treated overnight with 30 µL proteinase K (20 mg/mL, NewEngland Biolabs). Next day, immunoprecipitated DNAs were cleaned with the kit USB PrepEase DNA Clean-Up (USB). Next steps were rigorously followed according to the manufacturer's instructions (Affymetrix; http://cmgm.stanford.edu/pan/section_html/GE/protocols/Chromatin%20Immun

oprecipitation%20Assay%20Protocol.pdf). Immunoprecipitation enrichment was checked by real-time qPCR against promoter regions of *TIR1*, *IXR1*, *ROX1* and *HEM13* (Table 2), which are known to bind the Ixr1 protein and immunoprecipitated using the specific antibody (Bordineaud *et al.*, 2000; Castro-Prego *et al.*, 2010a; Castro-Prego *et al.*, 2010b). At least, >8-fold enrichment was obtained for IP samples compared to the IgG samples.

Eighteen GeneChip® *S.cerevisiae* Tiling 1.0R arrays from Affymetrix Inc. (Wycombe. United Kingdom) were used and processed in the GeneChip® System with Autoloader from Affymetrix Inc. (Wycombe, United Kingdom). Control Oligo B2 was included to provide alignment signals for image analysis. Image caption and preliminary data analysis were carried out with Affymetrix® Expression Console™ software (v1.1).

Table 2. Oligos used during ChIP-on-chip sample preparations.

Oligo name	Sequence	Gene	Strand[a]	Position[b]
AVV220	AGAACTTGGCGATTGCTGACA	*ROX1*	C	-408
AVV221	AAGACCGTTACATTACGCAAAGTG	*ROX1*	W	-275
AVV222	CATACACATCGTGCTTAGCGATC	*IXR1*	W	-526
AVV223	CCCATTCGTTCTCTCACCAAG	*IXR1*	C	-376
AVV224	CATAAAGGGTCTCTTTCACCTATACG	*TIR1*	W	-273
AVV225	CTTCACTTTTTTCTCTGTCAAGGG	*TIR1*	C	-178
AVV226	TCAAACCATTTCCTGCGGAG	*HEM13*	C	-539
AVV227	TGCCTATGACGGTAATCCCA	*HEM13*	W	-406
Primer_A	GTTTCCCAGTCACGGTCNNNNNNNNN	-	-	-
Primer_B	GTTTCCCAGTCACGGTC	-	-	-

[a]W: Watson strand; C: Crick strand.
[b]Numbering is considering +1 for the adenine in the first start codon.

2.3.- Statistical data analysis and data mining

ChIP-on-chip raw data data from Affymetrix GCOS software were analyzed using Affymetrix Tiling Analysis Software (TAS) v1.1.03 (http://www.affymetrix.com/support/developer/downloads/TilingArrayTools/inde

x.affx), and the .BPMAP file *Sc03b_MR_v04.bpmap*. A two-sample analysis was conducted using specific chromatin immunoprecipitation (from Ixr1-(c-Myc) or Rox1-(c-Myc) tagged samples) DNA samples as the 'treatment' group and three whole genome fragmented and amplified DNA samples as the 'control' group. Data were normalized using built-in quantile normalization and probe-level analysis with perfect match (PM) probes and run with a bandwidth of 250. Ixr1 and Rox1 protein occupancy profiles were visualized with Affymetrix Integrated Genome Browser (IGB). Interval analyses were done using TAS software with a minimum run of 10 and maximum gap of 250 bp, and *p-value* cutoff of 0.01. Bed file conversions were done using UCSC (*University of California Santa Cruz*) tools (https://genome.ucsc.edu). Bed file analyses were done using ChIpSeek tools (http://chipseek.cgu.edu.tw) (Chen *et al*., 2014).

Gene descriptions and comparative analyses of lists from differentially expressed genes were obtained through Yeast Mine (http://yeastmine.yeastgenome.org/yeastmine/begin.do). Functional distribution of genes in the differentially regulated clusters was also analyzed using FunSpec (http://funspec.ccbr.utoronto.ca/) developed by Robinson and co-workers (Robinson *et al*., 2002). The MIPS Functional Catalogue Database (FunCatDB) was used in the analyses (http://mips.helmholtz-muenchen.de/proj/funcatDB/). Motif analysis was performed using RSAT (*Regulatory Sequence Analysis Tools*) (http://rsat.ccb.sickkids.ca/) (Bailey *et al*., 1994; van Helden, 2003), YEASTRACT (*Yeast Search for Transcriptional Regulators And Consensus Tracking*) (http://www.yeastract.com) (Teixeira *et al*., 2014), the WebMOTIFS suite (http://fraenkel.mit.edu/webmotifs.html) (Romer et al., 2007) and MEME (*Multiple Em for Motif Elicitation*) suite (http://meme-suite.org) (Bailey and Elkan, 1994).

3. RESULTS AND DISCUSSION

3.1.- *In silico* anaylisis of regulatory motifs in the promoters of the genes, which are up-regulated in a *Δixr1* strain during normoxia.

It has been previously reported that Ixr1 is a repressor of hypoxic genes during normoxia (Lambert *et al*. 1994; Bourdineaud *et al*. 2000) and activator during hypoxia (Castro-Prego *et al*. 2010a; Castro *et al*., 2010b). The molecular mechanisms that support this regulatory control are however unknown. The action of Ixr1 upon the hypoxic gene *TIR1* was initially described to be Rox1 independent (Bourdineaud *et al*., 2000), however other data suggest a cross-regulation between Rox1 and Ixr1 (Castro *et al*., 2010b) that is strain dependent and it is not observed in a W303 genetic background (Liu & Barrientos, 2013). Besides, the gene *HEM13* is regulated both by Rox1 and Ixr1 binding to the same *cis*-regulatory sequence and, considering that *IXR1* expression increases during hypoxia (Castro-Prego *et al*., 2010b), while the available Rox1 protein is degraded in these conditions (Deckert *et al*., 1995), it has been suggested that Ixr1 replaces Rox1 in the change normoxia-hypoxia in the binding to this *cis*-regulatory element (Castro-Prego *et al*., 2010a). We have scanned the promoter sequences (-1000 to -1) of the genes up-regulated in normoxia when the Ixr1 gene is knocked-down (those described in Vizoso *et al*., 2012) looking for the already reported motifs of Rox1 and Ixr1 binding. The PSSM matrix from the JASPAR data base (http://jaspardev.genereg.net/) and the program "matrix scan full" (Turatsinze *et al*., 2008) from the RSA tools suit (http://www.rsat.eu/) were used as described in materials and methods. 79 genes among those analysed contain significant Rox1 binding sites and 54 significant Ixr1 binding sites (Figure 1a). Only 16 genes contained both sequences (outlined in red in the Figure 1a and 1b) and their relative positions in reference to the A base from the ATG codon is also shown in (Fig 1c). A positional association between the Rox1 and Ixr1 binding sites is not observed (Fig 1c). A specific functional enrichment in

the subgroups having only Rox1 or Ixr1 predicted sites was neither observed using YeastMine (http://yeastmine.yeastgenome.org/) or Funspec (http://funspec. med.utoronto.ca/) programs (data not shown).

We also used available bioinformatics methods to discover new motifs for Ixr1 binding that could better explain the regulatory control exerted by Ixr1 during normoxia. First, we analysed the promoters with the WebMOTIFS suite (Romer *et al*., 2007), which performs individual searches with the programs AlignACE (Hughes *et al*., 2000), MDscan (Liu *et al*., 2002), MEME (Bailey and Elkan, 1994) and Weeder (Pavesi *et al*., 2006) and then coordinately evaluates the whole set of results and their statistical significance. Using high-stringent criteria as defined in the suit, only one motif was selected. The found sequence, ***AAGGGG***, showed a 6.97 enrichment (enrichment score of over-representation of motif) and median enrichment Z score of 4.98 (median of the enrichment Z scores for all the motifs in this cluster, considering motifs from different programs together). Using the program oligo-analysis (van Helden *et al*., 1998) from RSA tools (http://www.rsat.eu/) a very similar motif, ***AAGGGGC***, was found. 105 genes contain one or more sequences fitting this motif in their promoters (Figure 2).

This motif was compared to other regulatory motifs in the databases, using the TOMTOM motif comparison tool (Gupta *et al*., 2007) through MEME (http://meme.nbcr.net) and the results are summarized in figure 3. The most significant similarity is observed with the already known motifs binding the Msn2 and Msn4 transcriptional factors, which are related to the general stress response, although they have also been associated to the short-term response to oxygen deprivation (Lai *et al*., 2005).

(a)
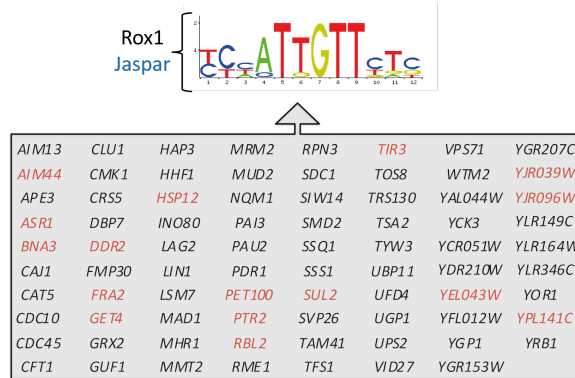
(b)
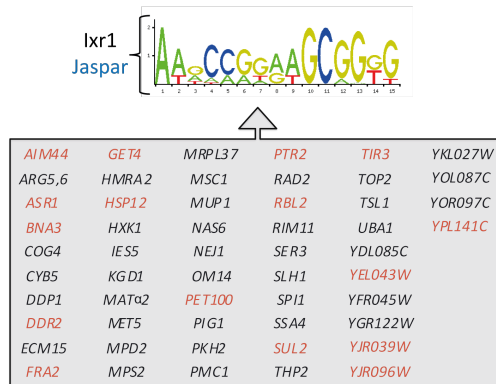
(c)

Figure 1. Genes up-regulated during normoxia after Ixr1 deletion and showing binding sites for Rox1 (a) or Ixr1 binding (b). Those genes having both regulatory sequences are in red. The relative position of Rox1 (left) and Ixr1 binding sites (right) is shown in (c).

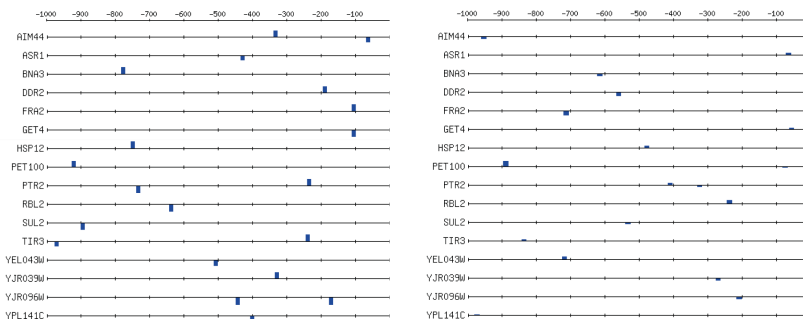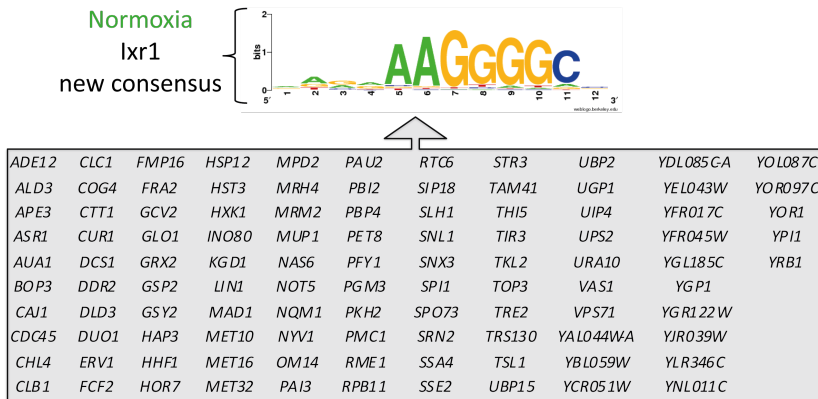|										
ADE12	CLC1	FMP16	HSP12	MPD2	PAU2	RTC6	STR3	UBP2	YDL085C-A	YOL087C
ALD3	COG4	FRA2	HST3	MRH4	PBI2	SIP18	TAM41	UGP1	YEL043W	YOR097C
APE3	CTT1	GCV2	HXK1	MRM2	PBP4	SLH1	THI5	UIP4	YFR017C	YOR1
ASR1	CUR1	GLO1	INO80	MUP1	PET8	SNL1	TIR3	UPS2	YFR045W	YPI1
AUA1	DCS1	GRX2	KGD1	NAS6	PFY1	SNX3	TKL2	URA10	YGL185C	YRB1
BOP3	DDR2	GSP2	LIN1	NOT5	PGM3	SPI1	TOP3	VAS1	YGP1	
CAJ1	DLD3	GSY2	MAD1	NQM1	PKH2	SPO73	TRE2	VPS71	YGR122W	
CDC45	DUO1	HAP3	MET10	NYV1	PMC1	SRN2	TRS130	YAL044W-A	YJR039W	
CHL4	ERV1	HHF1	MET16	OM14	RME1	SSA4	TSL1	YBL059W	YLR346C	
CLB1	FCF2	HOR7	MET32	PAI3	RPB11	SSE2	UBP15	YCR051W	YNL011C	

Figure 2. Genes up-regulated during normoxia after Ixr1 deletion and showing the new discovered consensus for Ixr1 binding.

3.2.- *In silico* analysis of regulatory motifs in the promoters of the genes, which are down-regulated in a *Δixr1* strain during hypoxia

We have scanned the promoter sequences (-1000 to -1) of the genes down-regulated in hypoxia when *IXR1* gene is knocked-down (those described in Vizoso *et al.*, 2012) looking for the already reported motifs of Rox1 and Ixr1 binding using the PSSM matrixes from the JASPAR database and the program "matrix scan full" from the RSA tools suit as already described for the analysis of the promoters up-regulated in normoxic conditions.

In hypoxic conditions 48 down regulated genes after Ixr1 deletion contain significant Rox1 binding sites and 54 significant Ixr1 binding sites (Figure 1). Only 14 genes contained both sequences (highlighted in red in figure 4). A positional association, between the Rox1 and Ixr1 binding sites in these 14 promoters having both consensus sequences, was not observed (data not shown).

The sequence **AAGGGGC** deduced analysing the genes up-regulated after Ixr1 deletion in normoxia is also found in the genes down-regulated after Ixr1 deletion in hypoxia in 61 of their promoters (Figure 4).

106

	p-value	*E-value*	*Q-value*	Overlap	Orientation	Consensus sequence alignment
MSN2	1.98033×10^{-5}	0.00293088	0.00583151	7	D	
MSN4	0.000194008	0.0287133	0.0285651	7	D	
RPH1	0.0016483	0.243949	0.161793	6	RC	
NDD1	0.0111524	1.65056	0.803487	7	D	
NRG1	0.0136428	2.01914	0.803487	6	RC	

Figure 3. Overlap of the consensus *AAGGGGC* with other consensus for known transcriptional regulators.

(a)

AIM13	GCY1	PAI3	SPS4	YJL017W
AME1	HSP12	PEX19	SUL1	YJR096W
ANT1	INH1	RPL29	TFB4	YKT6
ARC18	IPL1	RRP43	UTP11	YMR226C
ATP3	JLP1	RTN2	VPS51	YOR246C
CDC28	LIP1	SAM2	YDL144C	YOR338W
CDC8	MAD2	SCD6	YFL012W	YRB1
CTP1	MRPL49	SFG1	YGR153W	YUH1
CWC2	NDJ1	SGT2	YHP1	
ERD1	NTG2	SPO12	YHR138C	

(b)

ADI1	CTP1	HSP12	RAD2	UAF30	YNL155W
AIM5	CYS4	HYR1	RPL25	UBC9	YOL107W
ANT1	DDP1	KIN28	RPL29	UFE1	YOR097C
AQY2	DID4	LDB18	SCD6	VPS20	YOR246C
ARG82	ERD1	MATα2	SCP1	YFR045W	
ATP3	EST3	MET5	SGT2	YGR201C	
AYT1	GIM5	MGE1	SLD2	YHR138C	
CAP1	GPX2	MRP2	SPC110	YJR096W	
CDC28	GRX1	MRPL35	SPO12	YLL053C	
CRC1	GTT1	MRPL37	SPS100	YMR226C	

(c)

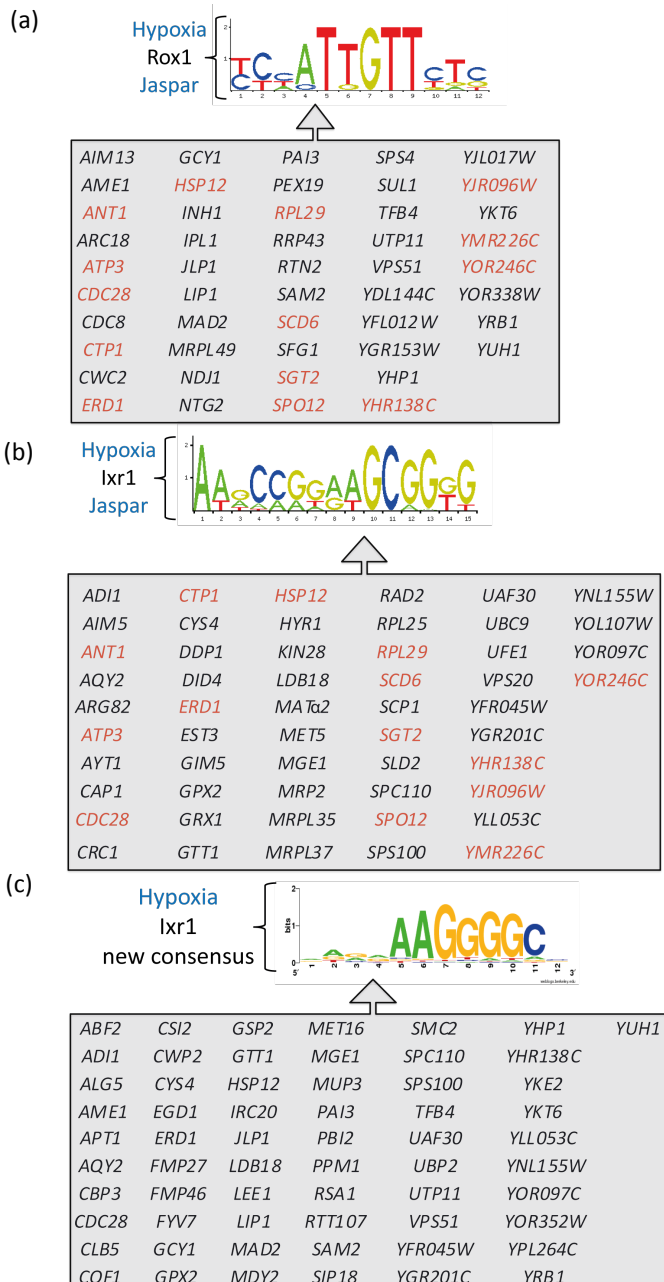| ABF2	CSI2	GSP2	MET16	SMC2	YHP1	YUH1
ADI1	CWP2	GTT1	MGE1	SPC110	YHR138C	
ALG5	CYS4	HSP12	MUP3	SPS100	YKE2	
AME1	EGD1	IRC20	PAI3	TFB4	YKT6	
APT1	ERD1	JLP1	PBI2	UAF30	YLL053C	
AQY2	FMP27	LDB18	PPM1	UBP2	YNL155W	
CBP3	FMP46	LEE1	RSA1	UTP11	YOR097C	
CDC28	FYV7	LIP1	RTT107	VPS51	YOR352W	
CLB5	GCY1	MAD2	SAM2	YFR045W	YPL264C	
COF1	GPX2	MDY2	SIP18	YGR201C	YRB1	

Figure 4. Genes down-regulated during hypoxia after Ixr1 deletion and showing the Rox1 (a), Ixr1 (b) or the new discovered consensus for Ixr1 binding (c). Those genes having both regulatory sequences as in JASPAR are in red.

108

3.3.- The distribution of consensus sequences for Rox1 and Ixr1 binding in the pools of up-regulated genes during normoxia and down-regulated genes during hypoxia

We have investigated the relative distribution of the consensus sequences for Rox1 and Ixr1 binding in the pools of up-regulated genes during normoxia by Ixr1 deletion (Figure 5a) and down-regulated genes during hypoxia by Ixr1 deletion (Figure 5a). The intersections among the different sub-groups have been calculated through YeastMine (Balakrishnan *et al.*, 2012) and Venn diagrams summarize the results. Among the genes up-regulated in normoxia, the genes that have Ixr1, Rox1 or the new discovered (***AAGGGC***) motif in their promoters form subgroups with some overlap, but clearly differed and there are 60 genes up-regulated in the mutant during normoxia, which are out of these three subgroups (Figure 5a). Among the genes down-regulated during hypoxia, the genes that have Ixr1, Rox1 or the new discovered (***AAGGGC***) motif in their promoters also form subgroups with some overlap, but clearly differed and there are 51 genes down- regulated in the mutant during hypoxia, which are out of these three subgroups (Figure 5b).

These data indicate that the effects observed, after Ixr1 deletion during normoxia or hypoxia, are not associated to a unique and common regulatory sequence in the promoter of the regulated genes. At least three subgroups are differentiated and related to the presence of different consensus sequences, and a fourth subgroup is not associated to these consensuses.

However, the consensus attributed to Ixr1 recognition (http://jaspar.binf.ku.dk/cgibin/jaspar_db.pl?rm=browse&db=core&tax_group=fungi; JASPAR accession MA0323.1) and the new discovered motif in our analysis using the WebMOTIFS suite (Romer *et al.*, 2007) are not independent, and share a core formed by the common sequence ***AAG[G/C]GG***, which is also present in the

regulatory sequences for binding of Msn2 and Msn4 (Figure 3). Thus, considering all the genes, which present in their promoters this re-defined core (Figure 6) a high percent (58 % in normoxia and 58 % in hypoxia) of regulated genes may be considered as grouped in a common regulon.
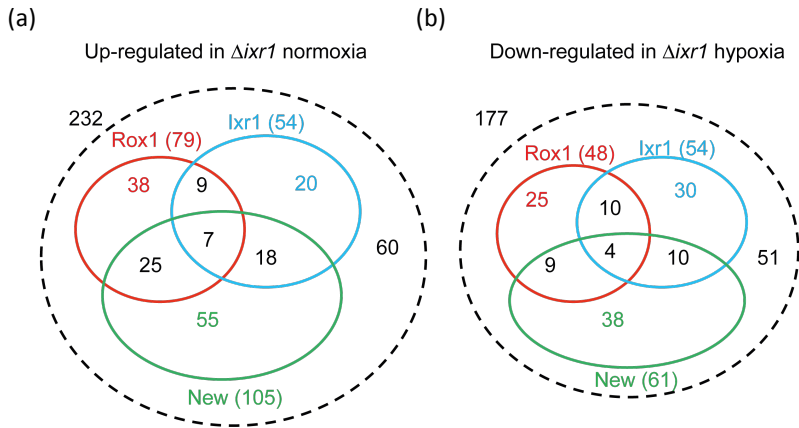

Figure 5. Venn diagrams showing the distribution of consensus sequences for Rox1 and Ixr1 binding in the pools of up-regulated genes during normoxia (a) and down-regulated genes during hypoxia (b). Matrix for Rox1 (JASPAR accession MA0371.1) and Ixr1 (JASPAR accession MA0323.1) as defined in JASPAR (http://jaspar. binf.ku.dk/cgi-bin/jaspar_db.pl?rm=browse&db=core&tax_group=fungi). New matrix for **AAGGGGC** calculated as defined in the text.

Msn2 and Msn4 are paralogs; both proteins share 41% identity and are similar in size and amino acid composition (Estruch & Carlson, 1993). Although physical interactions of these factors with Ixr1 have not been described, a negative genetic interaction between Msn4 and Ixr1 has been reported in two high-throughput independent analyses (Bandyopadhyay *et al*., 2010; Zheng *et al*., 2010). Msn2 and Msn4 have been associated to the rapamycin-sensitive TOR signaling pathway (Beck & Hall, 1999) like Ixr1 (Chen *et al*., 2013). This signalling pathway in *Saccharomyces cerevisiae* activates responses for metabolic adaptations required in response to nitrogen and carbon variations. It has been proposed that the TOR signalling pathway controls nutrient metabolism by sequestering several

transcription factors in the cytoplasm; thus, TOR inhibits expression of carbon-source-regulated genes by stimulating the binding of the transcriptional activators Msn2 and Msn4 to the cytoplasmic 14-3-3 protein Bmh2 (Beck & Hall, 1999). Upon several stress conditions, Msn2 and Msn4 are re-localized to the nucleus (Gorner *et al*., 1998).
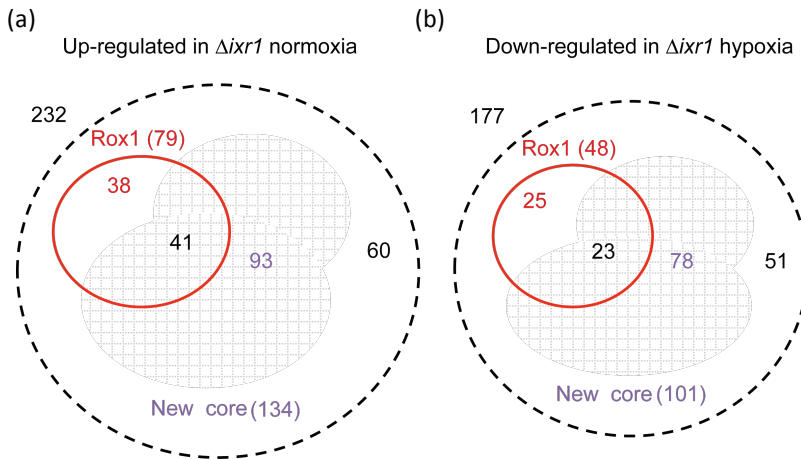

Figure 6. Venn diagrams showing the distribution of consensus sequences for Rox1 and Ixr1 binding in the pools of up-regulated genes during normoxia (a) and down-regulated genes during hypoxia (b). Matrix for Rox1 as defined in JASPAR (accession MA0371.1). The redefined new core for Ixr1 is defined by the string ***AAG[G/C]GG***.

Considering that the genes down-regulated in hypoxia in the *Δixr1* strain versus the isogenic Ixr1 strain, could be responding to hypoxia as a general stress response, we compared them with the known targets of Msn2 and Msn4 from the information in SGD. These targets in SGD are recompiled majorly from chromatin immunoprecipitation, and DNA arrays reported in the literature and databases. The intersection with a pool of 383 Msn2 targets was only 12 genes and with a pool of 53 Msn4 targets only 1 gene. Therefore the regulation observed after Ixr1 deletion during hypoxia is probably not a direct consequence of changes in

expression mediated by direct interaction with Msn2 or Msn4 with their known targets.

3.4.- Determination of Ixr1 binding during normoxia by ChIP-on-chip analysis and its association to transcriptional regulation mediated by Ixr1

The analysis was done according to the description in section materials and methods. The summary of peaks distribution and length is shown in Figure 7, rendering 1004 peaks (*p-value* < 0.01), 519 related to promoter regions, that represent 51.7% of the total peaks obtained.

The list of selected peaks is summarized in supplementary Table S1. Analysis of these selected peaks with YEASTMINE reveals the enrichment of two functional groups of genes; those related to cell-wall organization and to oxidation-reduction processes. *HEM13*, a normoxic target for Ixr1 regulation already characterized in previous analyses (Castro-Prego *et al*., 2010a), is included in this list.

The intersection between the list of genes regulated by Ixr1 in normoxia (Suppletory table S1 of chapter 1) and these found through immunoprecipitation mediated by Ixr1 in normoxia suggests that only a 11.2% of the observed transcriptional regulation could be attributed to direct and stable binding of Ixr1 to the target promoters (Table S2). Among the promoters of genes activated by Ixr1 and showing direct binding of Ixr1 are *ILV3*, *ILV5* and *ILV6*, which encode enzymes necessary for the biosynthesis of branched-chain amino acids like isoleucine and valine. Also *TDH2*, *TDH3* and *HPF1* from sugar metabolism and *FIT3* from ion transport. As previously explained (Chapter 1), the genes from the sulphate assimilation pathway are up-regulated in the mutant *Δixr1* during normoxic growth (repressed by Ixr1 protein) and down-regulated in the mutant during hypoxic growth (activated by Ixr1 protein), but with the difference that in this case no

112

direct and stable binding of Ixr1 to the promoter region of these genes is observed. The only gene promoter bound by Ixr1 and related to sulphur metabolism is the *MET28* promoter.
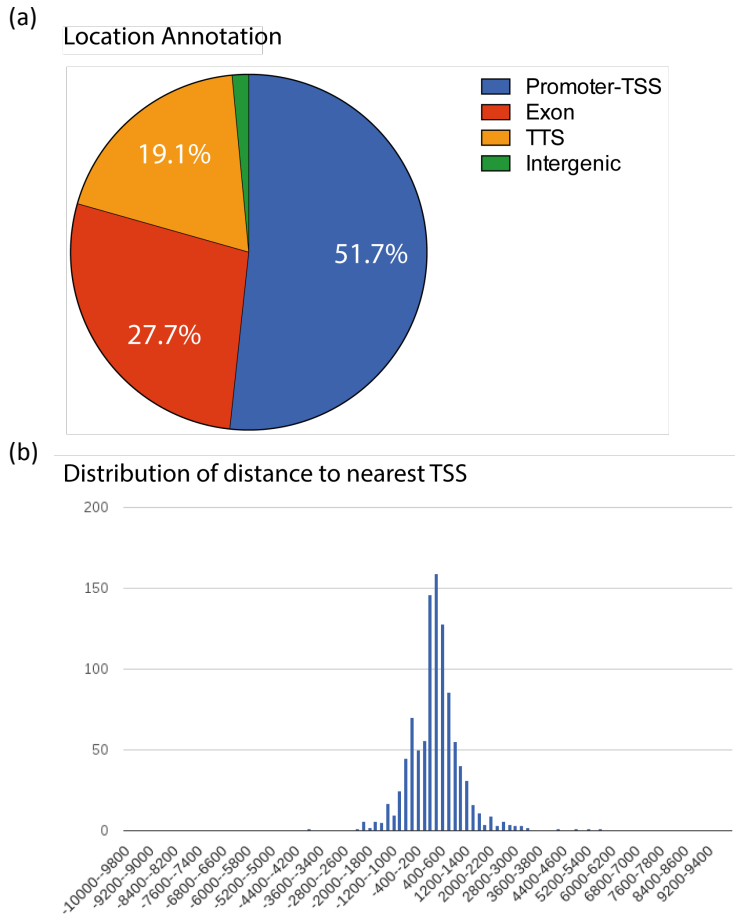
(a)

(b)

Figure 7. Location of Ixr1 binding peaks in Z1580 Ixr1-(c-Myc) tagged strain growth in conditions of aerobiosis and represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS).

Although Met4 is the major transcriptional activator of the genes from the assimilation of sulphate and sulphur compound biosynthesis, Met28 and other

regulatory proteins (Met31, Met32, Cbf1), which lack intrinsic transcriptional activation domains, act as adaptors for recruiting Met4 to appropriate promoters (Blaiseau *et al.*, 1997; Kuras *et al.*, 1996). Therefore it might be possible that the effect of Ixr1 observed in the transcriptome might be mediated indirectly by control of these adaptors.

3.5.- Determination of Ixr1 binding during hypoxia by ChIP-on-chip analysis and its association to transcriptional regulation mediated by Ixr1

The analysis was done according to the description in section materials and methods. The summary of peaks distribution and length is shown in Figure 8, rendering 540 peaks (*p-value* < 0.01), 271 related to promoter regions, that represents 50.2% of the total peaks obtained.

The list of selected peaks is summarized in supplementary Table S3. Analysis of these selected peaks with YEASTMINE reveals the enrichment of two functional groups of genes; those related to gluconeogenesis and to oxidation-reduction processes. The gene *HEM13,* a hypoxic target for Ixr1 regulation already characterized in previous analyses (Castro-Prego *et al*. 2010a), is included in this list.

An analysis of the intersection between the list of genes regulated by Ixr1 in hypoxia (Suppletory table S2 of chapter 1) and these found through immunoprecipitation mediated by Ixr1 in hypoxia suggests that, similarly to the situation described in normoxia, the direct and stable regulation by Ixr1 binding to the target promoters is limited to a reduced number of genes (14.02%, table S4). Curiously again the genes related to branched chain amino acids biosynthesis (*ILV2* and *ILV3*) are among those showing direct binding of Ixr1.

(a) Location Annotation
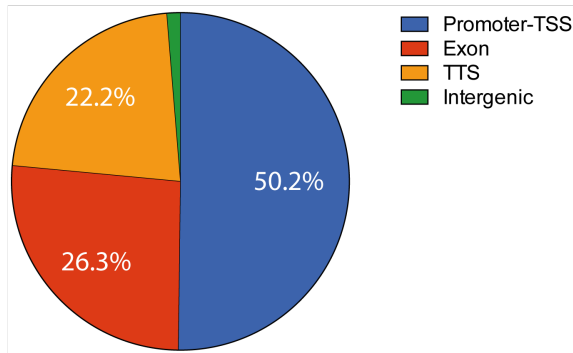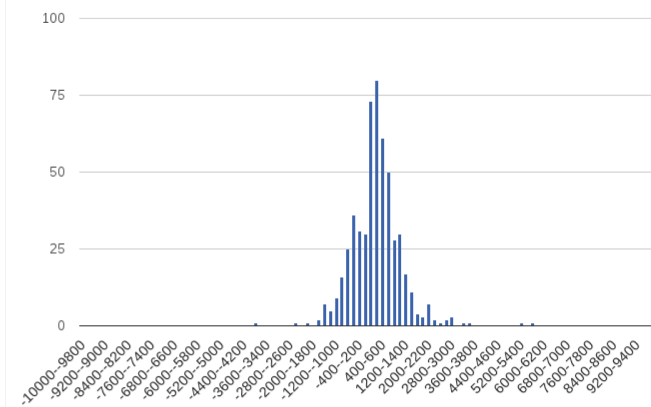

(b) Distribution of distance to nearest TSS

Figure 8. Location of Ixr1 binding peaks in Z1580 Ixr1-(c-Myc) strain growth in conditions of hypoxia and represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS).

3.6.- Comparative of Ixr1 and Rox1 binding during normoxia

As previously explained, some genes are simultaneously regulated by Ixr1 and Rox1 during normoxia. This is the case of *HEM13*, which is repressed both by Rox1 and Ixr1 during normoxia and the two regulators compete for binding to regulatory signals in the *HEM13* promoter (Castro-Prego *et al*. 2010a). A wide-genome analysis of Rox1 binding was not available from previous analyses, despite

the binding of its unique HMG-box domain had been previously explored for DNA binding *in vitro* (GEO accession number GSE13751). However these results are not directly comparable to the binding of the complete protein *in vivo*. Therefore we carried out the analysis for Rox1 binding in the same conditions and genetic background as previously selected for assaying Ixr1 binding. The summary of peaks distribution and length is shown in Figure 9, rendering 1486 peaks (*p-value* < 0.01), 547 related to promoter regions, that represents 36.8% of the total peaks obtained.

The list of selected peaks is summarized in supplementary Table S5. Analysis of these selected peaks with YEASTMINE reveals the enrichment of two functional groups of genes; those related to ribosomal proteins and translation and those related to alpha amino acids biosynthesis.

The intersection of peaks obtained for Rox1 binding and transcriptome data comparing wild type and *Δrox1* strains previously reported (Kwast *et al*., 2002) indicates high consistence between both data sets with 96 genes in common (Table S6). Genes for ergosterol biosynthesis are highly represented in this intersection.

3.7.- Binding of Ixr1 after depletion of Rox1 during normoxia

Rox1 and Ixr1 have in common the presence of HMG-box domains for binding DNA targets. Therefore in order to test the hypothesis whether there was a competition for DNA binding between these two transcriptional factors a ChIP-on-chip analysis of Ixr1 binding was carried after deletion of *ROX1* gene. The summary of peaks distribution and length is shown in Figure 10, rendering 342 peaks (*p-value* < 0.01), 217 related to promoter regions, that represents 63.5% of the total peaks obtained.

(a) Location Annotation
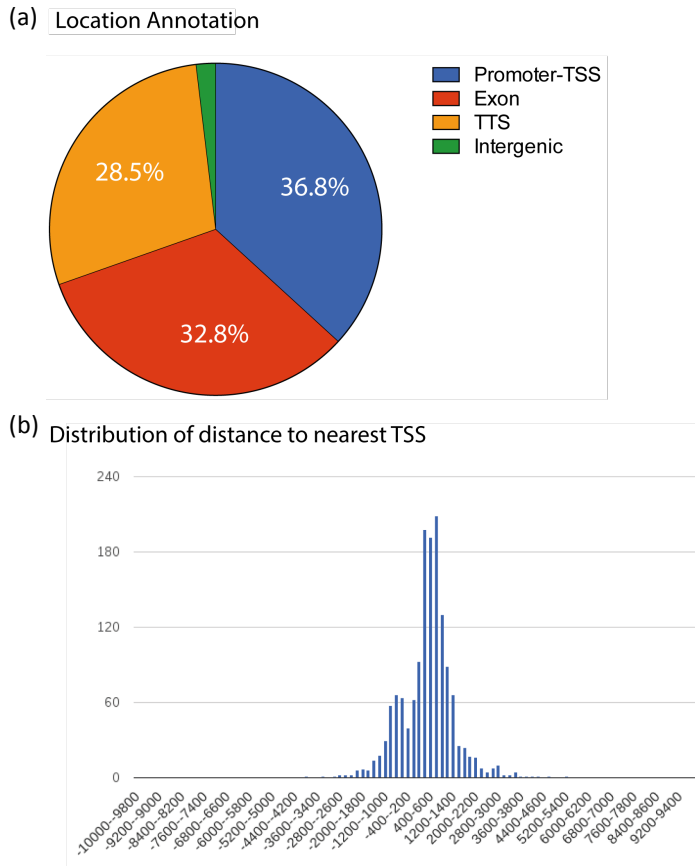

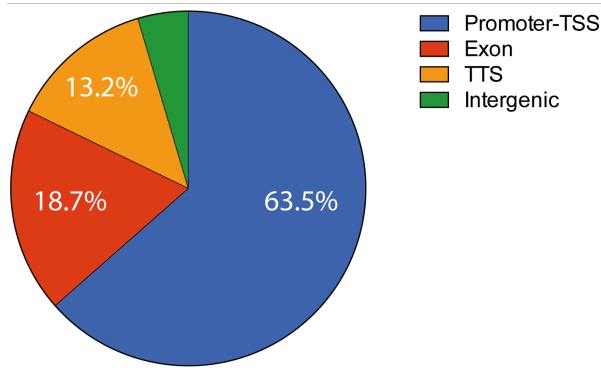(b) Distribution of distance to nearest TSS

Figure 9. Location of Ixr1 binding peaks in Z1465 Rox1-(c-Myc) strain growth in conditions of aerobiosis and represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS).

117

(a) Location Annotation
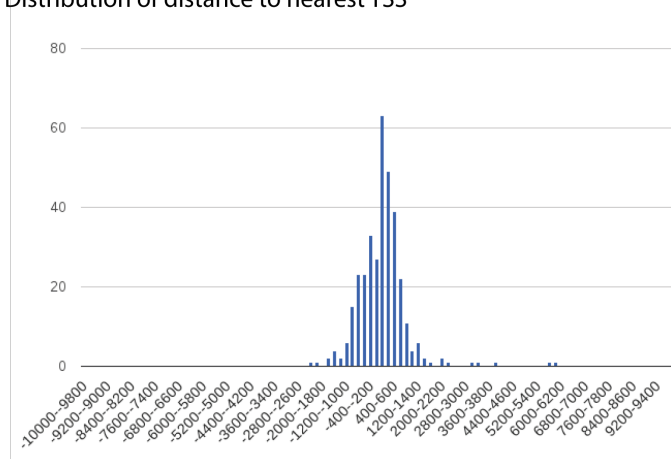

(b) Distribution of distance to nearest TSS

Figure 10. Location of Ixr1 binding peaks in Z1580 Ixr1-(c-Myc) *Δrox1* strain growth in conditions of aerobiosis and represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS).

The list of selected peaks is summarized in supplementary table S7. Analysis of these selected peaks with YEASTMINE reveals the enrichment of genes related to ribosomal proteins and translation. It is interesting to remark that these two functional groups are coincident with those found for Rox1 binding in normoxia. Taken into account that Rox1 is expressed at higher levels during normoxia than during hypoxia (Lowry *et al*., 1990), and the opposite occurs in reference to Ixr1 (Castro-Prego *et al*, 2010b.), these results could reflect that even

low levels of Ixr1 may occupy the sites for Rox1 binding in its absence. Therefore it is also possible that in the transition from aerobiosis to hypoxia, accompanied by a decrease of Rox1 and an increase of Ixr1 levels, a similar replacement might occur.

3.8.- *In silico* analysis of the DNA sequences obtained from the peaks after immunoprecipitation and localized in promoter regions

As reported in the previous sections, the majority of the DNA binding sequences found for Rox1 or Ixr1 by ChIP-on-chip analysis are not related to a functional transcriptional regulation mediated by stable binding of the transcriptional factors, as demonstrate the intersections between regulated genes and physically bound sites. In this section we describe the results obtained by *in silico* analysis of the DNA sequences obtained from the peaks after immunoprecipitation but limiting this analysis to peaks localized in promoter regions.

We first investigate the presence of the core consensus for Rox1 binding (***ATTGTT***) and the core consensus for Ixr1 binding (***AAG[G/C]GG***) in the selected peaks according to promoter positioning. We used the "dna- pattern" search program through the RSA tools facilities (http://rsat-tagc.univ-mrs.fr/rsat/). The results are summarized in figure 11.

The results summarized in figure 11 show that, although Rox1 is considered a transcriptional factor with a sequence-specific binding to target regulated promoters during normoxia, only 21.9 % of the analyzed sequences contain the core ***ATTGTT*** conserved in the consensus defined for the binding of this factor. Therefore the remaining detected interactions (78.1%) are not specific of sequence, or might be specific but occurring through other consensus not discovered, or they are reached indirectly and implicating other protein/s forming a protein complex for the interaction with DNA, in which other partner is

119

responsible of the specificity. These alternatives are shown in the scheme of figure 12a. The comparison of the percent of identification of the core **ATTGTT** (Rox1-specific) and **AAG[G/C]GG** (Ixr1-specific) among the peaks after immunoprecipitation against Rox1-tagged protein reveals an enrichment of 2.3 folds (21.9/9.7) in favour of binding to the specific Rox1 consensus (Figure 11).

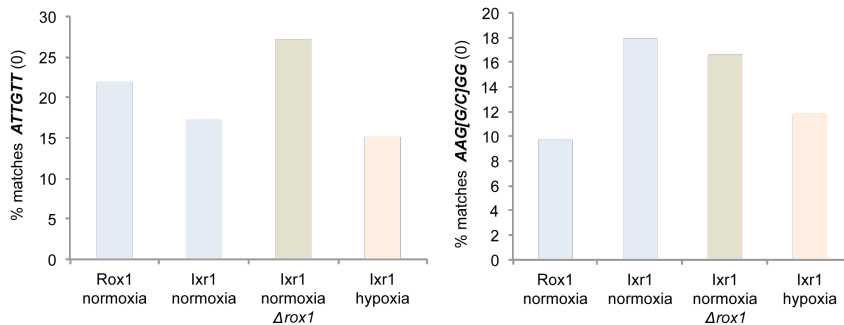|	ChIP-on-chip											
PATTERN	max	count	%	max	count	%	max	count	%	max	count	%
ATTGTT (0)	547	120	21.9	519	90	17.3	217	59	27.2	271	41	15.1
ATTGTT (1)	547	474	86.7	519	434	83.6	217	191	88	271	213	78.6
	Normoxia			Normoxia			Normoxia *Δrox1*			Hypoxia		
	Rox1 ChIP						**Ixr1 ChIP**					
PATTERN	max	count	%	max	count	%	max	count	%	max	count	%
AAG[G/C]GG (0)	547	53	9.7	519	93	17.9	217	36	16.6	271	32	11.8
AAG[G/C]GG (1)	547	439	80.3	519	431	83	217	185	85.3	271	211	77.9
	Normoxia			Normoxia			Normoxia *Δrox1*			Hypoxia		

Figure 11. ChIP-on-chip sequences of peaks located in promoter regions have been analyzed with the program DNA pattern in RSA tools looking for the core of the (*Figure 11 continued*) consensus for Rox1 binding (**ATTGTT**) or Ixr1 binding (**AAG[G/C]GG**) with 0 or 1 substitutions. *Max* represents the number of analyzed peak sequences; count the number of consensus matches and % the percent of counts.

The results shown in figure 11 about the consensuses found in the peaks obtained by immunoprecipitation against the Ixr1-tagged protein are consistent with the fact that Ixr1 can bind to DNA both with sequence-dependent specificity

and in a way independent of sequence specificity. Indeed, more than 80% of binding may be considered independent of the presence of the core ***AAG[G/C]GG*** in normoxia or hypoxia. The same mechanisms proposed to explain Rox1 binding are valid for explaining Ixr1 binding (Figure 12b). However, for Ixr1 binding, the presence of the cores ***ATTGTT*** or ***AAG[G/C]GG*** are very similar (17.3% and 17.9% during normoxia and 15.1% and 11.8% during hypoxia). This could indicate that the recognition by sequence specificity is minor than the observed for Rox1, and also that both core sequences are recognized similarly by Ixr1. A remarkable feature is that, during normoxia, the elimination of Rox1 increases the percentage of Ixr1 binding to the core ***ATTGTTT*** (from 17.3% to 27.2%) but not to the core ***AAG[G/C]GG*** (17.9% and 16.6% respectively).

In order to find other sequence-specific consensuses for Rox1 and Ixr1 binding not yet defined in the literature or already derived from the sequences found in the promoters of regulated genes (as in previous sections of this chapter), we carried several *in silico* analyses of the sequences derived from the ChIP-on-chip results, using different available programs. Results obtained with the program "oligo analysis" through RSA tools are summarized in figure 13. Results obtained with the program "peak-motifs" through RSA tools are depicted in figure 14.

The results presented in figure 13 show that the "*de novo*" consensus more frequently found in the sequences analyzed is common to immunoprecipitations carried against Rox1 or Ixr1 tagged proteins during normoxia (frequency: 20.5 % for Rox1 and 12.5 % for Ixr1) and shares the core ***AGCAG[G/C]***. The frequency of appearance of a similar consensus containing this core diminishes to 3.7% in the analysis of Ixr1 binding to promoter regions during hypoxia (figure 13). The core ***AGCAGC*** aligned to the JASPAR database with TOMTOM reveals similarity to the sequences bound by the yeast transcription factors Ace2 (*p-value* =0.0056) or Swi5 (*p-value*= 0.0061).
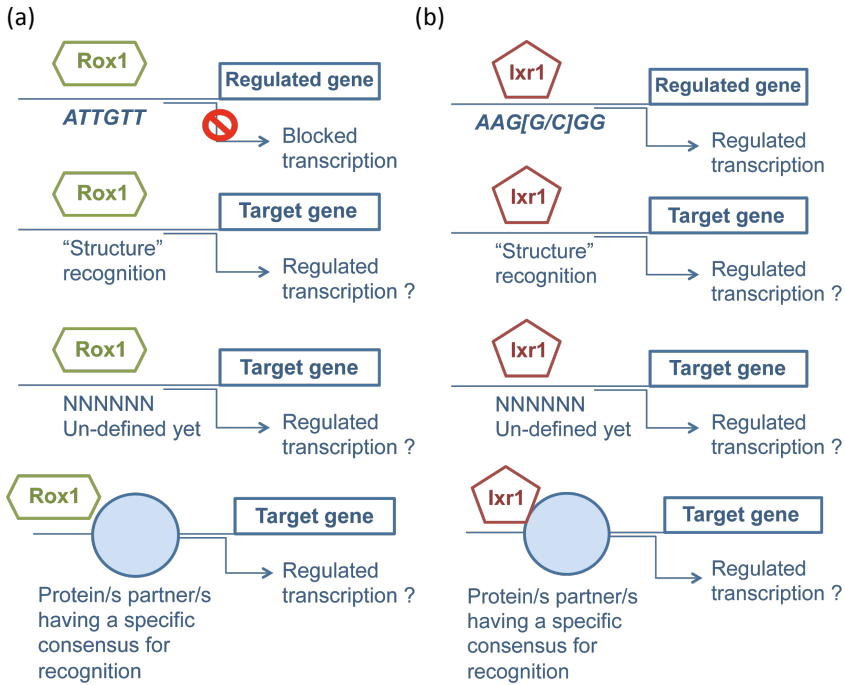
121

Figure 12. Scheme of possible mechanisms of physical interactions of Rox1 (a) and Ixr1 (b) with DNA, related or not related to transcriptional regulation.

The discovered motifs using the program "peak-motifs" through the RSA-tools suite are summarized in figure 14. The parameters for this search were established with a restricted length for each peak of ± 50 bp in each side of peak center. Coincidences with the sequence for Ace2 and Swi5 binding are also reported in the data sets of Rox1 immunoprecipitation during normoxia and Ixr1 immunoprecipitation during hypoxia. Besides, multiple repetitions for Rap1 binding are found in all the tested conditions.
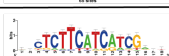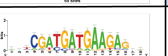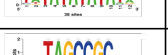
	Max	Consensus D	Logo D	Logo R	Matches	%
Rox1 **Normoxia**	547	kycTT**AGCAGC**Agyc			25	4,6
	547	ky**AGCACC**wd			112	20,5
	547	wcCTTACCGky			33	6,0
	547	ykh**CGATGA**Tww			24	4,4
Ixr1 **Normoxia**	519	ykw**CGATGAA**GAGsw			28	5,4
	519	yaGC**AGCAGC**rgcas			65	12,5
	519	wkCTCTTCATCATCGabt			18	3,5
	519	wrCAACAGCwk			48	9,2
	519	rsCACCACCAcb			23	4,4
Ixr1 **Hypoxia**	271	wkC**AGCAGC**Gst			10	3,7
	271	ygA**CGATGAA**rw			7	2,6
	271	mtTCAATGGCwt			13	4,8
	271	taTATATATAta			38	14,0
	271	ysGCGGCTAmh			16	5,9

Figure 13. Summary of results obtained through the olygo-analysis program on the peaks positioned at promoter regions.

3.9.- General outlines about Ixr1 binding in normoxia and hypoxia deduced from the analyses

In silico analysis of promoter sequences from the genes regulated by Ixr1 in normoxia and hypoxia shows that they include consensus sequences previously characterized for Rox1 (Lowry *et al*., 1990) and Ixr1 binding (MacIsaac *et al*., 2006). The interaction of Ixr1 with the target sequence defined for Rox1 binding is consistent with the feature that both transcriptional factors bind DNA through

HMG motifs (Balasubramanian *et al*., 1993; Deckert *et al*., 1995) and the core **ATTGTT** present in the Rox1 consensus for binding is recognized by the HMG motifs of many different proteins (Weir *et al*., 1993; Weiss & King, 1995; Kim *et al*., 2015). In the *HEM13* promoter region, binding of Ixr1 to the Rox1 site had been previously proved *in vitro* and *in vivo* (Castro-Prego *et al*., 2010a) and our data indicate that this feature may be generalized to other promoters. However, other sequence with no similarity to the Rox1 consensus was also defined as specific for Ixr1 binding (MacIsaac *et al*., 2006). Our results indicate that a redefined version of this Ixr1-consensus, restricted to the core **AAG[G/C]GG**, is even more frequent in the regulated promoters than the **ATTGTT** core.

ChIP-on-chip	Max peaks analyzed	Consensus	% matches	Max sites per peak	Mean sites per peak	Matching to known TF
Rox 1 Normoxia	547	KGKGKGTGTGGKKKK	1,28	32	23,29	Rap1, Aft2, Rpn4, Met32
	547	scTAGCAGCvk	2,38	2	1,08	Swi5, Hac1, Rfx1, Mal63
	547	srGCGGCTAkm	3,47	2	1,11	Ume6, Mal63, War1, Car1, Spt1
Ixr1 Normoxia	519	GtGTGGGTGTGGkkk	1,16	19	13,7	Rap1, Aft2, Met32
	519	kvGTGGTGrw	2,12	3	1,17	Rap1, Aft2, Met32
	519	kKkkGTGTGkGkGTGkkGkGkGk	1,16	27	21,83	Rap1
Ixr1 Hypoxia	271	KGTGTGGGTGTGGKK	2,95	18	10,75	Rap1, Aft2, Aft1
	271	mTAGCAGCGb	4,8	2	1,08	Swi5, Hac1, Rfx1, Mal63
	271	KKGTGTGGGTGTGGKKKKK	2,58	20	13,14	Rap1
	271	hmCAGCAGCGby	5,9	2	1,12	Swi5, Hac1, Ace2
	271	KKKGTGTGKGTGTGKKKK	2,58	26	17,71	Rap1

Figure 14. Summary of results obtained through the "peak motifs" program on the peaks positioned at promoter regions and sequences restricted to ± 50 bp from the center of the peak.

Only a fraction (ranging from ≈12% to ≈18%) of the genes that have either the **ATTGTT** or the **AAG[G/C]GG** core in their promoters are functionally regulated by Ixr1 as deduced from the comparison of tables from transcriptome and ChIP-on-chip analysis. This indicates that also transient or indirect effects are important in controlling transcriptional regulation by Ixr1.

Both during normoxia and hypoxia Ixr1 binds to DNA without observed enrichment in promoter regions (ranging from ≈50% to ≈60% of the total significative peaks), which is in accordance with the predictable behaviour of its two different HMG-boxes, one sequence specific and another not sequence specific (Castro-Prego et al., 2010a). Our data reveal that even in the interactions detected through the promoters, not all of them, but rather a small fraction in the approximate range of 20-30%, may be considered sequence-specific as deduced from the presence of the ***ATTGTT*** or the ***AAG[G/C]GG*** cores.

Search of sequences that can produce Ixr1 binding by protein-protein interactions with other transcriptional regulators indicates that matches to previously known sequences for binding to yeast transcriptional regulators occurs in the promoter- specific peaks. However it is not possible to find a particular enrichment for the consensus sequence of only one transcriptional factor or at least a reduced number of them. The most frequent consensuses found include the core ***AGCA[G/C]C*** which is also present in the target sequences of yeast transcription factors Ace2 and Swi5 among others.

5.- REFERENCES

Bailey T.L. & Elkan C. (1994) "Fitting a mixture model by expectation maximization to discover motifs in biopolymers" Proc Int Conf Intell Syst Mol Biol. **2**:28-36.

Balakrishnan R., J. Park, K. Karra, B.C. Hitz, G. Binkley, E.L. Hong, J. Sullivan, G. Micklem, J.M. Cherry (2012) "YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit" Database (Oxford).

Bandyopadhyay S., M. Mehta, D. Kuo, M.K. Sung, R. Chuang, E.J. Jaehnig, B. Bodenmiller, K. Licon, W. Copeland, M. Shales, D. Fiedler, J. Dutkowski, A. Guénolé, H. van Attikum, K.M. Shokat, R.D. Kolodner, W.K. Huh, R. Aebersold, M.C. Keogh, N.J. Krogan, T. Ideker (2010) "Rewiring of genetic networks in response to DNA damage" Science **330**(6009):1385-9.

Beck T., M.N. Hall (1999) "The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors" Nature **402**(6762):689-92.

Blaiseau P.L., A.D. Isnard, Y. Surdin-Kerjan, D. Thomas (1997) "Met31p and Met32p, two related zinc finger proteins, are involved in transcriptional regulation

of yeast sulfur amino acid metabolism" Molecular and Cellular Biology **17**(7):3640-8.

Bourdineaud J.P., G. De Sampaïo, G.L. Lauquin (2000) "A Rox1-independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the SRP1/TIR1 gene" Molecular Microbiology **38**(4):879-90.

Bruhn, S.L., P.M. Pil, J.M. Essigmann, D.E. Housman & S.J. Lippard (1992) "Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin" Proceedings of the National Academy of Sciences USA 89(6):2307-11.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso, M.E. Cerdán (2010a) "Ixr1p regulates oxygen-dependent *HEM13* transcription" FEMS Yeast Research **10**(3):309–321.

Castro-Prego M., M. Lamas-Maceiras, P. Soengas, I. Carneiro, M.I. González Siso, M.E. Cerdán (2010b) "Regulatory factors controlling transcription of *Saccharomyces cerevisiae IXR1* (*ORD1*) by oxygen levels. A model of transcriptional adaptation from aerobiosis to hypoxia implicating *ROX1* and *IXR1* cross-regulation" Biochemical Journal **425**(1):235–243.

Chen H., J.J. Workman, A. Tenga, R.N. Laribee (2013) "Target of rapamycin signaling regulates high mobility group protein association to chromatin, which functions to suppress necrotic cell death" Epigenetics Chromatin 6(1):29.

Chen T.W., H.P. Li, C.C. Lee, R.C. Gan, P.J. Huang, T.H. Wu, C.Y. Lee, Y.F. Chang, P. Tang (2014) "ChIPseek, a web-based analysis tool for ChIP data. BMC Genomics" 15:539.

Deckert, J., R. Perini, B. Balasubramanian, and R.S. Zitomer (1995) "Multiple elements and auto-repression regulate Rox1, a repressor of hypoxic genes in Saccharomyces cerevisiae" Genetics **139**(3):1149-58.

Estruch F & Carlson M (1993) "Two homologous zinc finger genes identified by multicopy suppression in a *SNF1* protein kinase mutant of *Saccharomyces cerevisiae*" Molecular and Cellular Biology **13**(7):3872-81.

Gietz R.D., A. Sugino (1988) "New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites" Gene **74**(2):527-34.

Gorner W, Durchschlag E, Martinez-Pastor MT, Estruch F, Ammerer G., B. Hamilton, H. Ruis, C. Schüller (1998) "Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity" Genes & Development **12**(4):586-97.

Grosschedl R., K. Giese, J. Pagel (1994) "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures" Trends in Genetics **10**(3):94-100. Review.

Gupta S., J.A. Stamatoyannopoulos, T.L. Bailey, W.S. Noble (2007) "Quantifying similaritybetween motifs" Genome Biology **8**(2):R24.

Hodge M.R., K. Singh, M.G. Cumsky (1990) "Upstream activation and repression elements control transcription of the yeast COX5b gene" Molecular and Cellular Biology **10**(10):5510-20.

Hughes J.D., P.W. Estep, S. Tavazoie, G.M. Church (2000) "Computational identification of cis-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae*." Journal of Molecular Biology **296**(5): 1205-1214.

Jantzen, H.M., A. Admon, S.P. Bell, and R. Tjian (1990) "Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins" Nature **344**(6269):830-6.

Kuras L., H. Cherest, Y. Surdin-Kerjan, D. Thomas. (1996) "A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism" EMBO Journal **15**(10):2519-29.

Kwast K.E., L.C. Lai, N. Menda, D.T. James, S. Aref, P.V. Burke (2002) WGenomic analyses of anaerobically induced genes in Saccharomyces cerevisiae: functional roles of Rox1 and other factors in mediating the anoxic response" Journal of Bacteriology **184**(1):250-65.

Lai L.C., A.L. Kosorukoff, P.V. Burke, K.E. Kwast (2005) "Dynamical remodeling of the transcriptome during short-term anaerobiosis in *Saccharomyces cerevisiae*: differential response and role of Msn2 and/or Msn4 and other factors in galactose and glucose media" Molecular and Cellular Biologiy **25**(10):4075-91.

Landsman D., M. Bustin (1993) "A signature for the HMG-1 box DNA-binding proteins. Bioessays" **15**(8):539-46. Review.

Lambert J.R., V.W. Bilanchone, M.G Cumsky (1994) "The ORD1 gene encodes a transcription factor involved in oxygen regulation and is identical to *IXR1*, a gene that confers cisplatin sensitivity to *Saccharomyces cerevisiae*" Proceedings of the National Academy of Sciences **91**(15):7345-9.

Lee T.I., N.J. Rinaldi, F. Robert, D.T. Odom, Z. Bar-Joseph, G.K. Gerber, N.M. Hannett, C.T. Harbison, C.M. Thompson, I. Simon, J. Zeitlinger, E.G. Jennings, H.L. Murray, D.B. Gordon, B. Ren, J.J. Wyrick, J.B. Tagne, T.L. Volkert, E. Fraenkel, D.K. Gifford, R.A. Young (2002) "Transcriptional regulatory networks in *Saccharomyces cerevisiae*" Science **298**(5594):799-804.

Liu J. & A. Barrientos (2013) "Transcriptional regulation of yeast oxidative phosphorylation hypoxic genes by oxidative stress" Antioxidant & Redox Signal. 9(16):1916-27.

Lowry, C.V., M.E. Cerdan, R.S. Zitomer (1990) "A hypoxic consensus operator and a constitutive activation region regulate the ANB1 gene of *Saccharomyces cerevisiae*" Molecular and Cellular Biology **10**(11):5921-6.

MacIsaac K.D., T. Wang, D.B. Gordon, D.K. Gifford, G.D. Stormo, E. Fraenkel (2006) "An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*". BMC Bioinformatics 7:113.

Oosterwegel, M., M. van de Wetering, D. Dooijes, L. Klomp, A. Winoto, K. Georgopoulos, F. Meijlink, H. Clevers (1991) "Cloning of murine TCF-1, a T cell-specific transcription factor interacting with functional motifs in the CD3-e and TCR-a enhancers" The Journal of Experimental Medicine **173**(5):1133-42.

Pavesi G., P. Mereghetti, F. Zambelli, M. Stefani, G. Mauri, G. Pesole (2006) "MoD Tools: regulatory motif discovery in nucleotide sequences from co-regulated or homologous genes." Nucleic Acids Research **34**(Web Server issue): W566-70.

Robinson M.D., J. Grigull, N. Mohammad, T.R. Hughes (2002) "FunSpec: a web-based cluster interpreter for yeast" BMC Bioinformatics. 2002 Nov 13;3:35.

Romer K.A., G.R. Kayombya, E. Fraenkel (2007) "WebMOTIFS: automated discovery, filtering and scoring of DNA sequence motifs using multiple programs and Bayesian approaches." Nucleic Acids Research Web Server Issue.

Shobhit G., J.A. Stamatoyarmopolous, B. Timothy, W.S Noble (2007) "Quantifying similarity between motifs" Genome Biology **8**(2):R24,2007.

Teixeira M.C., P. T. Monteiro, J.F. Guerreiro, J.P. Gonçalves, N.P. Mira, S.C. dos Santos, T.R. Cabrito, M. Palma, C.Costa, A.P. Francisco, S.C. Madeira, A.L. Oliveira, A.T. Freitas, I. Sá-Correia (2014). "The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae." Nucleic Acids Research. 42(Database): D161-D166.

Thomas J.O. & A.A. Travers (2001) "HMG1 and 2, and related 'architectural' DNA-binding proteins" Trends in Biochemical Sciences 26(3):167-74. Review.

Turatsinze J.V., M. Thomas-Chollier, M. Defrance, J. van Helden (2008) "Using RSAT to scan genome sequences for transcription factor binding sites and cis-regulatory modules" Natures Protocols **3**(10):1578-88.

van Helden J., B. André, J. Collado-Vides (1998) "Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies" Journal of Molecular Biology 281(5):827-42.

van Helden J. (2003) "Regulatory sequence analysis tools" Nucleic Acids Research **31**(13):3593-6.

Waterman, M.L., W.H. Fischer, K.A. Jones (1991) "A thymusspecific member of the HMG protein family regulates the human T-cell receptor Ca enhancer" Genes & Development. **5**: 656- 669.

Liu X., D.L. Brutlag, J.S. Liu (2002) "An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments" Nature Biotechnology **20**(8): 835-9.

Vizoso-Vázquez A., M. Lamas-Maceiras, M. Becerra, M.I. González-Siso, E. Rodríguez-Belmonte, M.E. Cerdán (2012) "Ixr1p and the control of the *Saccharomyces cerevisiae* hypoxic response" Applied Microbiology and Biotechnology **94**(1):173-84.

Zheng J., J.J. Benschop, M. Shales, P. Kemmeren, J. Greenblatt, G. Cagney, F. Holstege, H. Li, N.J. Krogan (2010) "Epistatic relationships reveal the functional organization of yeast transcription factors" Molecular Systems Biology. 2010 Oct 5;6:420.

Zitomer R.S., B.D. Hall (1976) "Yeast cytochrome c messenger RNA. *In vitro* translation and specific immunoprecipitation of the *CYC1* gene product" Journal of Biological Chemistry" **251**(20):6320-6.

Zitomer R.S. & C.V. Lowry (1992) "Regulation of gene expression by oxygen in *Saccharomyces cerevisiae*" Microbiology Reviews **56**(1):1-11. Review.

Chapter 3

Deciphering Ixr1 function in the response of yeast cells to cisplatin

SUMMARY

Ixr1 is a HMGB protein that binds to adducts formed between cisplatin and DNA and its deletion causes increased resistance to cisplatin in yeast. Previous data indicated that Ixr1 and other proteins with HMG-domains might induce cell death by blocking repair of the major cisplatin-DNA adducts *in vivo,* but the mechanisms of cisplatin resistance in yeast *ixr1* mutants remain mostly unknown. Results exposed in this work reveal that cisplatin resistance in *Δixr1* mutants may be also attributed to changes in the regulation of ribosome biogenesis pathways, which are less down-regulated by cisplatin in the mutant that in the wild type, thus increasing the proliferative capacity of the mutant cells *versus* wild type. In this sense, the interconnections between Ixr1 mediated regulation and the TOR signaling pathway, highly connected to the control of ribosome biogenesis, are discussed. A higher enhancement of sulfur-compounds metabolism is also observed in the mutant than in the wild type. This may also contribute to increase the availability of bio-molecules bearing -SH chelating groups to immobilise cisplatin as well as to promote glutathione biosynthesis, thus favouring anti-oxidant reactions or extrusion of cisplatin-glutathione complexes out of the cell.

1.- INTRODUCTION

Cisplatin is used for the clinical treatment of testicular and ovarian cancers because it has high effect against them and is it also used for treating esophageal, cervical, head and neck, bladder, and small cell lung cancer (Giaccone, 2000). However others, like colorectal and non-small cell lung cancers have intrinsic resistance to cisplatin. Even those initially sensitive, like ovarian or small cell lung cancer, might acquire resistance after the initial treatment (Pérez, 1998). The molecular mechanisms of cisplatin action inside the tumor cells require the binding of the drug to DNA and non-DNA targets, the activation of intracellular signal pathways and finally the induction of cell-death through apoptosis, necrosis, or both. Diverse and extensive data about the cytotoxic action of cisplatin and the

mechanisms of drug resistance have been thoroughly reviewed (Fuertes *et al*., 2002; Siddik, 2003; Basu & Krishnamurthy, 2010; Tanida *et al*., 2012).

Saccharomyces cerevisiae is a good eukaryotic model to find genes related to cisplatin-sensitivity or resistance (Fox *et al*., 1994; Huang *et al*., 2005; Schenk *et al*., 2001, 2003). Although 22 cisplatin-related yeast genes have been identified in wide-genome screenings, (Huang *et al*., 2005) little is known about their specific role in the cellular response to cisplatin. The gene *IXR1* encodes a protein, belonging to the family of chromatin non histone proteins with two HMG-domains, that binds to the adducts that cisplatin forms with DNA (Brown *et al*., 1993) and its deletion causes increased resistance to cisplatin in yeast (McA´Nulty & Lippard, 1996). The hypothesis about Ixr1 and other HMG-domain proteins may block repair of the major cisplatin-DNA adducts *in vivo,* thus inducing cell death, was postulated (McA´Nulty & Lippard, 1996). It is based in the evidence that *IXR1* deletion does not increase the resistance of *S. cerevisiae* mutant cells, which already carry mutations in the genes *RAD2*, *RAD4* and *RAD14* related to excision repair mechanisms (McA´Nulty & Lippard, 1996). More recently it has been reported that *IXR1* mutations increase the rate of spontaneous mutagenesis mediated by replication errors (Fedorov *et al*. 2010) and that Ixr1 is required for the maintenance of an adequate supply and balance of dNTPs for DNA synthesis and repair (Tsaponina *et al*. 2011). Besides, Ixr1 also regulates the yeast hypoxic response (Lambert *et al*. 1994; Bourdineaud *et al*. 2000; Castro *et al*. 2010a; Castro *et al*. 2010b; Vizoso-Vázquez *et al*., 2011).

The increased resistance of cisplatin observed in yeast strains carrying *SKY1* and *IXR1* deletions has in common several changes produced on DNA repair pathways (Chao, 1991; Liu, 1996; Reed, 1998; Schenk, 2002). The transcriptional response to cisplatin and its dependence on Sky1 function has been recently reported (Rodríguez Lombardero *et al*., 2014). Sky1 is a *S. cerevisiae* rich serine-arginine (SR) protein-specific kinase whose enzymatic activity is necessary for the

cytotoxic effect of cisplatin (Schenk, 2001). There is a high increase in mRNA levels of genes related to sulfate assimilation and metabolism of sulfur compounds upon cisplatin treatment of the yeast strain W303 (Rodríguez Lombardero *et al.*, 2014). The genes for purine synthesis are also up-regulated after the cisplatin treatment (Rodríguez Lombardero *et al.*, 2014). These transcriptional changes might represent a cellular response favoring chelation of cisplatin with sulfur-containing amino acids and also facilitating DNA repair by the stimulation of purine biosynthesis. The transcription pattern of stimulation of sulfur-containing amino acids and purine synthesis decreased, or even disappeared, in a *Δsky1* strain (Rodríguez Lombardero *et al.*, 2014). The effect of Sky1 and Ixr1 deletions on cisplatin resistance is additive in the W303 strain, which might be interpreted in the sense that Ixr1 and Sky1 control cisplatin resistance by two independent mechanisms, each contributing partially to the final phenotype (Rodríguez Lombardero *et al.*, 2011). Therefore in this work we analyze the effects of Ixr1 deletion on the transcriptional response to cisplatin in the yeast strain W303.

Considering that Ixr1 has two HMG-box domains that can recognize specific consensus sequences in the promoters of regulated genes (Bordineaud *et al.*, 2000; Castro-Prego *et al.*, 2010a; Castro-Prego *et al.*, 2010b; Tsaponina *et al.*, 2011), but also bind preferentially to altered DNA structures as the adducts formed by DNA platination (McA´Nulty & Lippard, 1996) we have also explored the specific binding of Ixr1 to DNA by ChIP-on-chip experiments in the yeast strain W303 untreated or treated with cisplatin.

Previous data from our laboratory (yet unpublished) have demonstrated that Ixr1 physically interacts with the transcriptional activator Swi6. Swi6 has the ability to activate transcription but does not have DNA binding activity. Swi6 altogether with Swi4 form part of the SBF (<u>S</u>CB <u>*Binding*</u> <u>*Factor*</u>) complex that regulates the transcription of the cyclins Cln1 and Cln2 which are necessary in the transition G_1/S (Koch *et al.*, 1995). Besides, Swi6, altogether with Mbp1, forms part

of the MCB (*MluI* *C*ell cycle *B*ox) complex that regulates the transcription of genes related to DNA synthesis, as well as those encoding for the cyclins Clb5 and Clb6, controlling the chromosomal replication during the S phase (Verma *et al.*, 1992). Besides, Swi6 is involved in the response to DNA damage and it is phosphorylated by Rad53 in this condition (Sidorova *et al.*, 1997). Ixr1, as well as Rad53, controls the biosynthesis of dNTPs, which are necessary for DNA synthesis and repair, through the control of the gene *RNR1* and in response to DNA damage (Tsaponina *et al.*, 2011). In this work we have also tested whether the deletion of Swi6 affects the interaction of Ixr1 with DNA in response to DNA damage caused by cisplatin treatment.

The interest of knowing the mechanisms of action of Ixr1 in relation to the yeast response to cisplatin is due to previous observations showing that the human gene *HMGB1*, which has partial sequence similarity with *IXR1*, is over-expressed in cisplatin-resistant cell lines from human epidermoid cancer and transcriptionally regulated by the specific factor CTF/NF-1 (Nagatani *et al.*, 2001). Therefore, this study in yeast could open new research frontiers to future studies in human cell lines, which might be important to cope with a serious problem in the chemotherapy of cancer, as it is cisplatin resistance.

2.- MATERIAL AND METHODS

2.1.- Cell culture and treatments

The *S. cerevisiae* strain W303 (*MAT**a** ade2-1 can1-100 leu2-3,112 trp1-100 ura3- 52*) and its derivative W303-*Δixr1* previously described (Rodríguez Lombardero *et al.*, 2012) have been used in transcriptomic experiments. The *S. cerevisiae* strain Z1580 (*MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ psi+ IXR1:myc9::TRP1*) where obtained from Young's lab (Lee *et al.*, 2002). The knockout strain Z1580-*Δswi6* was obtained by one-step replacement with the

URA3 marker. The plasmid Yeplac195 (Gietz & Sugino, 1988) was used as a template to amplify a linear fragment containing the *URA3* gene and two flanking regions of homology to the ORF end of *SWI6* by PCR using the primers AVV284 and AVV285 (see table 1). After transformation of the Z1580 strain with amplified fragments, cells were selected in complete media without uracil. The correct replacement in the *S. cerevisiae* genome was verified by PCR as described previously (Tizón *et al.*, 1999) using primers designed inside the *URA3* ORF, ECV700AV and ECV701AV, and the flanking regions of *ROX1* and *SWI6* external to the recombination event, using the primers AVV286 and AVV287 (see table 1). The handling of yeast cells was carried out according to standard procedures.

Table 1. Oligos used to obtain and check Z1580-*Δswi6 mutant strains.*

Oligo name	Sequence	Gene	Strand[a]	Position[b]
ECV699AV	AAGAGATGAAGGTTACGATTGGT	*URA3*	C	+592
ECV700AV	TTGTACTTGGCGGATAATGC	*URA3*	W	+234
AVV284	atggcgttggaagaagtggtacgatacttaggacctcCCTT TAGCTGTTCTATATGCTGC	*SWI6*	C	+1
AVV285	tcttgcatttcgtcagtgtcaatgtcttgtgggtcttcagCCA CCTGACGTCTAAGAAACC	*SWI6*	W	+2387
AVV286	CTAGTGCGGTCATCTCTTGCG	*SWI6*	W	-394
AVV287	TCATTTACCGTCATGGTAAGGAC	*SWI6*	C	+2766

[a]W: Watson strand; C: Crick strand.
[b]Numbering is considering +1 for the adenine in the first start codon.

Three biological replicates of cultures and treatments were run. The yeast cells were pre-cultured overnight in 10 mL of complete medium (SD) containing per liter: 6.7 g of bacto-yeast nitrogen base without amino acids from Difco (Franklin Lakes, New Jersey, USA); 40 mg each of histidine, leucine, adenine, uracil, lysine and tyrosine, 10 mg each of arginine, methionine and threonine, 30 mg tryptophan; 60 mg each of phenylalanine and isoleucine; 2% glucose (w/v). For transcriptomic experiments, the following day the cells were inoculated at initial OD_{600} of 0.4 in 70 mL SD and grown in 250 mL Erlenmeyer flasks at 30 ^{o}C and with agitation at 250 rpm. When cells reached OD_{600} of 0.6, the cultures from each strain were divided in two aliquots of 25 mL (control and cisplatin treatment). A

stock solution of cisplatin 6 mM in dimethyl sulfoxide (DMSO) was prepared and the drug was added to the treated cultures at a final concentration of 600 µM. An equivalent volume of DMSO was added to the control cultures. The treatment was done at 30 ^{o}C and with agitation at 250 rpm during four hours in darkness. The concentration of cisplatin and the time course of the treatment were previously established in trial experiments with the selected yeast strains (Rodríguez Lombardero *et al.*, 2012).

On the other hand, for ChIP-on-chip experiments the cells were inoculated at initial OD_{600} of 0.1 in 200 mL SD and grown in 1 L Erlenmeyer flasks at 30 ^{o}C and with agitation at 250 rpm. When cells reached OD_{600} of 0.6, a stock solution of cisplatin 20 mM in dimethyl sulfoxide (DMSO) was prepared and the drug was added to the treated cultures at a final concentration of 600 µM. An equivalent volume of DMSO was added to the control cultures. The treatment was done at 30 ^{o}C and with agitation at 250 rpm during four hours in darkness. The concentration of cisplatin and the time course of the treatment were applied as determined for transcriptome analysis.

2.2.- RNA preparation and transcriptomic microarray analysis

RNA was extracted from a number of cells corresponding to OD_{600} of 3 with the Aurum^{TM} Total RNA Mini Kit (Bio-Rad) and by following the manufacturer's instructions. Concentration and purity of RNA was evaluated by measuring the ratio R= A_{260}/A_{280} (always in the range 1.7<R<2.1). RNA integrity was evaluated by the RIN parameter (RNA Integrity Number) with the 2100 Bioanalyzer (Agilent Technologies, Inc. Santa Clara, CA 95051-7201USA) according with manufacturer instructions. RIN was near to the value 9 in all the samples, which is considered high-quality extraction (Schroeder *et al.*, 2006).

Twelve GeneChip® Yeast-Genome-2.0 arrays from Affymetrix Inc. (Wycombe. United Kingdom) were used and processed in the GeneChip® System

with Autoloader from Affymetrix Inc. (Wycombe. United Kingdom). We started from 10 ng of total RNA from each sample for successive cDNA, aRNA generation, labeling with biotin and fragmentation using the GeneChip® 3' IVT Express Kit. RNA fragmentation was monitored with the 2100 Bioanalyzer (Agilent Technologies, Inc. Santa Clara, CA 95051-7201USA), selecting conditions producing fragments from 35-200 nt with a majority among 100-120 nt. Hybridization, washes and staining were done with the GeneChip® HT Hybridization, Wash and Stain Kit. (Ambion, Inc. Affymetrix). These kits include RNA Poly-A controls (lys, phe, thr and dap) from *Bacillus subtilis* to monitor the target labeling process and they serve as sensitivity indicators of target preparation and labeling efficiency. They also include the Hybridization Controls, which are comprised of a mixture of biotinylated and fragmented RNA of *bioB*, *bioC*, *bioD* (genes from the biosynthesis of biotin in *Escherichia coli*) and *Cre* (recombinase from bacteriophage P1). These controls monitor the hybridization, washing and staining steps. Control Oligo B2 is included to provide alignment signals for image analysis.

Image caption and preliminary data analysis were carried out with Affymetrix® Expression Console™ software (v1.1).

2.3.- Chromatin immunoprecipitation and ChIP-on-chip analyses

Chromatin immunoprecipitation experiments were carried out as described previously (Lee *et al*., 2002) with minor modifications. Shortly, 200 mL of yeast culture were collected at OD_{600} of ≈ 0.9-1. Cross-linking was performed by adding 1% formaldehyde to the culture and incubating at room temperature for 20 min. 125 mM of glycine was then added and culture was incubated 5 min. Cells were then harvested and washed four times with 50 mL Tris-HCl buffer saline (20 mM Tris-HCl, pH 7,5, 150 mM NaCl) at 4 $^{\circ}$C. The cell breakage was performed in 800 μL of lysis buffer (50 mM HEPES-KOH, pH 7,5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0,1% sodium deoxycholate, 2X complete protease inhibitor cocktail,

Roche, and 2X complete phosphatase inhibitor cocktail, Roche) with glass beads, and the cell extracts were sonicated for 5 min in 10 sec on/59 sec off cycles (chromatin was sheared into an average size of 400 bp). The immunoprecipitations were performed with magnetic Dynabeads (Invitrogen) following the manufacturer's instructions and using anti-(c-Myc) antibodes (sc47694; Santa Cruz Biotechnology) for specific Ixr1p–(c-Myc) immunoprecipitation. Negative controls, with rabbit IgG immunoprecipitation, were also performed. Samples were washed three times with 1 mL of Lysis buffer, three times with 1 mL of Lysis buffer high salt (50 mM HEPES-KOH, pH 7,5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0,1% sodium deoxycholate), three times with 1 mL of Wash buffer (10 mM Tris-HCl, pH 8, 250 mM LiCl, 0,5% NP-40, 0,5% sodium deoxycholate, 1 mM EDTA), and once with 1 mL of TE Buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA). Immunoprecipitations were then eluted in 250 μL of elution buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS) and treated overnight with 30 μL of proteinase K (20 mg/mL, NewEngland Biolabs). Next day, immunoprecipitated DNA were cleaned with the kit USB PrepEase DNA Clean-Up (USB). Next steps were rigorously followed in the manufacturer's instructions (Affymetrix) (http://cmgm.stanford.edu /pan/section_html/GE/protocols/Chromatin%20Immunoprecipitation%20Assay%2 0Protocol.pdf). Immunoprecipitation enrichment was checked by real-time qPCR against promoter regions of *TIR1, IXR1, ROX1* and *HEM13* (table 2), which are known to be bound by Ixr1 protein and immunoprecipitated using the antibody being investigated (Bordineaud *et al*., 2000; Castro-Prego *et al*., 2010a; Castro-Prego *et al*., 2010b). In all cases, >8-fold enrichments were obtained for IP samples compared to the IgG samples.

Eighteen GeneChip® S.cerevisiae Tiling 1.0R arrays from Affymetrix Inc. (Wycombe. United Kingdom) were used and processed in the GeneChip® System with Autoloader from Affymetrix Inc. (Wycombe, United Kingdom). Control Oligo B2 is included to provide alignment signals for image analysis. Image caption and

preliminary data analysis were carried out with Affymetrix® Expression Console™ software (v1.1).

Table 2. Oligos used during ChIP-on-chip sample preparations.

Oligo name	Sequence	Gene	Strand[a]	Position[b]
AVV220	AGAACTTGGCGATTGCTGACA	*ROX1*	C	-408
AVV221	AAGACCGTTACATTACGCAAAGTG	*ROX1*	W	-275
AVV222	CATACACATCGTGCTTAGCGATC	*IXR1*	W	-526
AVV223	CCCATTCGTTCTCTCACCAAG	*IXR1*	C	-376
ACC224	CATAAAGGGTCTCTTTCACCTATACG	*TIR1*	W	-273
AVV225	CTTCACTTTTTTCTCTGTCAAGGG	*TIR1*	C	-178
AVV226	TCAAACCATTTCCTGCGGAG	*HEM13*	C	-539
AVV227	TGCCTATGACGGTAATCCCA	*HEM13*	W	-406
Primer_A	GTTTCCCAGTCACGGTCNNNNNNNNN	-	-	-
Primer_B	GTTTCCCAGTCACGGTC	-	-	-

[a]W: Watson strand; C: Crick strand.
[b]Numbering is considering +1 for the adenine in the first start codon.

2.4.- Statistical data analysis and data mining

Transcriptomic array data were normalized and summarized using the RMA algorithm from Affymetrix. The data were analyzed using the web suite Babelomics (v4.3) (Medina *et al.*, 2010). Statistical analyses to identify DEGs were made by using the LIMMA (Linear models for microarray data) test (Smyth, 2005). The FDR (False Discovery Rate) was estimated to correct values for multiple comparisons (Benjamini & Hochberg, 1995). Statistical significant DEGs were considered those with a FDR<0.01 and a fold change ≥1.4 in the comparisons.

ChIP-on-chip raw data from Affymetrix GCOS software were analyzed using Affymetrix Tiling Analysis Software (TAS) v1.1.03 (http://www.affymetrix.com/support/developer/downloads/TilingArrayTools/index.affx), and the .BPMAP file *Sc03b_MR_v04.bpmap*. A two-sample analysis was conducted using specific chromatin immunoprecipitation (from Ixr1-(c-Myc) tagged samples) DNA samples as the 'treatment' group and three whole genome

fragmented and amplified DNA samples as the 'control' group for both untreated control and cisplatin treated cultures. Data were normalized using built-in quartile normalization and probe-level analysis with perfect match (PM) probes and run with a bandwidth of 250. Ixr1 protein occupancy profiles were visualized with Affymetrix Integrated Genome Browser (IGB). Interval analyses were done using TAS software with a minimum run of 10 bp and a maximum gap of 250 bp, and *p-value* cutoff of 0.01. Bed file conversions were done using UCSC (*University of California Santa Cruz*) tools (https://genome.ucsc.edu). Bed file analyses were done using ChIpSeek tools (http://chipseek.cgu.edu.tw) (Chen *et al*., 2014).

Gene descriptions and comparative analyses of lists from differentially expressed genes were obtained through Yeast Mine (http://yeastmine.yeastgenome.org/yeastmine).

Functional distribution of genes in the differentially regulated clusters was analyzed using FunSpec (http://funspec.ccbr. utoronto.ca/) developed by Robinson and coworkers (Robinson *et al*., 2002) and PANTHER (Thomas *et al*., 2006; Mi *et al*., 2013) (http://pantherdb.org). The MIPS Functional Catalogue Database (FunCatDB) was used in the analyses (http://mips.helmholtz-muenchen.de/proj/ funcatDB/). For these analyses a *p-value* lower than 0.01 was selected. These *p-values* represent the probability that the intersection of one given list of genes with any given functional category occurs by chance. In the report of the analyses carried out with FunSpec, *k* is the number of genes from the input cluster in given category and f is the total number of genes in given category. Hierarchical and k-means clustering was performed using Multiple Array Viewer package (MeV, v10.2), using the 'organize genes' option and default options of 'Euclidean distance' and 100 runs. Motif analysis was performed using HOMER (***H***ypergeometric ***O***ptimization of ***M***otif ***E***n***R***ichment) (Heinz *et al*., 2010) in the ChIPSeek suite (http://chipseek.cgu.edu.tw/index.py) (Chen *et al*., 2014), MEME (*Multiple Em for Motif Elicitation*) suite (http://meme.sdsc.edu/meme/), RSAT (*Regulatory Sequence*

Analysis Tools) (http://rsat.ccb.sickkids. ca/) (Bailey _et al._, 1994; van Helden _et al._, 2003) and YEASTRACT (_Yeast Search for Transcriptional Regulators And Consensus Tracking_) (http://www.yeastract.com) (Teixeira _et al._, 2006). Regulatory pathways were constructed by YEASTRACT and Cytoscape (http://www.cytoscape.org) (Smoot _et al._, 2011).

3.- RESULTS AND DISCUSSION

3.1.- Experimental design

The experimental design proposed to achieve the objectives is summarized in the scheme of figure 1.
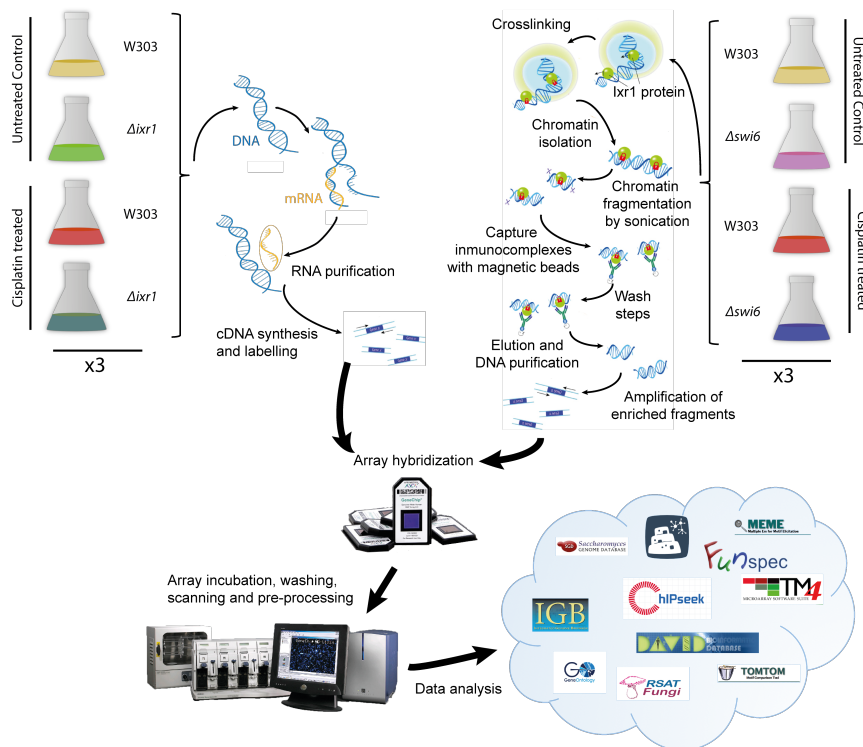
Figure 1. Schematic representation of strategies and procedures followed in the present study.

3.2.- Effects of *IXR1* deletion on the W303 transcriptome

Variations observed in the transcriptome in the *Δixr1* strain compared to the isogenic W303 strain, carrying the wild type *IXR1* allele, are evaluated as explained in Materials and Methods. After RMA normalization of raw data of triplicate biological repeats (figure 2a), the paired samples *Δixr1* mutant and isogenic W303 strains were analyzed using a Student's t-test with Bonferroni adjusted *p-value* and expressed in fold change values. The analysis revealed that 499 out of a total of *S. cerevisiae* 5744 probe sets in the GeneChip® Yeast-Genome-2.0 arrays were significantly changed (Figure 2b). There are 197 genes up-regulated (Table S1) and 302 down-regulated (Table S2) in the isogenic strain in which the *IXR1* gene has been deleted. Functional distribution of up-regulated genes analyzed with FUNSPEC (Table 3) shows that enriched functional groups include genes that take part in oxidation-reduction processes [GO:0055114], response to stress [GO:0006950], lipid metabolism ([GO:0006696], [GO:0008610], [GO:0006631]) or carbohydrate metabolism ([GO:0005978], [GO:0000025], [GO:0005975]). Most of the genes related to response to stress in Table 3 are genes expressed during anaerobic or hypoxic conditions.

Functional distribution of down regulated genes after *IXR1* deletion (Table 4) shows that enriched groups are those related to ribosome biogenesis ([GO:0042254], [GO:0006364], [GO:0000462], [GO:0006361], [GO:0000027], [GO:0000447]), and translation [GO:0006412], as well as other related to the metabolisms of amino acids [GO:0009085], membrane transport [GO:0055085] or ion channels [GO:0006810].

In a previous work, the effect of *IXR1* deletion was analyzed using a different *S. cerevisiae* strain, BY4741, and a different platform of arrays (Vizoso *et al*., 2012). Although in part attributable to the different growth media composition and performance of the platforms used in the two analyses, the observed

differences might also show that the expression and function of Ixr1 is highly dependent on the yeast strain analyzed as previously reported (McA'Nulty & Lippard, 1996; Castro *et al*., 2010a; Rodríguez Lombardero *et al*., 2012). Coincident up-regulated genes include those functionally related to oxidative and hypoxic stress. This is in accordance with the function previously attributed to Ixr1 as a repressor of hypoxic genes during normoxic conditions (Lambert *et al*., 1994; Bourdineaud *et al*., 2000; Castro *et al*., 2010a; Castro *et al*., 2010b; Vizoso *et al*., 2012) or its participation in the response to oxidative stress generated by H_2O_2 treatment (Castro *et al*., 2010a).
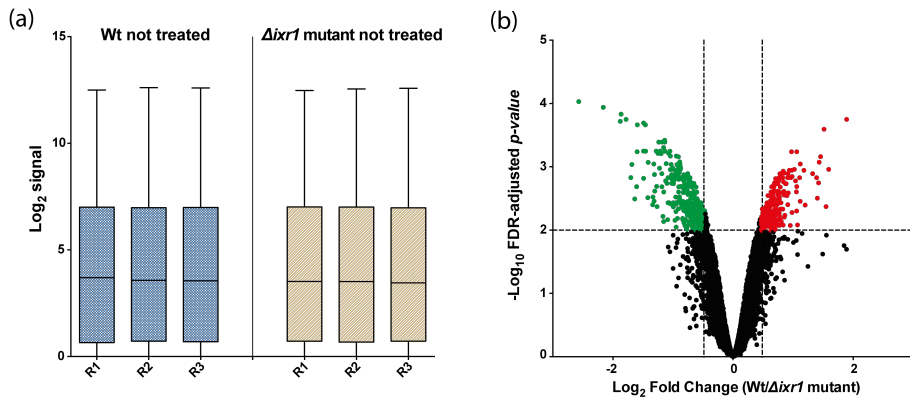
Figure 2. (a) Boxplots of untreated W303 (blue) and *Δixr1* (beige) strains after RMA normalization. Boxes represent the 25th to the 75th percentile of intensity values for the 5744 probe sets expressed in logarithm. R1, R2 and R3 correspond to the three biological replicates. (b) Volcano plot of *IXR1* deletion effect depicting individual probe *p-value* ($-\log_{10}$) *versus* expression fold change ($\log_2$). Dotted lines represent, in the logarithmic corresponding scale, the *p-value* selection threshold (<0.05) and the fold change cut-off threshold (>1.40). Up-regulated genes are represented in red and down-regulated genes in green.

Among the up-regulated DEGs after *IXR1* deletion, but only reported in the W303 genetic background, are those related to ergosterol biosynthesis (*ERG28, NCP1 MCR1, ERG5, ERG24* and *ERG10*). This could be explained due to differences already detected between the strain BY4741 and W303 (Davis & Rine, 2006). The expression of genes related to ergosterol biosynthesis is regulated by a mechanism dependent of the two homologous transcriptional factors Ecm22 and Upc2, which

bind to their target promoters during normoxia or hypoxia respectively (Davies & Rine, 2006). However, this regulatory mechanism is functional in the *S. cerevisiae* W303 strain but not in the BY4741 strain (Davies & Rine, 2006). Differences observed about Ixr1-dependent regulation of ERG genes in the two strains might be caused by the regulatory interplay between these factors, since aerobic and hypoxic expression of *IXR1* also depends alternatively on the transcriptional activators Ecm22 and Upc2 (Castro *et al.*, 2010a). Interestingly, among the new discovered targets of Ixr1 regulation, we found *ERG24*, whose mutation causes mitochondrial abnormality (Dimmer *et al.*, 2002) and decreased resistance to cisplatin (Liao *et al*. 2007).

Table 3. Functional gene groups over-represented among genes whose expression in the *Δixr1* mutant strain was higher than in the W303 strain

Category	P-value	In category from cluster	κ	f
Oxidation-reduction process [GO:0055114]	4.01×10^{-05}	*GDH3 ETR1 ADH7 NDE2 CTA1 YEL047C DSF1 NCP1 GRE3 GUT2 YIR035C MCR1 FOX2 XYL2 IDP2 ERG5 YIM1 ADH2 IDP3 ERG24 YNR073C YPL113C*	22	272
Ergosterol biosynthetic process [GO:0006696]	5.86×10^{-05}	*ERG28 NCP1 MCR1 ERG5 ERG24 ERG10*	6	23
Response to stress [GO:0006950]	7.94×10^{-05}	*PAU3 DFM1 TIR1 PAU5 WSC4 GRE3 TIR3 PAU15 PAU16 PAU19 HSP33 PAU21 GRE1 HSP32 PAU22*	15	152
Glycogen biosynthetic process [GO:0005978]	0.0003881	*GLG2 GLG1 GSY2 GAC1*	4	12
Lipid biosynthetic process [GO:0008610]	0.001116	*ETR1 ERG28 NCP1 PHS1 SFK1 ERG5 ERG24*	7	52
Fatty acid metabolic process [GO:0006631]	0.001607	*YAT1 FOX2 CAT2 CRC1 PIP2*	5	28
Carbohydrate metabolic process [GO:0005975]	0.002688	*MAL32 CTS2 PKP2 MAL12 PCL7 XYL2 CRR1 FBP1 YLR446W*	9	94

Table 4. Functional gene groups over-represented among genes whose expression in the *Δixr1* mutant strain was lower than in W303 strain

Category	*P-value*	In category from cluster	*κ*	*f*
Transmembrane transport [GO:0055085]	8.52 x 10^{-08}	FUI1 FLR1 YMC2 CTP1 PHO89 RGT2 GGC1 BAP3 CAN1 FCY2 NUP49 MUP1 MEP1 TPO2 TNA1 YHK8 HXT1 QDR2 FLX1 OPT1 YJR124C DAL5 NUP100 STE6 SUL2 PHO84 HXT2 VBA1 AQR1 TAT2 ENB1 TPO4 SAM3 MEP3 OPT2	35	303
Response to pheromone [GO:0019236]	0.0005437	FIG2 STE2 GPA1 BAR1 STE6 SST2 AGA1	7	34
Translation [GO:0006412]	0.0006239	EFB1 RPS16B RPS11A RPS18A RPL27B RPL23B RPL22B RPL30 RPS2 RPL26B MES1 RPS24B RPL16A MEF2 RPS22A RPL14A RPS27A RPL15A RPL22A NAM2 RPS29A RPL31B RPS18B BRX1 WRS1 RPL33B RPS23B	27	318
RNA methylation [GO:0001510]	0.0007692	NOP1 MRM1 TGS1	3	5
rRNA processing [GO:0006364]	0.0008012	SRD1 NOP1 SNM1 CGR1 RPL30 RPS2 NSR1 IPI1 DBP8 UTP18 DHR2 EMG1 UTP13 TSR2 IPI3 POP1 POP3 DBP6 NIP7	19	195
Transport [GO:0006810]	0.001557	FUI1 FLR1 YMC2 CTP1 PHO89 RGT2 GGC1 BRE4 BAP3 TRS85 TCA17 CAN1 FCY2 NUP49 MUP1 MEP1 TPO2 MTM1 TNA1 YHK8 HXT1 QDR2 FLX1 TOK1 ALB1 MRS3 OPT1 DAL5 SFT1 NUP100 LST4 STE6 FRE6 SUL2 GRX8 IKI3 PML39 PHO84 HXT2 VBA1 SFB2 AQR1 TAT2 ENB1 ORT1 PNS1 TPO4 SNF8 FLC1 SAM3 DSS4 MEP3 OPT2	53	815

Table 4. Continued

Category	*P-value*	In category from cluster	*κ*	*f*
Mitochondrial citrate transport [GO:0006843]	0.001896	*CTP1 YHM2*	2	2
Regulation of telomere maintenance via telomerase [GO:0032210]	0.001896	*CDC13 TGS1*	2	2
Ribosome biogenesis [GO:0042254]	0.002918	*REI1 NOP1 CGR1 RPS2 IPI1 DBP8 UTP18 ALB1 DHR2 EMG1 UTP13 SFP1 IPI3 DBP6 BRX1 NIP7*	16	170
Lysine biosynthetic process [GO:0009085]	0.003905	*LYS21 LYS20 LYS14*	3	8
Maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000462]	0.004218	*RPS16B RPS11A RPS24B DHR2 RPS27A UTP13 TSR2 RPS23B*	8	60
Transcription initiation from RNA polymerase I promoter [GO:0006361]	0.005524	*RRN6 RRN7*	2	3
Ribosomal large subunit assembly [GO:0000027]	0.00555	*RSA4 IPI1 IPI3 DBP6 BRX1 NIP7*	6	38
Nitrogen utilization [GO:0019740]	0.00567	*MEP1 DAL81 MEP3*	3	9
Endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000447]	0.007175	*RPS18A DBP8 UTP18 EMG1 UTP13 RPS18B*	6	40

Comparing data in Table 4 with those previously published (Vizoso *et al*., 2012), coincident down-regulated genes in the two analyses after *IXR1* deletion

include those functionally related to the metabolism of amino acids or to membrane transport. The most relevant feature is the association of several genes shown in Table 4 with cellular functions dealing with RNA processing and ribosome biogenesis, even among significantly (*p-value* <0.05) down-regulated genes with a fold change >-1.4 (data not shown, [GO:0042254]; *p-value* ≈3 x 10^{-07}). Noteworthy, it has been previously shown that cisplatin treatment also modulates these functions (Jordan & Carmo-Fonseca, 1998; Jin *et al*., 2008; Rodríguez Lombardero *et al*., 2014) and, considering that *IXR1* is a gene that increases sensibility to the drug, we further analyzed this feature.

Ribosome assembly is a major undertaking for cells and it is subject to stringent inspection, requiring a significant fraction of the resources devoted to macromolecular synthesis and trafficking. All three RNA polymerases participate in this process. RNA polymerase I, and to a lesser extent also RNA polymerase III, transcribe the rRNAs. Because production of ribosomes is so closely tied to the growth and proliferation of cells, deregulation of ribosome assembly has profound consequences on the health of organisms. Complete loss-of-function mutations in most assembly factors and ribosomal-proteins are lethal in yeast, and embryonic lethal in higher organisms (Woolford & Baserga, 2013). Table S3 contains several genes related to ribosome biogenesis and translation that are significantly repressed in the *Δixr1* mutant in comparison with the W303. These genes are distributed along all the sequential steps that occur in ribosome biogenesis, from pre-rRNA transcription to final ribosome assembly. Thus, from initial steps, in table S3 there are genes that encode several subunits of RNA polymerase I. Among the three RNA polymerases in eukaryotic cells, RNA polymerase I is the busiest. This enzyme transcribes the pre-rRNA that is processed to yield mature 18S, 5.8S, and 25S rRNAs. This represents 60% of the total cellular RNA transcription (French *et al*. 2003; Kos & Tollervey, 2010).

Transcription of the 35S primary transcript by RNA polymerase I occurs in

the cell nucleolus, that is formed around the ≈150 tandem repeats of the rDNA transcription unit found on chromosome XII in *S. cerevisiae*. The initial 6.6-kb pre-rRNA includes RNAs destined for both the SSU (18S) and the LSU (5.8S and 25S) rRNAs. Four general transcription factors or transcription factor complexes aid in the recruitment of RNA polymerase I to this promoter. They include UAS-binding upstream activity factor (UAF) (composed of Rrn5, Rrn9, Rrn10, and histones H3 and H4), TATA-binding protein (TBP), core factor (CF) (composed of Rrn6, Rrn7, and Rrn11 proteins and analogous to SL1 in mammals), and Rrn3 (TIF1A in mammals). Remarkably, *SPT15* that encodes TBP, *RRN6* and *RRN7* are present in Table 4.

There are also several genes in Table 4 related to 35S primary transcript pre-processing and formation of the final 40S and 60S subunits. This includes genes encoding RNA-methylases, like *TGS1*, *MRM1*, *TRM44*, *NOP1* and *MGE1*, or RNP-methylases like *HMT1*. *TGS1* encodes a trimethylguanosine synthase that causes hypermethylation of m(7)G to the m(2,2,7)G 5' cap of snRNAs, snoRNAs (Mouaikel *et al.*, 2002), it is necessary for ribosome biogenesis (Colau *et al.*, 2004) and also methylates the telomerase Tlc1 RNA (Franke *et al.*, 2008). Mrm1 is an rRNA(2'-O-ribose)-methyltransferase that, together to Mrm2, functions in mitochondrial 21S rRNA processing (Pintard *et al.*, 2002). *TRM44* encodes a tRNA(2'-O-ribose)-methyltransferases causing tRNA-modification (Motorin & Grosjean, 1999). The protein encoded by *EMG1* (Nep1 protein) methylates the hypermodified ψ1191 base of 18S rRNA and has an additional essential function during ribosome biogenesis (Eschrich *et al.*, 2002). Finally, the protein encoded by *HMT1* (Rmt1 protein) causes methylation of hnRNPs affecting their activity and nuclear export (Shen *et al.*, 1998), and it also methylates the U1 snRNP protein Snp1 and the ribosomal protein Rps2 (Lipson *et al.*, 2010). Besides, *IXR1* probably has a function in the control of the methionine/S-adenosylmethionine (AdoMet) ratio. At least this is the expected result of its control upon the two genes, *SAM4* encoding the S-adenosylmethionine-homocysteine methyltransferase and *MHT1*, which participates in the conversion of (AdoMet) to methionine (Thomas *et al.*, 2000). As

a whole, the presence of all these genes in table 4 indicates a deep impact of Ixr1 function in 40S and 60S ribosomal subunit maturation.

Transcriptional regulatory factors related to ribosome biogenesis are also down regulated in the W303 *Δixr1* strain. These are (table 4) *SRD1* encoding a zinc finger protein involved in the processing of pre-rRNA to mature rRNA (Badis *et al*., 2008), and *SFP1* that encodes for other zinc finger protein that regulates transcription of ribosomal proteins and other genes related to ribosome biogenesis (Marion *et al*., 2004), as well as responses to DNA-damage, nutrient availability and cell cycle progression (Xu & Norris, 1998). The protein Sfp1 binds DNA directly at highly active RP genes and probably indirectly through Rap1 protein at others (Gordan *et al*., 2009). Noteworthy, the mutation of Sfp1 causes a decrease in cisplatin resistance (Liao *et al*., 2007).

In accordance with all these effects of *IXR1* deletion upon RNA processing and ribosome biogenesis, a decrease of specific ribosomal proteins is also observed (GO term GO:0003735) in table 4.

3.3.- A transcriptome analysis of the function of Ixr1 in the response to cisplatin

Studies on the effect of cell exposure to cisplatin upon the *S. cerevisiae* transcriptome have been previously reported (Caba *et al*., 2005; Rodríguez Lombardero *et al*., 2014). Functional groups over-represented among up-regulated genes after cisplatin treatment are those related to sulfate assimilation and metabolism of sulfur containing compounds, as those related to the transport of sulfate or sulfur containing amino acids or those necessary for methionine, cysteine and S-adenosylmethionine biosynthesis. Also, several genes related to purine nucleotide biosynthesis are also up-expressed. The major functional group over-represented among down-regulated genes is related to ribosome biogenesis and include *RPS8A, RPS9B, RPS16B, RPS8B* and *RPS1B*, necessary for maturation of SSU-rRNA from tricistronic rRNA transcript, or *RPS17B, NSR1, RPS0A, RPS26A* and

RPS18B, participating in ribosomal small subunit assembly and/or rRNA export from the nucleus (Rodríguez Lombardero *et al*., 2014).

In order to identify genes and mechanisms involved in the resistance of *Δixr1* strain to cisplatin treatment, the changes in mRNA levels produced by addition of 600 μM cisplatin in *Δixr1* mutant cells grown in SD media were recorded, normalized and analyzed for statistical significance as described in the Matherial and Methods section (figure 3a). The analysis revealed that 195 out of a total of *S. cerevisiae* 5744 probe sets in the GeneChip® Yeast-Genome-2.0 arrays were significantly changed (Figure 3b). There are 96 genes up-regulated (Table S3) and 99 down-regulated (Table S4) in the *Δixr1* mutant strain after cisplatin treatment.

Functional distribution of up-regulated genes analyzed with FUNSPEC (Table 5) shows that enriched functional groups are those related to sulfate assimilation and metabolism of sulfur containing compounds; such as those related to cysteine and methionine biosynthetic processes ([GO:0019344] and [GO:0009086], respectively), sulfate assimilation [GO:0000103] or sulfate transport [GO: 0008272]. Among overrepresented functional groups there are also several related to transmembrane transport [GO:0055085] and oxidation-reduction processes [GO:0055114]. Besides, some genes related to purine nucleotide biosynthesis are up-regulated (*ADE1*, *ADE17, ADE13 and MTD1*), as shown in Table 5. These results are similar to those obtained from the W303 strain after cisplatin treatment (Rodríguez Lombardero *et al*., 2014). Nevertheless, the transcriptional activation of genes related to metabolism of sulfur compounds is stronger in the *Δixr1* strain than in the W303 strain, as indicates the figure 3c. Is interesting to note that, at difference with its W303 isogenic strain, *Δixr1* strain shows up-regulated several genes related to iron homeostasis [GO:0055072] and siderophore transport [GO:0015891], including *SIT1, ARN1, SMF3, FRE4 and FIT2.*

Table 5. Functional gene groups over-represented among genes whose expression in the *Δixr1* mutant strain treated with cisplatin was higher than in *Δixr1* mutant strain untreated.

Category	*p-value*	In Category from Cluster	*κ*	*f*
Cysteine biosynthetic process [GO:0019344]	$<1 \times 10^{-14}$	*CYS3 MET10 CYS4 MET28 MET3 MET5 MET14 MET17 MET16*	9	12
Methionine biosynthetic process [GO:0009086]	$<1 \times 10^{-14}$	*MET8 MET6 MET10 STR3 MET28 MET3 MET5 HOM6 MET14 MET1 YLL058W MHT1 MET17 ADI1 MET2 MET22 MET16*	17	31
Sulfate assimilation [GO:0000103]	2.14×10^{-13}	*MET8 MET10 MET3 MET5 MET14 MET1 MET22 MET16*	8	11
One-carbon metabolic process [GO:0006730]	5.98×10^{-10}	*GCV3 GCV1 SAM2 MTD1 SHM2 SAM1 GCV2*	7	15
Transmembrane transport [GO:0055085]	2.19×10^{-7}	*SEO1 SUL1 PHO89 BAP3 AGP3 TPN1 MUP1 VHT1 MUP3 HXT5 OPT1 MCH2 YCT1 MMP1 SUL2 YML018C YOL163W*	18	303
Oxidation-reduction process [GO:0055114]	1.29×10^{-6}	*MET8 HBN1 MXR1 AAD6 MET10 FMO1 SER33 MET5 HOM6 MTD1 ADI1 GCV2 ZWF1 FRE4 OYE3 MET16*	16	272
Siderophore transport [GO:0015891]	2.69×10^{-6}	*SIT1 ARN1 FRE4 FIT2*	4	8
Trans-sulfuration [GO:0019346]	1.14×10^{-5}	*CYS3 STR3 CYS4*	3	4
Iron ion homeostasis [GO:0055072]	2.87×10^{-5}	*SIT1 ARN1 SMF3 FRE4 FIT2*	5	26
Purine nucleotide biosynthetic process [GO:0006164]	0.0001054	*ADE1 MTD1 ADE13 ADE17*	4	18
Glutathione catabolic process [GO:0006751]	0.001206	*DUG2 DUG3*	2	4
Serine family amino acid biosynthetic process [GO:0009070]	0.001206	*SER33 SER1*	2	4
Folic acid-containing compound biosynthetic process [GO:0009396]	0.006905	*MTD1 FOL1*	2	9

Functional distribution of down-regulated genes analyzed with FUNSPEC (Table 6) shows that enriched functional groups include genes that take part in translation [GO:0006412], amino acid biosynthetic process [GO:0008652], glutamate biosynthetic process [GO:0006537] or nucleosome assembly [GO:0006334]. This last group includes genes encoding for histones. Interestingly, interplay between Ixr1 levels and histone dosage was previously described. Tsaponina and co-workers showed that when histone levels were high in *Δrad53* mutants because of a defect in histone degradation mechanisms, Ixr1 levels were low; meanwhile, the reduction of histone levels restored Ixr1 levels (Tsaponina *et al.*, 2011). These authors proposed a competition model in which high levels of histones displace Ixr1, or alternatively directly conducted Ixr1 to proteolysis.

A striking difference between *Δixr1* and its isogenic W303 strain is the lower impact of cisplatin treatment on the transcriptional repression of genes related to the ribosome biogenesis and translation, as shows figures 3c and 3d. This could be attributable to the feature that levels of expression are already low before cisplatin treatment due to the drastic effect of *IXR1* deletion on these genes, as previously described (table 4).

The comparison of bonferroni-corrected *p-values* of main significative gene ontologies obtained among down-regulated genes of W303 and *Δixr1* strains after cisplatin treatment shows that the effect of cisplatin is minimized by Ixr1 depletion (table 7), and the expression of most genes involved in the formation of the large and small ribosomal subunits remains unchanged after cisplatin treatment in the *Δixr1* strain .

Table 6. Functional gene groups over-represented among genes whose expression in the *Δixr1* mutant strain treated with cisplatin was lower than in *Δixr1* mutant strain untreated.

Category	p-value	In Category from Cluster	κ	f
Translation [GO:0006412]	1.61×10^{-6}	RPS9B SSB1 RPS8B RPL30 RPL24A RPL9A RPS26A RPS0A RPS20 RPL2B RPL15A RPL22A RPL31B RPS1B RPP2A RPS7A RPL5	17	318
Lysine biosynthetic process [GO:0009085]	2.47×10^{-6}	LYS20 LYS4 LYS12 LYS9	4	8
Propionate metabolic process [GO:0019541]	2.65×10^{-5}	PDR12 CIT3 PDH1	3	5
Maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000462]	0.0001802	RPS9B RPS8B PRP43 RPS20 FAF1 RPS1B	6	60
Cellular amino acid biosynthetic process [GO:0008652]	0.0004275	ARO4 LYS20 LYS4 LYS12 LYS9 LEU9 SAM4	7	98
Glutamate biosynthetic process [GO:0006537]	0.0006984	GDH3 CIT2 GDH1	3	13
GMP biosynthetic process [GO:0006177]	0.001156	IMD3 GUA1	2	4
Nucleosome assembly [GO:0006334]	0.001894	HTB2 HTA2 HTA1	3	18
Ammonium transport [GO:0015696]	0.002837	MEP2 ATO2	2	6
Allantoin catabolic process [GO:0000256]	0.003936	DUR1,2 DAL2	2	7
Phosphate metabolic process [GO:0006796]	0.0052	PHO3 PHO5	2	8
rRNA processing [GO:0006364]	0.00597	RPS9B SSB1 RPL30 PRP43 NSR1 RPS0A DBP2 RPS7A	8	195
Negative regulation of translation [GO:0017148]	0.008206	RPL30 ASC1	2	10
Regulation of translational fidelity [GO:0006450]	0.008206	SSB1 RPL31B	2	10

(a)

(b)

(c)

(d)

(e)
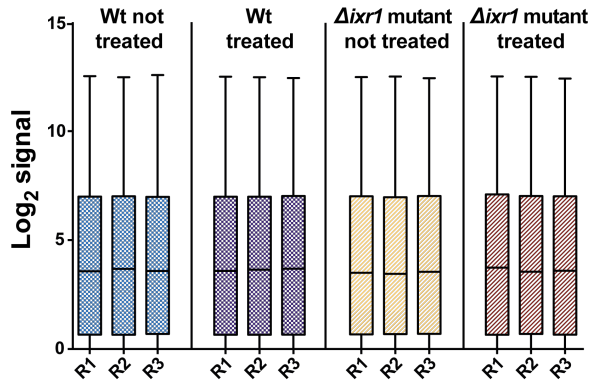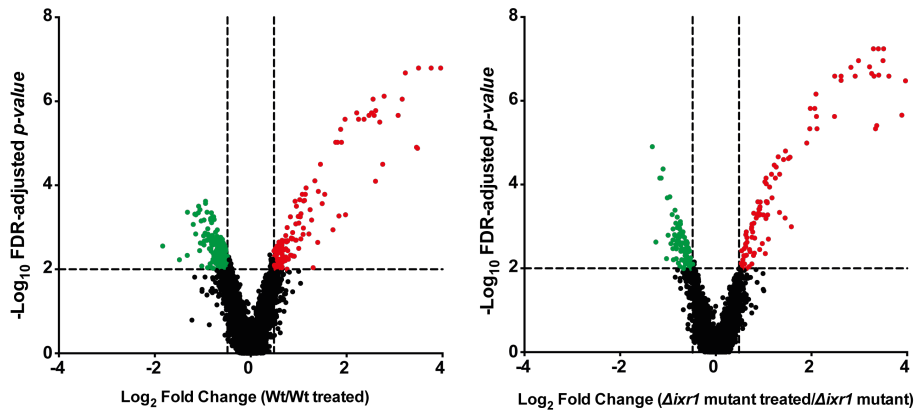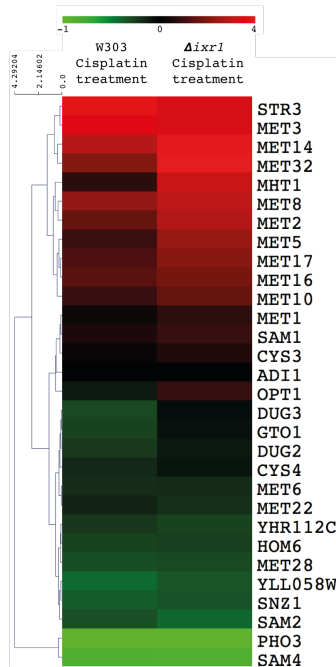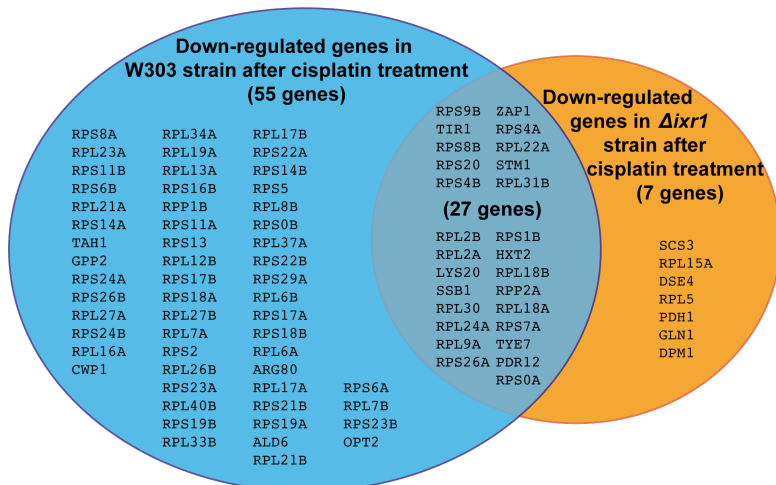
Figure 3. (a) Boxplots of untreated W303 (blue), W303 cisplatin-treated (violet), untreated *Δixr1* (beige) and *Δixr1* cisplatin-treated (red) strains after RMA normalization. Boxes represent the 25th to the 75th percentile of intensity values for the 5744 probe sets expressed in logarithm. R1, R2 and R3 correspond to the three biological replicates. (b,c) Volcano plots of cisplatin effect in W303 and *Δixr1* strains, depicting individual probe *p-value* ($-\log_{10}$) *versus* expression fold change ($\log_2$). Dotted lines represent, in the logarithmic corresponding scale, the *p-value* selection threshold (<0.05) and the fold change cut-off threshold (>1.40). Up-regulated genes are represented in red and down-regulated genes in green. (c) Heat map representing relative changes ($\log_2$) of expression of genes statistical significant related to sulfur metabolism [GO:0006790] in W303 after cisplatin treatment and compared with *Δixr1* strain treated. (d) Venn diagram of DEGs after cisplatin treatment related to translation [GO:0006412] that are dependent of Ixr1 (139 DEGs), showed in blue for W303 strain and in red in *Δixr1* strain (green). (e) Heat map representing relative changes ($\log_2$) of expression of genes statistical significant down-regulated in W303 after cisplatin treatment and compared with *Δixr1* strain treated. Genes signaled with the red bar are those whose repression after cisplatin treatment is minimized by *IXR1* deletion.

Category	GO_ID	GO_Term	Bonferroni-corrected *p-value*	
DNA/RNA	GO:0033279	Ribosomal subunit	6E-45	1E-07
	GO:0006412	Translation	3E-23	1E-04
	GO:0006417	Regulation of translation	2E-15	3E-06
	GO:0010608	Posttranscriptional regulation of gene expression	1E-14	7E-06
	GO:0042254	Ribosome biogenesis	7E-09	1E+00
	GO:0000462	Maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5,8S rRNA, LSU-rRNA)	7E-06	8E-01
	GO:0016072	rRNA metabolic process	3E-05	1E+00
	GO:0034470	ncRNA processing	5E-04	1E+00
	GO:0019843	rRNA binding	3E-05	6E-01
Localization	GO:0022626	Cytosolic ribosome	6E-57	3E-11
	GO:0015935	Small ribosomal subunit	2E-22	5E-03
	GO:0015934	Large ribosomal subunit	7E-18	1E-02
	GO:0005886	Plasma membrane	1E+00	1E+00

W303 down-regulated   Δixr1 down-regulated

>0.01 <0.01 <1E-04 <1E-10 <1E-16 <1E-21 <1E-30

GO bonferroni-corrected *p-values*

Table 7. Statistical GO significance comparison of W303 and *Δixr1* mutant responses to cisplatin treatment.

In a previous study we had tested the consequences of *SKY1* deletion, upon transcription in the same W303 genetic background (Rodríguez Lombardero *et al.,* 2012). Although both *Δixr1* and *Δsky1* null mutations increase cisplatin resistance, the identified DEGs comparing with W303 in absence of cisplatin show low gene overlap. The *Δixr1* and *Δsky1* observed variations only have in common 19 genes among the significant up-regulated genes (figure S1a) and 17 genes among the significant downregulated genes (figure S1b). Although *SKY1* deletion also produced a down-regulation of genes related to ribosome biogenesis, similarly as above reported for *IXR1* deletion, the affected genes are different in the two mutant strains. Furthermore, the comparison of fold change profiles of the different steps related to ribosome biogenesis and translation in W303, *Δixr1* and *Δsky1* strains after cisplatin treatment shows important regulation differences among them (figures S2a-l). For the majority of genes, repression by cisplatin is even higher in the *Δsky1* mutant than in the W303 strain (change towards green in

the heat map of the figure S2) and, for this reason, a function of Sky1 in mediating this repression was discarded (Rodríguez Lombardero *et al.,* 2012). The repression caused by cisplatin is however significantly diminished for several genes in the *Δixr1,* as shown in the heat map of figure 3d and also in the comparative analysis among wild type, *Δsky1* and *Δixr1* mutant strains in figure S2; in this last by a change towards red in the color scale that represents the ratio of expression in the treated *versus* untreated cells.

3.4.- Differentiating direct and indirect transcriptional responses to *IXR1* deletion by ChIP-on-chip analysis

In order to differentiate direct and indirect effects caused by Ixr1 upon transcription in the response of yeast cells to cisplatin, we then further analyse DNA interactions of Ixr1 by ChIP-on-chip experiments (see experimental procedure section). From control cultures, 1191 peaks were obtained (p-value < 0.01), 544 related to promoter regions (based on UCSC Genome Browser annotations) (Figure 4a and table S13). Analysing coincidences between genes found by ChIP-on-chip promoter-related-peaks and the whole set of up and down-regulated genes in the transcriptional analysis of the *Δixr1* strain *versus* wild type, the overlap was restricted to a very limited number of genes (figure 4b; table 9 and S3). This indicates that Ixr1 action upon transcription could be majorly mediated by direct mechanisms implying transient interactions or by other indirect mechanisms. It is interesting to note that Ixr1 binds directly to the promoters of several genes previously related to cisplatin sensibility (Birrell *et al*., 2002; Lee *et al*., 2005; Liao *et al*. 2007; Xia *et al*., 2007; Fedorov *et al*., 2010) such as *VMA8, GET2, RAD51, THR1, HMO1, RAD54, SOD1, MEC3, VRP1* or *TDA5* (table S13).

(a)

(b)

(c)
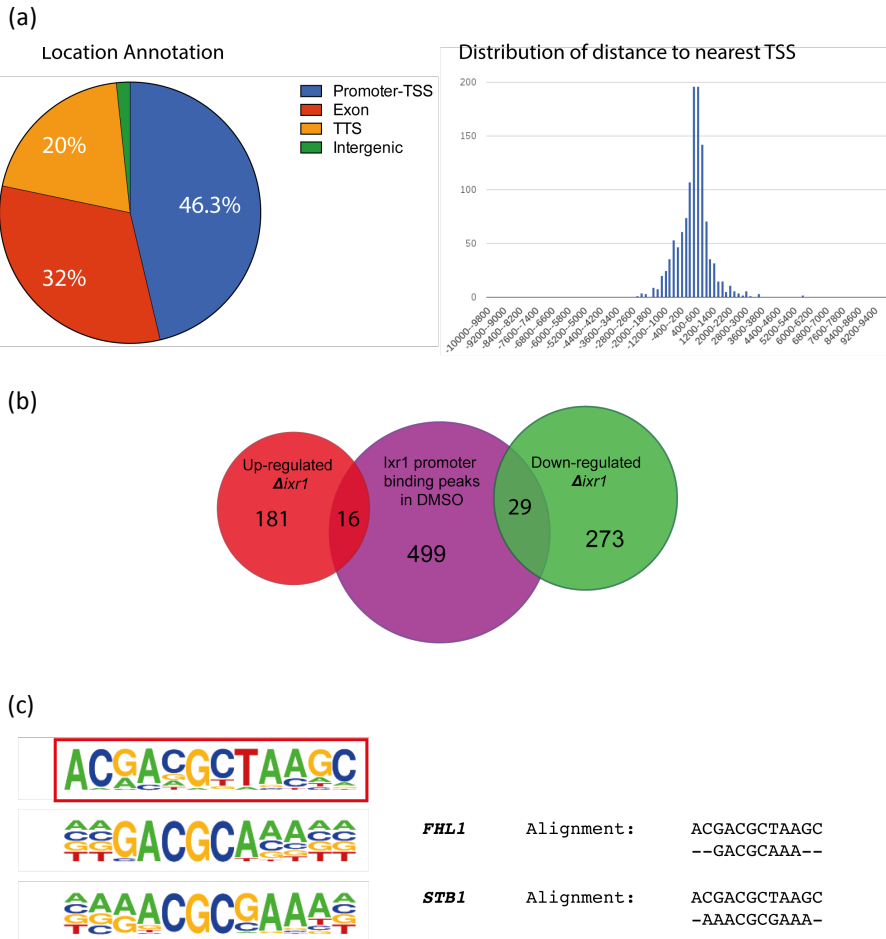
Figure 4. (a) Location of Ixr1 binding peaks after cisplatin treatment represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS). (b) Venn diagram of coincident up-regulated and down-regulated genes of *IXR1* deletion effect and genes with significant Ixr1 binding peaks in their promoter regions. (c) Logo representation of the *de novo* consensus sequence obtained (red rectangle) and aligned with their best-fit known consensus sequence.

We then analyzed the functional distribution of all genes with significant peaks in the promoter region with FUNSPEC (table 8) and enriched functional groups include genes that take part in transmembrane transport [GO:0055085] or response to stress [GO:0006950]. Among the genes related to transmembrane

transport, there are several targets transcriptionally regulated by Ixr1 (table 9, in bold) and related to nitrogen (*MEP1, DAL5*) and amino acid metabolism (*AVT6, ALP1, YMC2*), sulfate uptake (*SUL2*) and multidrug resistance (MDR) (*FLR1, TPO4*). Interestingly, both *MEP1* and *DAL5*, are implicated in the TORC1 signaling pathway through the GATA transcription activators Gln3 and Gat1/Nil1 (Georis *et al*., 2008; González *et al*., 2009).

De novo motif discovery with HOMER algorithm (**H**ypergeometric **O**ptimization of **M**otif En**R**ichment) (Heinz *et al*., 2010) in the ChipSeek suite (Chen *et al*., 2014) was also performed, setting a search restricted to ± 100 bp in each side of peak center. The consensus sequence **ACGACGCTAAGC** (*p-value* ≈1 x 10^{-12}) (figure 4c) was obtained in 29 peaks, that represent only the 5.25% of all the submitted peaks. The sequence was used to look for and compare with known motifs and, interestingly fits well with Fhl1 binding sequence (figure 4c). The forkhead factor Fhl1 binds directly to ribosomal protein (RP) genes actively transcribed, and indirectly through Rap1 motifs (Bar-Joseph *et al*., 2003; Lee *et al*., 2002). Its deletion causes a defect in 35S rRNA processing and slow growth (Hermann-Le Denmat *et al*., 1994).

Besides, Fhl1 interacts with Hmo1, a transcription factor regulating transcription mediated by RNA PolI (Ho *et al*., 2002; Ito *et al*., 2001). Fhl1, altogether with Ifh1 (coactivator) and Crf1 (corepressor), senses regulation of RP gene transcription via TOR and PKA signaling. Location of the repressor Crf1 in cytoplasm ensures the transcription of ribosomal protein genes. Crf1 is phosphorylated *in vivo* and *in vitro* by the TOR- and PKA- regulated kinase Yak1. Upon TOR inhibition, phosphorylated Crf1 rapidly translocates into the nucleus where it competes with Ifh1 for binding to Fhl1, which results in inhibition of RP transcription genes (Dietmar *et al*., 2004).

The *de novo* consensus sequence obtained also fits well with the

transcriptional activator Stb1 (figure 6c). Interestingly, Stb1 interacts with Swi6 to form the MBF (MCB-binding factor) complex with Swi4 and Mbp1 in a cell cycle-dependant manner to activate the transcription of genes required for DNA synthesis and G_1/S phase transition. Deletion of *STB1* caused an exacerbated delay in G_1 progression and the onset of start transcription in a *Δcln3* strain (Ho *et al.*, 1999; Costanzo *et al.*, 2003).

Table 8. Functional gene groups over-represented among genes with significant Ixr1-binding promoter peaks in the W303 strain untreated control.

Category	p-value	In Category from Cluster	κ	f
Peptidyl-lysine modification to hypusine [GO:0008612]	0.002077	HYP2 DYS1 LIA1	3	4
Transmembrane transport [GO:0055085]	0.005297	FLR1 YMC2 ADY2 MPH2 YEA6 SIT1 AVT6 ZRT1 VHT1 MEP1 ARN2 YHK8 HXT1 YIA6 YKE4 POR2 MIR1 DAL5 MPH3 JEN1 TRK2 MRS4 FPS1 GAL2 SUL2 SEC13 MMT1 YMR279C POR1 ALP1 VNX1 BIO5 NRT1 THI72 TPO4 YOR378W YPR011C AQY1	38	303
Response to stress [GO:0006950]	0.006189	PAU7 HSP26 TPS1 ARO4 BSD2 FMP45 HOR2 MDJ1 RIM15 TOS3 CTT1 WSC4 XBP1 POG1 KAR2 DAN4 YLR046C ICT1 FMP41 MCK1 IRA2 SSE1	22	152
Proteasome regulatory particle assembly [GO:0070682]	0.006357	RPT3 RPN14 RPT1 RPT4	4	10
Glycerophospholipid metabolic process [GO:0006650]	0.006751	PLB2 PLB1	2	2
Phospholipid biosynthetic process [GO:0008654]	0.009059	SCT1 SCS3 PCT1 OPI1 INO1 OPI3 ICT1 PAH1	8	37
Response to unfolded protein [GO:0006986]	0.009348	ULI1 KAR2 SCJ1 ZIM17	4	11
Maltose metabolic process [GO:0000023]	0.009348	MPH2 MAL13 IMA5 MPH3	4	11

Table 9. DEGs in *Δixr1 versus* wild-type with significant Ixr1 DNA-binding peaks in control cultures.

Gene	Change fold in *Δixr1*[a]	Distance to TSS[b]	Gene	Change fold in *Δixr1*[a]	Distance to TSS[b]
YAL064W	2.170	-1370	EFB1	-2.317	747
AVT6	1.436	446	**FLR1**	-1.942	-39
WSC4	1.626	165	**YMC2**	-1.853	555
YIR035C	1.773	-284	RSA4	-1.640	-608
ECM4	1.727	-749	GPP2	-1.671	304
CAF16	1.761	317	YHK8	-1.989	525
IGD1	1.915	-328	**HXT1**	-2.476	566
BNA6	1.480	434	YDL144C	-1.753	559
YAL064W-B	1.728	-3	RPA14	-1.695	108
SET4	1.799	-343	MTO1	-1.894	177
ALP1	1.663	469	**MEP1**	-2.379	284
PHM7	1.976	449	ATF2	-2.519	-896
CRC1	1.947	-158	**DAL5**	-3.027	-166
YPL067C	1.701	168	RPL15A	-1.604	356
SVS1	2.172	542	YLR046C	-1.443	-596
YPR015C	3.711	-629	ERG3	-1.781	469
			RPL22A	-1.501	-994
			SUL2	-1.515	-610
			TDA5	-2.063	485
			PLB2	-1.481	517
			NAT4	-1.981	-67
			POP3	-2.228	246
			PLB3	-1.690	-383
			AAD15	-1.572	-181
			HMS1	-1.447	-32
			TPO4	-1.517	-359
			SAM4	-1.843	-683
			RPL22B	-1.622	390
			YIL047C-A	-1.644	-55

[a]values in lineal scale. Data obtained from tables S3 and S4.
[b]*Translation Start Site*. Data obtained from table S14.
Genes related to transmembrane transport [GO:0055085] are highlight in bold.

3.5.- Differenciating direct and indirect roles of Ixr1 in the response to cisplatin treatment by ChIP-on-chip analysis

It has been proposed that Ixr1 shields cisplatin-modified DNA from nucleotide excision repair proteins, thus leading to higher cisplatin sensitivity in wild-type yeast strains (McA'Nulty and Lippard, 1996; McA'Nulty *et al*., 1996). Nevertheless, it was also proposed that a low level of constitutive genome integrity

check-point activation is responsible, at least in part, for the broad DNA damage tolerance seen in the *Δixr1* mutants, not only with cisplatin but also with other DNA-damaging drugs with different mechanisms of action such as 4-nitroquinoline 1-oxide (4-NQO), the alkylating agent methyl methanesulfonate (MMS) or the oxidizing agent tert-butyl hydroperoxide (t-BHP) (Tsaponina *et al.*, 2013).

Analysis of genomic localization of Ixr1 after cisplatin treatment by ChIP-on-chip rendered 457 peaks (*p-value* < 0.01), 315 related to promoter regions (Figure 5a) (table S14). Analysis of the functional distribution of all genes with significant peaks in the promoter region with FUNSPEC (Table 10) shows enriched functional groups similar to the obtained in the untreated control, including genes that take part in oxidation-reduction processes [GO:0055114], cell wall organization [GO:0007047] and transmembrane transport [GO:0055085] of spermidine [GO:0015848], carbohydrates [GO:0008643], acetate [GO:0006846] or water [GO:0006833]. The comparison of tables S13 (untreated control) and S14 (cisplatin treated) showed that both lists overlaps in 130 genes (figure 7b, table s15), and FUNSPEC analysis shows that are also principally enriched in genes related to response to stress [GO:0006950], cell wall organization [GO:0007047] and transmembrane transport [GO:0055085]. Among the 31 genes related to transmembrane transport, there are 15 targets that do not change with the presence of cisplatin, and that are related with the transport of carbohydrates (*MPH2, HXT1, MPH3, JEN1, FPS1, GAL2*) ions (*ADY2, ZRT1, VHT1, MEP1, YHK8, MMT1, TPO4*) and others (*THI72, AQY1*); and 16 targets that are specific to cisplatin treatment with functions related to transport of amino acids (*SAM3, PUT4, LYP1, AQR1, YCT1, BAP3*), nucleosides (*FUI1, FCY2*), ions (*FTR1, QDR2, PHO90, TPO1, MCH5*) and carbohydrates (*HXT3, HXT5, HXT7*).

Table 10. Functional gene groups over-represented among genes with significant Ixr1-binding promoter peaks in the W303 strain treated with Cisplatin

Category	*p-value*	In Category from Cluster	*κ*	*f*
Transmembrane transport [GO:0055085]	3.581 x 10⁻⁵	*FUI1 ADY2 MPH2 BAP3 HXT7 HXT3 FCY2 FTR1 ZRT1 VHT1 MEP1 YHK8 HXT1 HXT5 QDR2 PHO90 MPH3 JEN1 TPO1 FPS1 YCT1 GAL2 MMT1 AQR1 LYP1 THI72 TPO4 MCH5 PUT4 SAM3 AQY1*	31	303
Cellular cell wall organization [GO:0007047]	0.000257	*UTR2 NAG1 CRH1 SCW4 WSC4 PIR3 ECM4 EXG1 FKS3 GAS1 CHS1 HPF1 ECM23*	13	89
Cellular response to water deprivation [GO:0042631]	0.0004073	*CTT1 YJL144W YNL190W*	3	4
Maltose metabolic process [GO:0000023]	0.001254	*MPH2 MAL13 IMA5 MPH3*	4	11
Carbohydrate metabolic process [GO:0005975]	0.001504	*ROT2 GLK1 UTR2 MIG2 HXK2 CRH1 SCW4 YHR210C IMA5 PGM1 EXG1 GAS1*	12	94
Oxygen transport [GO:0015671]	0.00224	*YHB1 YNL234W*	2	2
Regulation of sporulation [GO:0043937]	0.00224	*PTP3 PTP2*	2	2
Acetate transport [GO:0006846]	0.00224	*ADY2 FPS1*	2	2
Cell wall chitin metabolic process [GO:0006037]	0.00224	*UTR2 CRH1*	2	2
Oxidation-reduction process [GO:0055114]	0.00229	*BDH1 BDH2 YDL124W HEM13 RNR1 SER3 CTT1 ERG1 RNR4 YHB1 YIR035C SOD1 SDH2 AHP1 PUT1 HMX1 HMG1 ADI1 ADH3 NDE1 YMR315W FRE4 ADH1 OYE3*	24	272
Glucose transport [GO:0015758]	0.003202	*MTH1 HXT1 HXT5*	3	7
Water transport [GO:0006833]	0.006509	*AQY2 AQY1*	2	3
Glucose import [GO:0046323]	0.006509	*GLK1 HXK2*	2	3
Asparagine biosynthetic process [GO:0006529]	0.006509	*ASN2 ASN1*	2	3
Proline catabolic process [GO:0006562]	0.006509	*PUT1 PUT4*	2	3
Chromatin silencing [GO:0006342]	0.007216	*MSI1 GAS1 HST3 HST2*	4	17
NADH oxidation [GO:0006116]	0.009871	*ADH3 NDE1 ADH1*	3	10
Nucleobase transport [GO:0015851]	0.009871	*FUI1 FCY2 THI72*	3	10

On the other hand, among the 184 Ixr1 binding peaks exclusively related to the cisplatin treatment (table S16) there are several genes related to stress response like *YRO2, SED1, CPR1, YHB1, YJL144W, SSC1, HMS2, MSN4, PIR3, HOR7, YNL190W, DDR2* or *PTP2*. There are also several genes related to G-protein mediated signal transduction (*GPA2, RGS2, STE4*) and with the drug/toxin transport (*QDR2, TPO1, AQR1, PDR5, PDR12*). Interestingly, QDR are multidrug-H$^+$-antiporters, which belong to the major facilitator superfamily (MFS). They were previously described to be involved in cisplatin resistance, showing up-regulation after cisplatin treatment and cisplatin sensitivity in null mutants (Tenreiro *et al.*, 2005). Besides *QDR2*, there are also two genes previously related to cisplatin resistance: *TRP1*, related to tryptophan biosynthesis; and *FCY2*, a purine and cytosine permease, whose expression is increased upon DNA replication stress; both null mutants confer cisplatin resistance (Huang *et al.*, 2005).

We then made *de novo* motif discovery with HOMER algorithm (Heinz *et al.*, 2010) in the ChipSeek suite (Chen *et al.*, 2014). There were obtained the consensus sequences **TCCGCGCG** and **GTGCCTGCRA** (figure 7c), obtained in 117 peaks (37% of all the submitted peaks; *p-value* ≈1 x 10^{-14}) and in 47 peaks (14.75% of all the submitted peaks; *p-value* ≈1 x 10^{-13}), respectively. Interestingly, consensus sequence 1 (figure 7c) fits well with the consensus of Rsc30 binding; This is a protein with a zinc cluster domain that binds DNA in a sequence specific manner to regulate, along with Scs3, the activity of the RSC chromatin remodelling complex. It has been reported that RSC function affects the transcription of ribosomal protein genes and genes involved in the integrity of the cell wall (Angus-Hill *et al.*, 2001). Consensus sequence 1 also fits well with the Sut1 binding site. Sut1 is a transcription factor involved in the activation of genes related to sterol uptake under anaerobic conditions and that is caused by release of the Cyc8(Ssn6)-Tup1 co-repressor from the target promoters (Regnacq *et al.*, 2001).

167

(a)

(b)

(c)
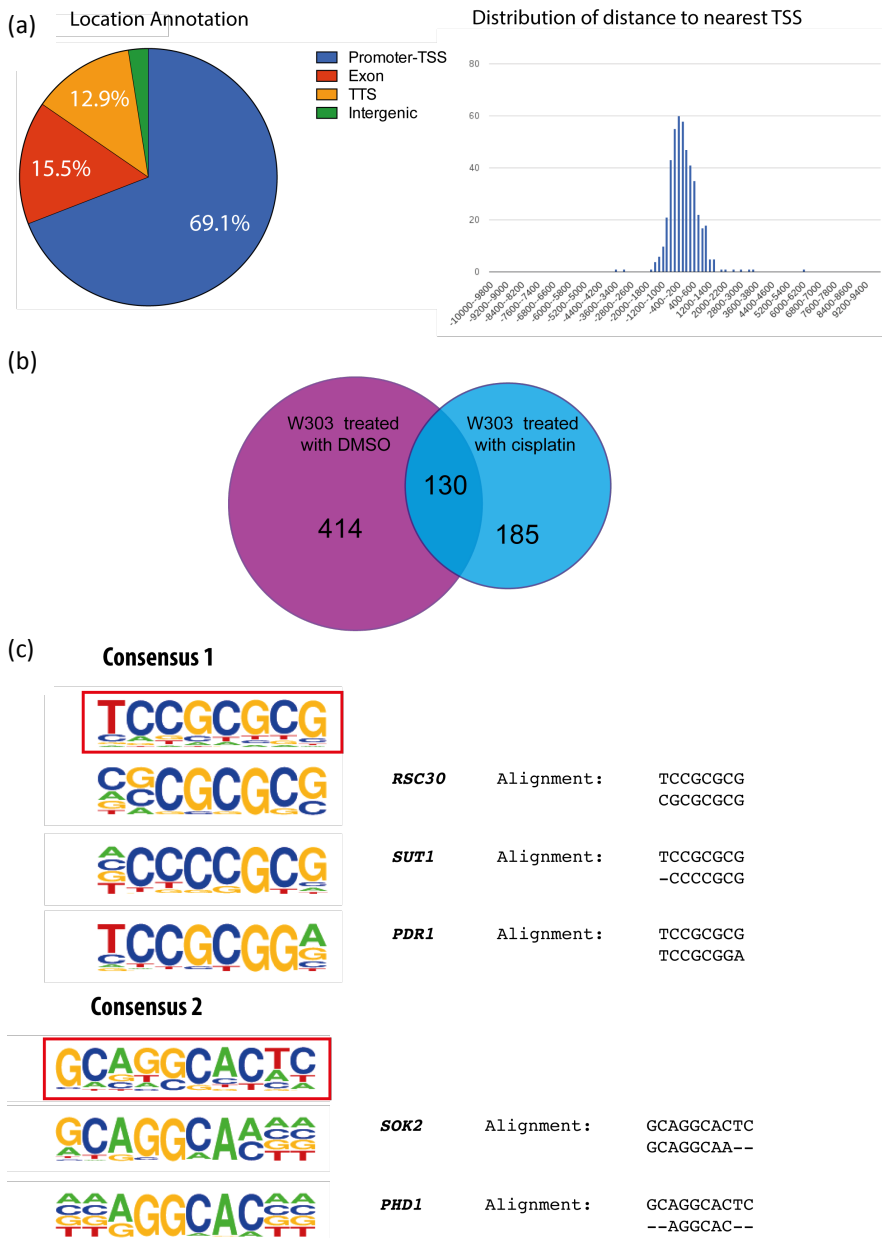
Figure 5. (a) Location of Ixr1 binding peaks after cisplatin treatment represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS). (b) Logo representation of the *de novo* consensus sequences obtained (red rectangles) and aligned with their best-fit known consensus sequence.

On the other hand, consensus sequence 2 (Figure 7c) fits well with binding of the paralog proteins Sok2 and Phd1. Sok2 is involved in the cAMP-dependent protein kinase signaling to negatively regulate yeast pseudo-hyphal differentiation (Pan and Heitman, 2000). Contrary, Phd1 acts as a transcriptional activator that enhances pseudo-hyphal growth (Gimeno and Fink., 1994).

3.6.- Effects of *SWI6* deletion upon Ixr1 DNA binding

Swi6 is a transcriptional cofactor that forms part of SBF and MBF complexes, each containing a distinct DNA-binding subunit, Swi4 and Mbp1, respectively. Each heterodimeric complex binds to a specific DNA sequence found in the promoters of G_1-specific genes. Although some genes are influenced by both factors, the regulation of most genes depends upon one of the two factors, the identity of which is correlated with the frequency of the specific binding motif (Spellman *et al*., 1998; Lyer *et al*., 2007; Simon *et al*., 2001). SBF regulates genes encoding proteins involved in the initiation of the G_1-S transition and progression into S phase, whereas MBF is primarily involved in the regulation of genes involved in DNA replication and repair. SBF and MBF are bound to G_1-specific promoters prior to G_1-specific transcription activation (Koch *et al*., 1993; Harrington *et al*., 1996; Cosma *et al*., 2001). The Swi6 subunit is present in the nucleus and only binds DNA from early G_1 through early S phase. SBF is inactive as a transcriptional activator during much of this time owing to interactions with the Whi5 and Stb1 proteins that recruit the Rpd3(L) histone deacetylase. The cyclin-dependent kinase CdC28 relieves this inhibition as cells progress past START, the commitment point for the G_1/S transition (Stillman, 2013).

The absence of Swi6 causes slow growth, it has large abnormally-shaped cells and cell cycle progression is delayed at G_1 and the G_2/M transition. It is unable to utilize various nitrogen sources and shows impaired fermentation (Chiu *et al*.,

2011; VanderSluis *et al*., 2014). It is also sensitive to metals, several DNA-damaging agents and oxidative stress (Hartman & Tippery, 2004; Jiang *et al*., 2014).

There are several evidences that indicate that Ixr1 could participate along with Swi6 in transcriptional regulation. First, we have known previously that Ixr1 physically interacts *in vivo* with the transcriptional activator Swi6, based on observations in our lab using the two-hybrid system (unpublished data, not shown). Second, Stb1 is a DNA binding protein that physically interacts with Swi6 (Ho *et al*., 1999) and the consensus sequence **ACGACGCTAAGC** obtained in the previously described ChIP-on-chip analysis using Ixr1 immunoprecipitation (this work, figure 4c) is similar to Stb1 binding site, thus correlating the presence of Ixr1, Swi6 and Stb1 to common regulated promoter regions. Third, deletion of Ixr1 or Mbp1 (both Swi6 interactors) produces a deregulation in the expression of *RNR1* gene, which encodes the major subunit of the large ribonucleotide-diphosphate reductase in the dNTP synthesis pathway (Raithatha & Stuart, 2005; Tsaponina *et al*., 2011). Accordingly, we have obtained binding-peaks for Ixr1 in *RNR1* promoter region after cisplatin treatment (table S17), as expected considering that *RNR1* has been previously described as a target of the MBF complex (Lyer *et al*., 2001). Besides, Ixr1 shares five genes with MBF complex after untreated control (*RAD51, APT2, YDR442W, POP3, NRM1*) and other five genes after cisplatin treatment (*RNR1, YDR442W, EXG1, FKS3, NRM1*). In the same way, analyzing the 131 peaks for binding of the SBF complex obtained by Lyer and co-workers (Lyer *et al*., 2001), we observed that Ixr1 shares up to 14 genes after untreated control, or up to 25 other genes after cisplatin treatments respectively. It is interesting to note that among the 25 shared genes, there are 8 related to the cell wall organization or biogenesis [GO:0071554] (*SED1, UTR2, CRH1, SCW4, CWP1, YPS3, GAS1* and *FLC1*); cell wall organization is a feature intimately related to cell cycle and under the control of the SBF complex, which regulate up to 19 genes related to cell wall (Lyer *et al*., 2001).

(a)

(b)
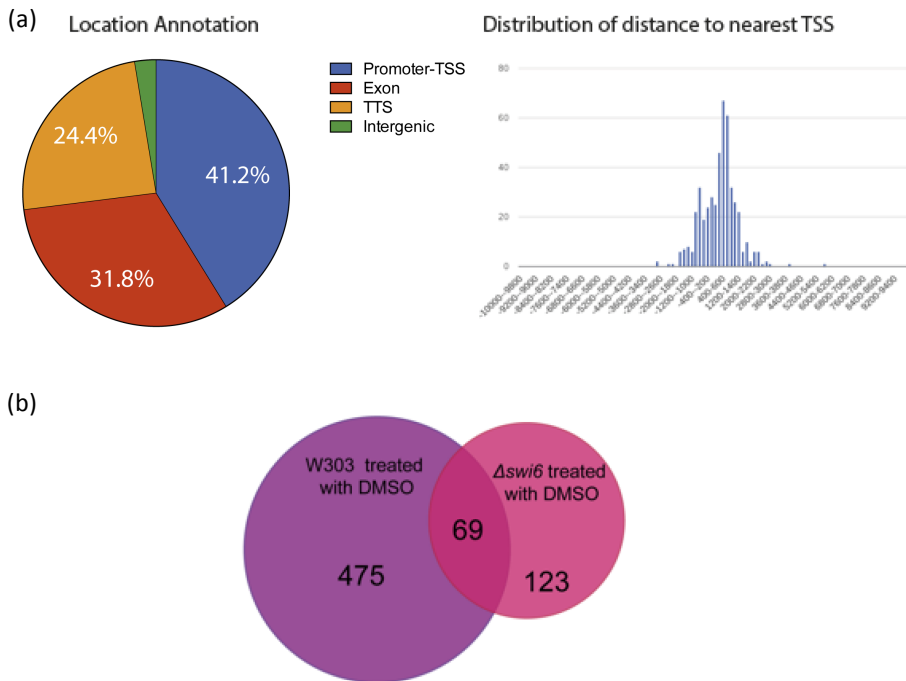
Figure 6. (a) Location of Ixr1 binding peaks in *Δswi6* strain after control treatment represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS). (b) Venn diagram of genes with significant Ixr1-binding promoter peaks that overlaps between W303 and *Δswi6* strain control treatments. (c) Logo representation of the *de novo* consensus sequence obtained (red rectangle) and aligned with their best-fit known consensus sequence.

To further analyze the transcriptional regulatory relationship between Ixr1 and Swi6, ChIP-on-chip experiments of Ixr1 in a *Δswi6* strain untreated (control) were performed. Analysis of genomic localization of Ixr1 by ChIP-on-chip rendered 471 peaks (p-value < 0.01), 194 related to promoter regions, that represents only 41% of the total peaks obtained (figure 6a) (table S17). Analysis of the functional distribution of all genes with significant peaks in the promoter region with FUNSPEC (table 11) revealed few enriched functional groups. Overlapping analysis between the lists of genes with Ixr1 binding peaks in their promoter regions in W303 and *Δswi6* strains untreated controls shows 69 coincident genes, indicating

Ixr1 binding that is independent of Swi6 function (figure 6b). Up to 12 of these genes are transmembrane proteins related to sugar transport (*MPH2, HXT1, MPH3, GAL2*), transport and homeostasis of metal ions (*ADY2, ZRT1, YHK8, MMT1, FRE4, SOD1, AHP1*) and others (*MEP1, YMR279C, THI72*). Interestingly, *TOS3* is also included and it encodes a protein kinase that phosphorylates and activates Snf1, an AMP-activated protein kinase that play key roles in the cellular response to nutrient stress (Nath *et al*., 2003). Snf1 activation is required for utilization of non-fermentable and non-preferred carbon sources, such as raffinose and glycerol-ethanol. *TOS3* (target of SBF) was first identified by a genomic screen for promoters that are bound by the SBF complex, which regulates transcription during the cell cycle (Lyer *et al*., 2001).

Table 11. Functional gene groups over-represented among genes with significant Ixr1-binding promoter peaks in the *Δswi6* strain untreated control.

Category	*p-value*	In Category from Cluster	κ	f
Pyrimidine nucleotide biosynthetic process [GO:0006221]	0.0001871	*URA3 DCD1 URA4 URA10*	4	11
Nitrogen utilization [GO:0019740]	0.001733	*ADY2 UGA1 MEP1*	3	9
Maltose metabolic process [GO:0000023]	0.003263	*MPH2 IMA5 MPH3*	3	11

Most of the Ixr1 target genes in common with SBF and MFB complexes are lost after *SWI6* deletion, with the exception of *TOS3*. It is interesting to note that Ixr1 binding to the promoter region of up to 37 positive transcriptional regulators (table S17) are lost in absence of Swi6. The functional distribution of these genes is related with the regulation of translation (*PET494, CAM1, HYP2, HST2, PHO4, MSS51, NAM8, STB1, FZF1*), response to nutrient levels (*RAS2, GAL4*), G_1/S cell cycle transition (*RPD3, BCK2*), proteolysis (*RPT3, RPT1, RPT4, SEC13,*

KAE1), chromosome organization (*ARP7, HOS1, NAP1, KAE1, BUD32*) and others
(*HAP3, CKS1, UTP5, EMI2, EDC1, OPI1, RIM101, SEC13, LEU3, PAH1, HOT1, RTG1,
SDH5, MED7*).

3.7.- Effects of *SWI6* deletion upon Ixr1 DNA binding after cisplatin treatment

We then mapped the Ixr1 binding sites in the genome of *Δswi6* strain after
cisplatin treatment. Analysis of genomic localization of Ixr1 by ChIP-on-chip
rendered 639 peaks (p-value < 0.01), 340 related to promoter regions, that
represents 53.5% of the total peaks obtained (Figure 7a; table S18). Analysis of the
functional distribution of all genes with significant peaks in the promoter regions
with FUNSPEC (Table 12) shows enriched functional groups of genes that take part
in the cell wall organization [GO:0007047], oxidation-reduction processes
[GO:0055114], carbohydrate metabolism [GO:0005975] and transport of
carbohydrates [GO:0008643], acetate [GO:0006846], oxygen [GO:0015671], water
[GO:0006833] and others (*FUI1, SSH1, ENA1, BAP3, NIC96, ZRT1, VHT1, MEP1,
TIM13, YHK8, QDR2, HOT13, MIA40, JEN1, NUP2, MMT1, AQR1, LYP1, TPO4, MCH5,
SAM3*).

Overlapping analysis between the list of genes with Ixr1 binding peaks in
their promoter regions in W303 and *Δswi6* strains after cisplatin treatment shows
166 coincident genes (figure 7b). Once again, up to 23 of these genes are
transmembrane transporters related to transmembrane transport (see table 8),
indicating the participation of Ixr1 in the expression of genes related to membrane
and cell wall in a way that is independent of Swi6 function or cisplatin treatment.
Also, it is important to note that also appears *TOS3*, an activator of the Snf1 kinase
complex.

In conditions of cisplatin treatment, in a similar way of previously
described in absence of the drug, deletion of *SWI6* produces Ixr1 displacement
from the promoter region of a large number of genes, 148, including several genes

related to oxidative stress response (*CTT1, GSH1, MSN4, IRA4*) or α-amino acid metabolism (*TRP1, ASN2, CYS4, DYS1, PUT1, YML082W, ARG80, LEU4, GLN1, ASN1*) among others (figure 7b and table 13).

Table 12. Functional gene groups over-represented among genes with significant Ixr1-binding promoter peaks in the *Δswi6* strain treated with cisplatin

Category	*p-value*	In Category from Cluster	κ	f
Cellular cell wall organization [GO:0007047]	4.08×10^{-05}	*KNH1 UTR2 CRH1 SCW4 PIR3 PIR1 MCD4 CCW12 FKS1 ECM19 FKS3 SUN4 CHS1 HPF1 ECM23*	15	89
Transmembrane transport [GO:0055085]	7.27×10^{-05}	*FUI1 SSH1 ADY2 MPH2 ENA1 BAP3 HXT7 HXT6 HXT3 NIC96 ZRT1 VHT1 MEP1 TIM13 YHK8 HXT5 QDR2 MPH3 HOT13 MIA40 JEN1 FPS1 GAL2 NUP2 HXT2 MMT1 AQR1 LYP1 TPO4 MCH5 SAM3 AQY1*	32	303
Oxidation-reduction process [GO:0055114]	0.001572	*ARA1 RIB7 ADH7 AAD3 YDL124W HEM13 RNR1 ALD5 SER3 ERG26 YGL039W YHB1 GRE3 SOD1 SDH2 AHP1 TSA1 HMG1 ADI1 ADH3 NDE1 YMR315W ZWF1 FRE4 DFR1 YPR127W*	26	272
Plasma membrane organization [GO:0007009]	0.002644	*HSP12 NCE102*	2	2
NADPH regeneration [GO:0006740]	0.004056	*PYC1 YMR315W ZWF1*	3	7
Glucose metabolic process [GO:0006006]	0.005483	*GLK1 TOS3 HXK2 PGM1 ZWF1*	5	23

(a)

(b)
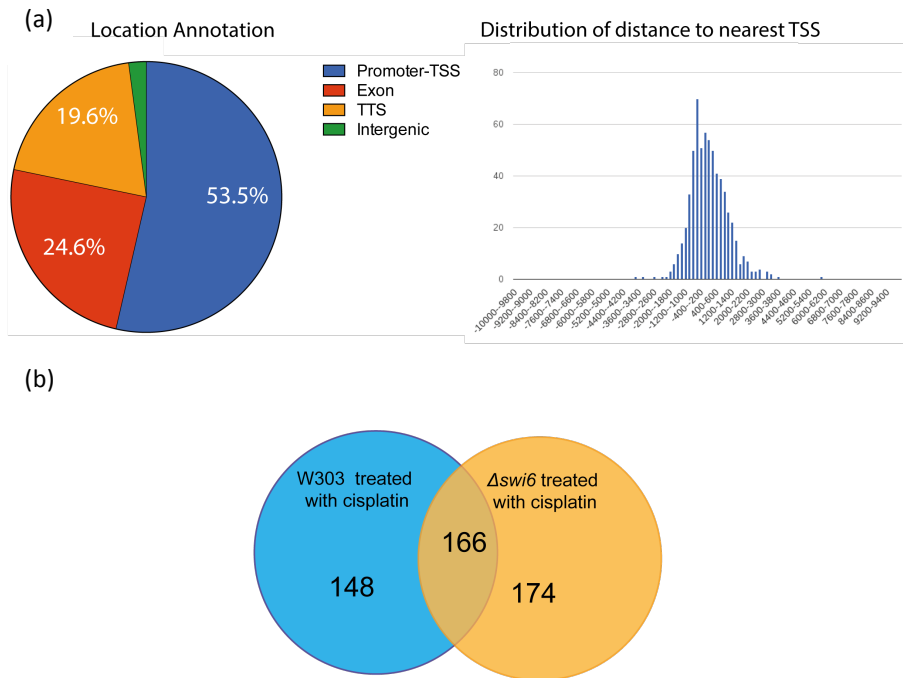
Figure 7. (a) Location of Ixr1 binding peaks in *Δswi6* stranin after cisplatin treatment represented by gene annotation pie chart and bar chart of peak distribution with respect to the transcription start site (TSS). (b) Venn diagram of genes with significant Ixr1-binding promoter peaks that overlaps between W303 and *Δswi6* strain treated with cisplatin.

3.8.- New clues about the function of Ixr1 in transcriptional regulation, nutrient signaling and in the response of yeast cells to cisplatin derived from this study

3.8.1. Ixr1 in transcriptional regulation under non stressed conditions

As derived from the transcriptome analyses presented in this chapter, during non-stressed conditions Ixr1 represses genes related to hypoxia and oxidative stress as already reported in individual (Lambert *et al.*, 1994; Bordineaud *et al.*, 2000; Castro-Prego *et al.*, 2010a; Castro-Prego *et al.*, 2010b) or wide-genome analysis (Vizoso-Vazquez *et al.*, 2012) and also other genes related to the

175

metabolism of lipids and carbohydrates (table 3). Besides, Ixr1 function is necessary for maintaining the normal levels of expression of genes related to ribosome biogenesis, amino acid metabolism as well as transmembrane transport and ion channels (table 4).

Table 13. Functional gene groups over-represented among genes with significant Ixr1-binding promoter peaks that overlaps between W303 and *Δswi6* strain treated with cisplatin

Category	*p-value*	In Category from Cluster	*κ*	*f*
Transmembrane transport [GO:0055085]	1.67×10^{-06}	*FUI1 ADY2 MPH2 BAP3 HXT7 HXT3 ZRT1 VHT1 MEP1 YHK8 HXT5 QDR2 MPH3 JEN1 FPS1 GAL2 MMT1 AQR1 LYP1 TPO4 MCH5 SAM3 AQY1*	23	303
Cellular cell wall organization [GO:0007047]	0.001688	*UTR2 CRH1 SCW4 PIR3 FKS3 CHS1 HPF1 ECM23*	8	89
Maltose metabolic process [GO:0000023]	0.002219	*MPH2 MAL13 MPH3*	3	11
Glucose metabolic process [GO:0006006]	0.002348	*GLK1 TOS3 HXK2 PGM1*	4	23
Transport [GO:0006810]	0.006276	*FUI1 ADY2 CDC48 MPH2 BAP3 HXT7 HXT3 ZRT1 VHT1 MEP1 YHK8 HXT5 QDR2 MPH3 JEN1 PTR2 SDH2 FPS1 AQY2 GAL2 MMT1 AQR1 LYP1 FRE4 CRC1 TPO4 MCH5 PDR12 FLC1 SAM3 NCE102 AQY1*	32	815
Oxidation-reduction process [GO:0055114]	0.008304	*YDL124W HEM13 RNR1 SER3 YHB1 SOD1 SDH2 AHP1 HMG1 ADI1 ADH3 NDE1 YMR315W FRE4*	14	272
Carbohydrate metabolic process [GO:0005975]	0.009158	*GLK1 UTR2 MIG2 HXK2 CRH1 SCW4 PGM1*	7	94

A comparison between the genes transcriptionally regulated by Ixr1 in non-stressed conditions, revealed through the transcriptome analysis, and confirmation of Ixr1 binding from ChIP-on-chip data (figure 6b) shows that the overlap is low (≈5%). This result is consistent with previous data reported in the literature across eukaryotes. The experimentally discovered binding sites for TFs are frequently only a small percentage (5-30%) of their actual regulated targets, as reported in yeast (Hughes & de Boer, 2013) and also in plants (Bolduc *et al*., 2012; Arenhart *et al*., 2014; Mönke *et al*., 2012; Para *et al*., 2014) and animals (Gorski *et al*., 2011; Bianco *et al*., 2014). Compassing to our description the nomenclature from Para and co-workers (Para *et al*., 2014), we will designate those showing Ixr1 binding and regulation as "stable targets"; those without Ixr1 binding but regulated as "transient targets"; and those with Ixr1 binding but not regulated as "poised targets". Besides, we will have also to consider possible examples of "indirect regulation" mediated by other secondary TF that is a "stable or transient target" of Ixr1 (Figure 8).

Among the "stable targets" (table 9) which are transcriptionally up-regulated, there are several genes related to cell wall organization (*WSC4, ECM4, SVS1*), amino acid transport (*ALP1, AVT6*) or the membrane transporter *CRC1*, required for the carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria during fatty acid beta-oxidation (van Roermund *et al*., 1999). On the other hand, among the "stable targets", which are transcriptionally down-regulated, are up to 14 genes encoding proteins intrinsic to membrane with different cellular functions (*ATF2, DAL5, MEP1, ERG3, YMC2, FLR1, HXT1, PLB2, PLB3, TPO4, SUL2, YLR046C, TDA5, YHK8*) or related to the ribosome biogenesis (*RPL15A, RPL22A, RPL22B, EFB1, RSA4*) (Figure 8a).

(a)

Stable binding with
direct regulation effect

Gene activation

[45 genes]

11.1%

33.3%

52.8%

2.8%

- cellular process (GO:0009987)
- metabolic process (GO:0008152)
- biological regulation (GO:0065007)
- localization (GO:0051179)

(b)

¿?

TF/Signal/Modification

lxr1

"Poised" stable binding

TARGET GENE

[499 genes]

0.3%
5.6%
4.7%
12.7%
10.3%
17.2%
47.9%

- cellular process (GO:0009987)
- metabolic process (GO:0008152)
- biological regulation (GO:0065007)
- localization (GO:0051179)
- cellular component organization or biogenesis (GO:0071840)
- developmental process (GO:0032502)
- response to stimulus (GO:0050896)

(c)

lxr1

"Transient" binding event

TARGET GENE

Gene activation

[57 genes]

6.7%
26.7%
3.3%
63.3%

- cellular process (GO:0009987)
- metabolic process (GO:0008152)
- biological regulation (GO:0065007)
- localization (GO:0051179)

(d)

Figure 8. Schematic representation of (a) "stable targets", (b) "poised targets", (c) "transient targets" and (d) indirect regulation by "transient targets" proposed based on the results obtained by transcriptomic and ChIP-on-chip experiments in the present work. Functional distribution of GO of different groups are showed (obtained by PANTHER).

Poised targets (Ixr1 binding without regulation) are among the most abundant in our analysis (≈50%) (Figure 8b). They might represent pre-charged pre-complexes at specific promoters waiting for a necessary TF/signal/modification not present in the analysed conditions as previously suggested (Para *et al.*, 2014). Besides, considering the peculiar characteristics of Ixr1, with two HMG-box domains, one predicted as SS and other NSS (Castro-Prego *et al.*, 2010b), they might be caused by unspecific binding of Ixr1 to DNA. However, Ixr1 binding to this poised targets is more stable than the produced with transient targets, which implies that other features are stabilizing the complex; This could be because a modified composition/structure of DNA produces a more tightened binding to Ixr1 or because other proteins in the chromatin harden these bonds. In this sense it is important to remark that the majority of binding events detected by ChIP-on-chip analyses are not observed after *SWI6* deletion (loosing ≈87%), including 37 genes

179

that encode the transcriptional activators described before (section 3.4 of this chapter).

Almost half (≈45%) of genes regulated by Ixr1 are "transient targets" (Figure 8c and 8d). This concept (as defined in Para *et al.*, 2014) assumes that the regulator, Ixr1 in our study, was able to interact with the regulated promoter, but the time course of the interaction was so fast that could not be captured by the ChIP event. Considering that Ixr1 is a HMGB protein, the high presence of "transient targets" in our Ixr1 analysis is consequent with the previous finding that HMG proteins are very dynamic in their interactions with DNA and quickly move along it (Gerlitz *et al.*, 2009; Stros, 2010). The interaction with Ixr1 will be however necessary to favor the binding of other TFs, co-regulators and/or the general transcriptional machinery in a "Hit-and-Run Model" for regulation, similarly to recently reported for bZIP1 that mediates nitrogen signaling in *Arabidopsis* (Para *et al.*, 2014). Indeed the hit-and-run model of transcription was first proposed by Schaffner in 1988 and it establishes that the TF organizes an efficient transcriptional complex, active for RNA polymerase II transcription, without the necessity of TF staying bound to the DNA for a long time (Schaffner, 1988). This model has the advantage that allows explaining a wide repertory of transcriptional activated events in different promoters that is orchestrated by a protein of very low expression, as is the case of Ixr1 under non-stressed conditions (Ghaemmaghami *et al.*, 2003). Further analyses for the identification of enriched sequences for binding of other TF in the promoters of these transient Ixr1 targets, positional analyses of the relationships between Ixr1 consensus and Ixr1-partners TF consensuses, as well as detection of interactions between Ixr1 and these TF partners will help to concrete this model.

In some cases direct binding of Ixr1 is not observed in the genes regulated in the transcriptome analysis, but Ixr1 regulates a transcriptional factor that in turn could explain the effect observed upon the target genes, called "Indirect

regulation" (Figure xd). In this sense, important general yeast regulators such as Sfp1 (Xu & Norris, 1998; Volkov *et al*., 2002; Marion *et al*., 2004; Cipollina *et al*., 2005), Abf1 (Morrow *et al*., 1989; Rhode *et al*., 1989; Della Seta *et al*., 1990; Buchman & Kornberg, 1990; Chambers *et al*., 1990; Gailus-Durner *et al*., 1996; Planta, 1997; Ozsarac *et al*., 1997; Reed *et al*., 1999; Beinoraviciūte-Kellner *et al*., 2005), Tec1 (Brückner *et al*., 2011), Sok2 (Pan & Heitman, 2000; Dastidar *et al*., 2012), Ume6 (Steber & Esposito, 1995; Elkhaimi *et al*., 2000; Williams *et al*., 2002) or Dal81 (Bricmont *et al*., 1991; Marzluf, 1997), as well as other specific regulators like Aca1, Hcm1, Rdr1, Rts2, Sut2, Tye7, Hfi1, Hms2, Srd1, Ada2, Gal80, Rgt1, Rds1, Nrg2 Or Tpo2, which are positively regulated by Ixr1, sum up to a total of 33 transcriptional regulators. Including the known targets of these TFs, which are effectively down- regulated in the transcriptome analysis carried with the *Δixr1* strain, the pool of secondary indirect targets of Ixr1 extends to 250 genes (Figure 8d). We have constructed the regulation network that connects these 250 genes (Figure 9). The major nodes are represented by yeast general regulators that, interestingly, are involved in the regulation of cellular growth and cell cycle progression in response to nutrient availability external stimuli or DNA damage. Thus, Sfp1 regulates response to nutrients and stress (Xu & Norris, 1998; Volkov *et al*., 2002; Marion *et al*., 2004; Cipollina *et al*., 2005); Abf1 is implicated in DNA replication and repair (Morrow *et al*., 1989; Rhode *et al*., 1989; Della Seta *et al*., 1990; Buchman & Kornberg, 1990; Chambers *et al*., 1990; Gailus-Durner *et al*., 1996; Planta, 1997; Ozsarac *et al*., 1997; Reed *et al*., 1999; Beinoraviciūte-Kellner *et al*., 2005); Tec1 links TORC1 and MAPK signaling pathways to coordinate control of cellular development in response to different stimuli (Brückner *et al*., 2011); Sok2 is a player in the PKA signal transduction pathway (Shenhar & Kassir, 2001; Malcher et a., 2011); Ume6 couples metabolic responses to nutritional cues with initiation and progression of meiosis (Steber & Esposito, 1995; Elkhaimi *et al*., 2000; Williams *et al*., 2002) and Dal81 is regulator of nitrogen degradation pathways (Bricmont *et al*., 1991; Marzluf, 1997) . Considering these data altogether, it seems that Ixr1

forms a complex regulatory network influencing strongly essential cell functions such RNA metabolism, ribosome biogenesis, translation, cell growth and cell cycle progression in response to nutrient availability, external stimuli or DNA damage.

(a)
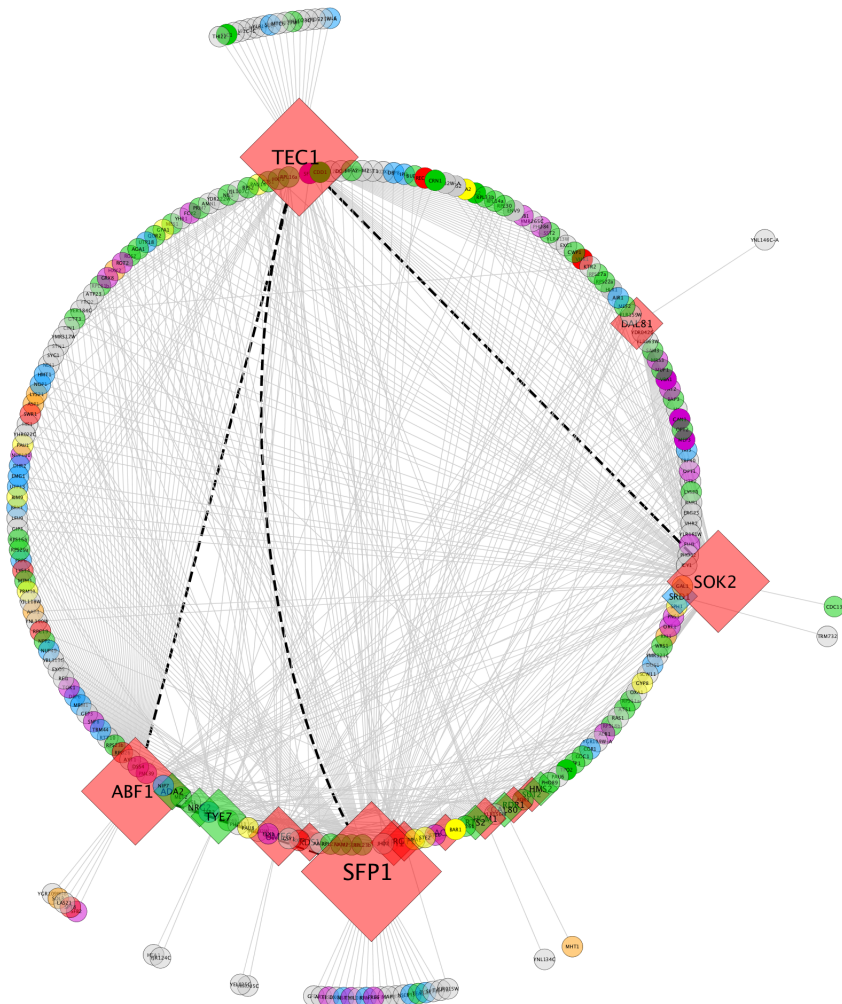
(b)
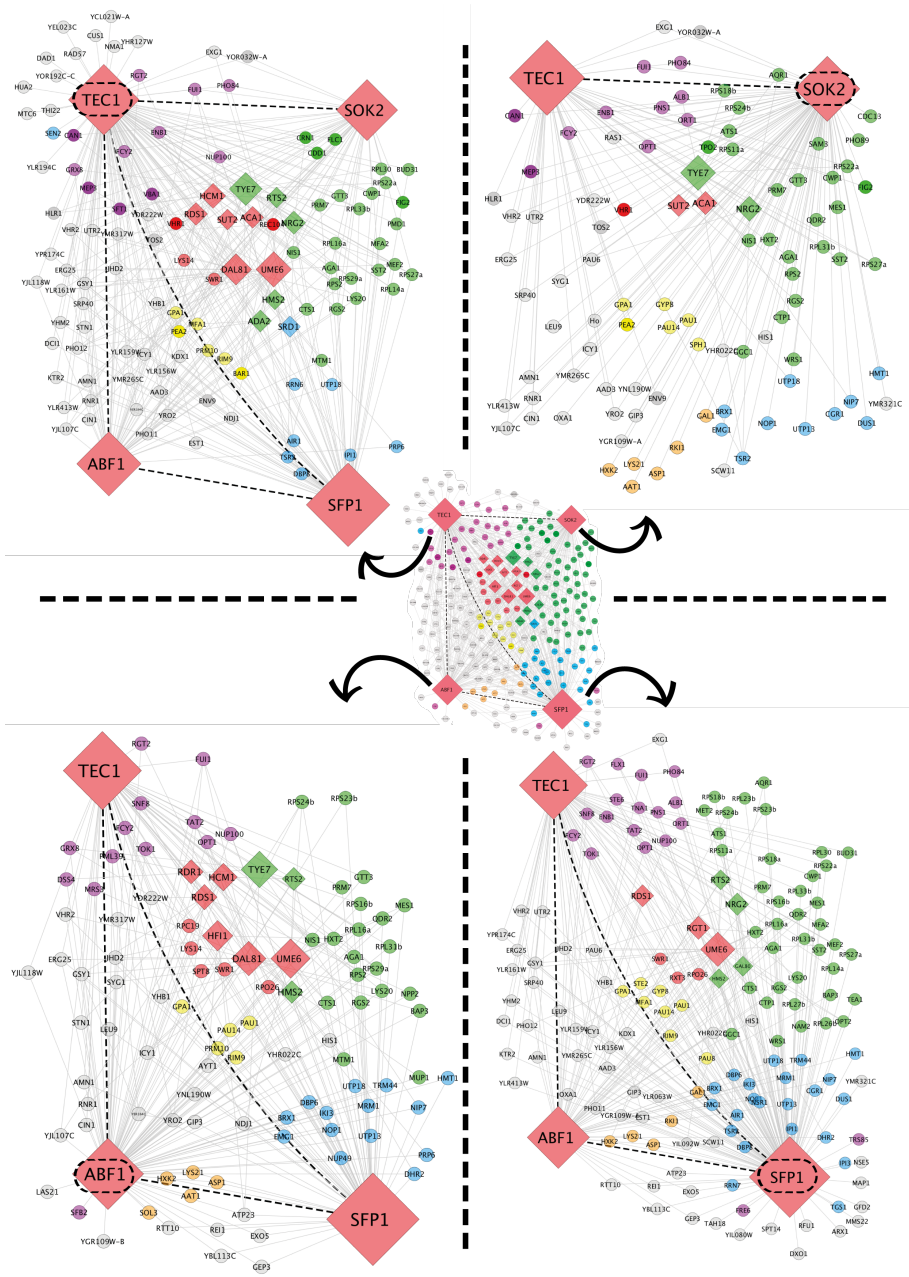
← **Figure 9.** Regulatory network showing the "regulator of regulators" role that Ixr1 has in the cell, represented in circular (a) or extended layouts (b) of the four major transcriptional regulators (Tec1, Sok2, Abf1 and Sfp1; indicated by dashed circles). Dark dashed lines indicate direct nodes between general regulators and rhombus correspond to genes that encode transcriptional regulators. Genes related to translation [GO:0006412] are coloured in green, ncRNA metabolism [GO:0034660] in blue, transcription [GO:0006351] in red, transport [GO:0006810] in violet, hexose metabolism [GO:0019318] in orange, sexual reproduction [GO:0019953] in yellow and others in grey.

3.8.2. Nutrient signaling: Is there a connection between Ixr1 and TOR signaling?

Transcriptome data obtained in this study reveal that deletion of *IXR1* affects expression of genes that may have a deep impact on 40S and 60S ribosomal subunit maturation (table 4), including two transcriptional regulators: Srd1, involved in rRNA maturation (Badis *et al*., 2008), and Sfp1, associated to transcription of ribosomal proteins and other genes related to ribosome biogenesis (Marion *et al*., 2004).

In *Saccharomyces cerevisiae* transcription is the major level of regulation of ribosome biogenesis and involves all three nuclear RNA polymerases. Pol I transcribes rDNA encoding the 35S rRNA precursor; Pol II transcribes the ribosomal protein (RP) genes and the non-ribosomal proteins necessary for ribosome biogenesis (RiBi) genes; finally Pol III produces 5S rRNA and tRNA. Approximately 90% of total cellular transcription is used for ribosome biogenesis in a rapidly growing cell (Warner *et al*., 2001). By other hand, the regulation of RP genes respond to many stress conditions including heat, oxidative stress, high osmolarity and DNA damage (Warner, 1999).

In yeast and pluri-cellular eukaryotes, the target of rapamycin (TOR) pathway is a signaling pathway that promotes anabolic processes necessary for cell growth (increasing amounts of proteins are needed for growth) and proliferation (cell division). Thus TOR activation promotes ribosomal biogenesis, while suppressing other catabolic processes as autophagy (Loewith *et al*., 2011).

Although TOR signals through two effector branches, the TOR complexes 1 and 2 (TORC1 and TORC2), there is functional specialization. TORC1 consists of Kog1, lst8 and either Tor1 or Tor2 and TORC2 consists of Avo1, Avo2, Avo3, lst8 and Tor2 (Martin & Hall, 2005). Tor1 and Tor2 are serine/threonine kinases that belong to the phosphatidylinositol-3 kinase (PI3K) family (Wedaman *et al.*, 2003). TORC1 has been related to nutrient signals and controls cell proliferation (Loewith *et al.*, 2002), meanwhile TORC2 is rapamycin insensitive and it is associated to the control of actin cytoskeleton and cell cycle (Jacinto *et al.*, 2004); it is also involved in cell wall integrity and in sphingolipid metabolism (Tabuchi *et al.*, 2006). In other words, TORC1 mediates the temporal control of cell growth by regulating the various signaling pathways that determine mass accumulation, and TORC2 mediates the spatial control of cell growth by regulating a Rho GTPase signaling pathway that ultimately affects the actin cytoskeleton (Martin & Hall, 2005). TORC1-mediated growth control in yeast is an important cellular event, not only because of ribosomes are required for growth, but also because ribosome biogenesis is a major consumer of cellular energy. In favourable conditions for optimal growth, yeast cells synthesize 2000 ribosomes per minute. This requires the coordinated activity of all three RNA polymerases transcribing several hundred of genes and therefore, a large portion of total cellular energy is committed to ribosome biogenesis, underscoring the need for tight regulation of ribosomal genes in response to nutrient and energy conditions (Warner, 1999).

Therefore, the interesting question about how Ixr1 interacts with the TOR signaling arises and this is of particular interest considering the recent implication of yeast HMG proteins with TOR signaling (Chen *et al.*, 2013). TORC1 is activated by nutrients and it has been suggested that abundance of branched amino acids, especially leucine, regulates TOR activity (Martin & Hall, 2005). Considering that our data from the transcriptome analysis indicate that Ixr1 in absence of stress is necessary for the synthesis of leucine, isoleucine and valine (Vizoso *et al.*, 2012), this could be in accordance with the importance of Ixr1 to maintain TORC1 active

185

and thus sustaining ribosome biogenesis. However this point may be controversial because apparently yeast cell adjusts its transcriptional program, metabolic machinery and growth rate on the basis of the nutrient status available in the media, rather than on the basis of metabolites actually produced by the cell (Figure 10) (Zaman *et al*., 2009; Slattery *et al*., 2008; Levy *et al*., 2007).

The best-characterized TORC1 target in yeast is Sch9, which is phosphorylated by Tor1 resulting in Sch9 activation (Urban *et al*., 2009). Sch9 is a kinase that phosphorylates other downstream factors involved in the regulation of ribosomal transcription, mRNA export, and protein translation (Huber *et al*., 2009). The Tor1 kinase can also be directly recruited to the promoter regions of the RNA polymerase I and III (Pol I and Pol III) to activate the transcription of 35S and 5S ribosomal DNA (rDNA) genes respectively (Li *et al*., 2006; Wei *et al*., 2009). The master regulator of Ribi and RP genes is Sfp1. This factor localizes inside the nucleus, in active growing cells but it rapidly translocates into the cytoplasm in response to carbon and nitrogen starvation, oxidative stress, as well as inactivation of TOR signaling (Jorgensen *et al*., 2004; Marion *et al*., 2004; Fingerman *et al*., 2003). TORC1 complex regulates Sfp1 function via phosphorylation at multiple residues. Sfp1, in turn, negatively regulates TORC1 phosphorylation of Sch9 in a feedback mechanism (Lempiäinen *et al*., 2009).

Along this work, we have found that the expression of the *SFP1* gene (≈-1.5 change fold), as well as SRD1 gene is (≈-3.7 change fold), are highly reduced as a consequence of *IXR1* deletion (table 4) and this could explain the dependence of the yeast cells on Ixr1 to maintain ribosome biogenesis. Besides, Sfp1 regulates other TFs and co-regulators affecting transcription mediated by RNA polymerases I, II and III, as further explained, and therefore, this effect of Ixr1 might be amplified in the final control on ribosome biogenesis.

Fhl1, a forkhead-like protein, together with its co-regulators Ifh1 (coactivator) and Crf1 (corepressor) has a dual role as an activator and a repressor in the transcription of ribosomal protein genes (Jorgensen *et al.*, 2004; Martin *et al.*, 2004; Schawalder *et al.*, 2004; Wade *et al.*, 2004; Rudra *et al.*, 2005). Fhl1 is constitutively bound to RP gene promoters and therefore its activity depends on their partners Ifh1 and Crf1. In growing cells, TOR maintains Crf1 inactive in the cytoplasm by repressing the Yak1 kinase, possibly via a PKA dependent route. After TOR inactivation, the phosphorylated Crf1 in the nucleus displaces Ifh1, thereby inhibiting transcription of RP genes (Martin *et al.*, 2004). Besides, the nuclear localization of both Fhl1 and Ifh1 is influenced by Sfp1 (Jorgensen *et al.*, 2004). The role of Crf1 in repression of RP genes is strain-dependent (Zhao *et al.*, 2006).

In this work we have found that Ixr1 is also a repressor of the *CRF1* gene expression (table 3). Ixr1 will in turn regulate Fhl1 and its two partners Ifh1 and Crf1 directly or through its effect on Sfp1 expression (Figure 10).

In the formation of the RNA-Pol II active transcriptional complex that initiates transcription of the RP genes, and in which Fhl1 and Ifh1 are included, also participate Sfp1, Hmo1, Esa1 and Rap1 (Bustin *et al.*, 2012). Interestingly, both Hmo1 and Rap1, form part of the "poised" target genes of Ixr1, meanwhile binding of Ixr1 was observed in the exonic region of Sfp1 gene (Figure 10). Esa1, is a histone acetylase subunit of NuA4 that has been implicated in the activation of RP genes in yeast (Reid *et al.*, 2000), while histone de-acetylation by Rpd3, a histone deacetylase subunit of the Rpd3-Sin3 complex causes repression (Rohde & Cardenas, 2003).

The RNA-Pol II active transcriptional complex that initiates transcription of the Ribi genes is composed of Sfp1, Dot6 and Stb3. Dot6, part of the "poised" target genes of Ixr1 found in our study, is also a subunit of the Rpd3L histone deacetylase complex (Shevchenko *et al.*, 2008), and Stb3 interacts with Sin3 a

component of both the Rpd3S and Rpd3L histone deacetylase complexes (Figure 10) (Liko *et al*. 2007; Silverstein & Ekwall, 2005).

The activation of rDNA transcription by RNA Pol I depends on the formation of a RNA polymerase I preinitiation complex (PIC) in which take part Hmo1, Tor1, Rrn3, CF and UAF (Bustin *et al*., 2012). Ixr1 may affect this process since Hmo1 is included in its "poised" targets. UAF is an RNA polymerase I specific transcription stimulatory factor composed of Uaf30, Rrn5, Rrn9, Rrn10, histones H3 and H4 (Hontz *et al*., 2008). Tor1 is imported into the nucleus by the importin Srp1 and after inhibition of TOR1 activity (with rapamycin), Tor1 is exported from the nucleus by the exportin Crm1 (Hong *et al*., 2006).

Rpd3, a histone deacetylase, has a negative effect upon transcription of both rDNA and RP genes (Bustin *et al*., 2012); it is a component of both the Rpd3S and Rpd3L complexes; regulates transcription, silencing, autophagy and other processes by influencing chromatin remodeling; Rpd3(L) recruitment to the subtelomeric region is regulated by interaction with the arginine methyltransferase, Hmt1. Again, both Rpd3 and Hmt1 are found to be Ixr1 targets in this study.

TORC1 also regulates cell growth and proliferation by epigenetic mechanisms that affect to histones H3 and H4 and a possible model has been suggested in which HMGB proteins are involved (Chen *et al*., 2013). Hmo1 is a chromatin associated HMGB protein that regulates transcription from RNA polymerase II promoters (Kasahara *et al*., 2008) and also RNA polymerase I promoters (Merz K, *et al*., 2008). DNA bridging and looping by Hmo1 provides a mechanism for stabilizing nucleosome-free chromatin (Murugesapillai *et al*., 2014). Chen and co-workers have probed that TORC1 signaling is required for steady-state global chromatin binding by HMGB proteins, Nhp10 and to a minor extent Hmo1,

and that the H3K37A mutation alone can negatively affect Nhp10 binding to specific TORC1-regulated genes (Chen *et al*., 2013).

Ixr1 deletion affects the expression of several genes related to chromatin modification (tables 3 and 4), with the down-regulation of several DNA acetylases (*HFI1, SPT8, ADA2, AYT1, NAT4*), DNA desacetylases (*UME6, RXT3*) and DNA metylases (*JHD2, SWD3*), and in this sense Ixr1 could also participate in epigenetic mechanisms controlled by TORC1 (Figure 10).

The TOR pathway also controls DNA damage responses by controlling dNTP production (Shen *et al*., 2007), and the TOR pathway effectors Sch9 and Sfp1 are known to be involved in both ribosome biogenesis and stress responses (Smets *et al*., 2008; Marion *et al*., 2004). These features highlight other links between the TOR pathway and Ixr1 function since Ixr1 also controls the dNTP pools (Tsaponina *et al*., 2011) and also mediates stress responses elicited by hypoxia or oxidative damage (Castro-Prego *et al*., 2010a). Several genes up or down regulated after *IXR1* deletion are related to dNTP production or stress response (Tables 3 and 4). In this sense, *RNR1* gene is down-regulated (-1.79 fold change) after Ixr1 depletion (Figure 10). It was previously described that Ixr1 was involved in the transcriptional activation of *RNR1* to a DNA-damage response independently of the Dun1-Crt1 signaling pathway (Tsaponina *et al*., 2011).

Rapamycin-induced down-regulation of ribosomal genes is suppressed by activation of the Ras-cAMP-PKA (protein kinase A) pathway and TOR controls the subcellular localization of PKA and the PKA-regulated kinase Yak1 (Schmelzle *et al*., 2004). Many interconnections between the two signaling pathways have been found (Bustin *et al*., 2012). The subcellular localization of Sfp1, the master regulator of Ribi and RP genes, is regulated by both the cAMP/PKA and TOR network in response to nutritional and stress inputs (Marion *et al*., 2004). Besides

TOR, via PKA, negatively regulates the kinase Yak1 and thereby maintains Crf1 in the cytoplasm allowing the expression of RP genes (Martin *et al.*, 2004).

Although nitrogen supply and principally the synthesis of glutamine from glutamate, using the ammonium imported by the Mep2 permease, activates the TOR signaling pathway (Vinod *et al.*, 2008), optimal growth also requires a carbon source. In yeast, the cAMP/PKA pathway, which regulates according to nutrients availability, growth, proliferation, metabolism, stress resistance, aging, and morphogenesis, is activated by glucose. Since TOR and cAMP/PKA are connected, changes in glucose availability may also affect the final TOR targets. The levels of cAMP are controlled by two distinct G-protein systems: the Ras pathway and the Gpr1-Gpa2 pathway (Santangelo, 2006). Ras1 and Ras2 are two small monomeric GTP-binding proteins capable to switch between an active GTP-bound state and an inactive GDP-bound form. When in their active conformation, Ras proteins stimulate cAMP production by direct binding to adenylate cyclase (Toda *et al.*, 1985). The GPCR system, composed by the Gpr1 receptor and its cognate G protein Gpa2, is other glucose-sensing system that works in parallel with Ras to activate PKA (Santangelo, 2006). Signaling through the GPCR system is strictly dependent on sugar uptake and phosphorylation (Rolland *et al.*, 2000; Rolland *et al.*, 2001).

Hexoses are transported by the multiple Hxts present in the yeast plasma membrane with different binding affinities, which allow a better use of available monosaccharides in the media. In the absence of glucose the Rgt1 repressor binds synergistically to multiple sites found in the upstream regions of most *HXT* genes blocking their transcription. Rgt1 transcription repressing activity requires Mth1 and Std1 (Schmidt *et al.*, 1999) and the recruitment of the general repressors Ssn6 and Tup1 (Polish *et al.*, 2005). Two Hxt homologous proteins, Snf3 and Rgt2, have glucose sensor function that take part in the activation of the *HXT* genes.
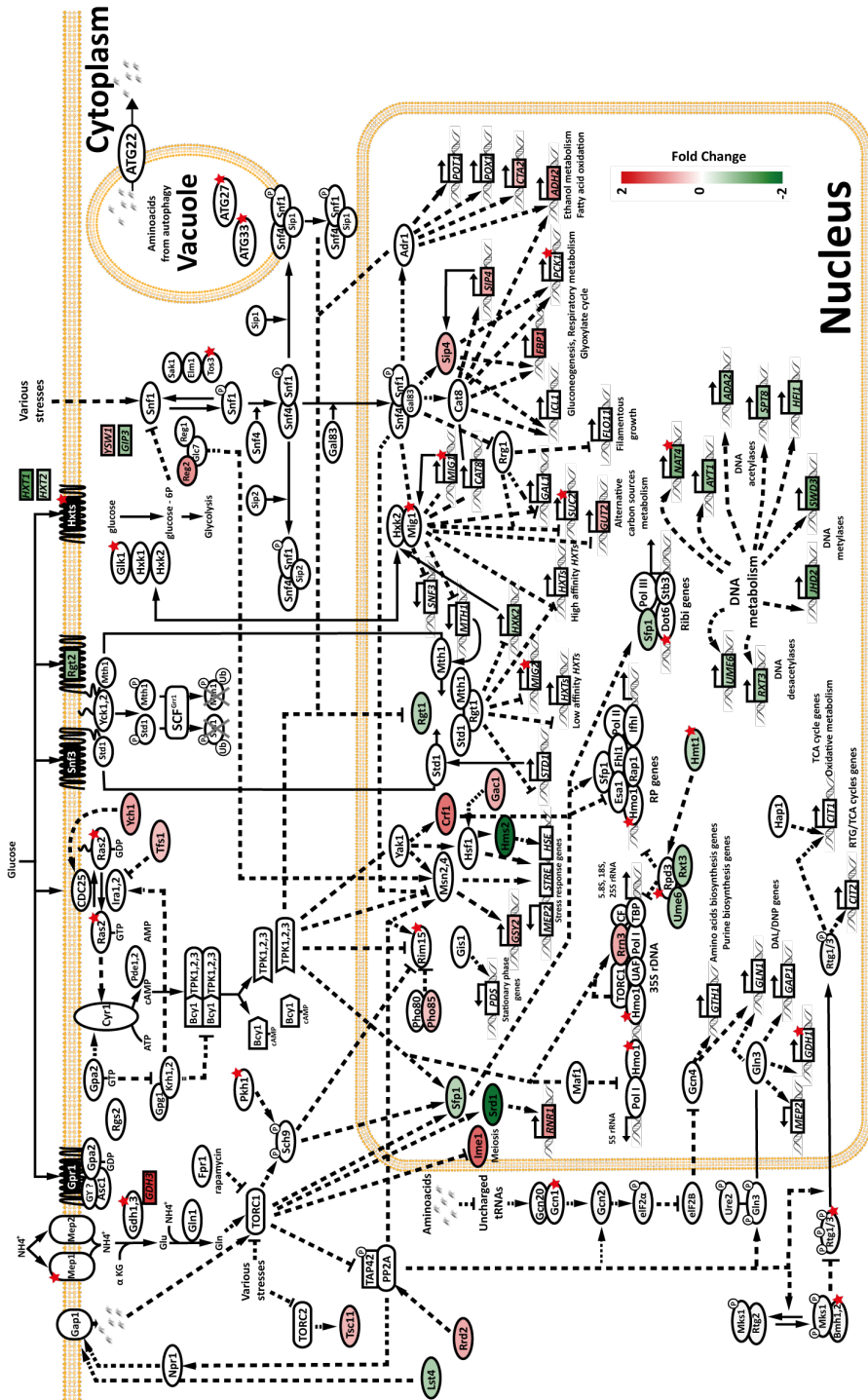
← **Figure 10**. Schematic view (modified from Busti *et al.*, 2010; and Busti *et al*., 2012) that shows how *IXR1* deletion alters the transcription of several genes encoding proteins involved in nutrient sensing pathways. Squares (genes) and circles (proteins) with colour shading are transcriptionally influenced by Ixr1. Colour code indicates fold change values, as indicated in the legend. Red stars indicate those genes in whose promoter region was detected an Ixr1 binding site by ChIP-on-chip experiment.

The C-terminal tails of Snf3 and Rgt2 may enhance signaling by facilitating the recruitment of the Mth1 and Std1 co-repressors to the plasma membrane thus avoiding their nuclear function (Moriya & Johnston, 2004). The glucose transporters have no regulatory function, being only required to maintain a critical level of intracellular glucose to sustain sugar phosphorylation, which is necessary for the activation of the GPCR system (Rolland *et al.*, 2000). In addition, neither of the two glucose sensors Snf3 and Rgt2 has a direct role in the cAMP signaling (Rolland *et al*., 2001). Deletion of Ixr1 also affect the expression of genes encoding hexose transporters (*HXT2* -1,4; *HXT1* -2,48) and Sok2, a player in the PKA signal transduction pathway in response to glucose (Shenhar & Kassir, 2001; Malcher et a., 2011) is also a target of Ixr1 (Figure 10).

In conclusion, data obtained in our study show that Ixr1 could be regarded as a master regulator of the regulation of cellular growth and proliferation via TOR and PKA signaling and in accordance to nutrient availability.

3.8.3. Ixr1 In the response to cisplatin

Cisplatin treatment has been considered as a source of oxidative damage that affects DNA and other biomolecules in the cell (Martins *et al.*, 2008; Pratibha *et al*., 2006) and therefore it can elicit a yeast response to stress. The stress responses are usually accompanied by down-regulation of genes related to ribosome synthesis and to reorganization of metabolic fluxes.

As a work hypothesis to explain the effect of cisplatin treatment upon ribosome biogenesis we may postulate that oxidative damage caused by cisplatin

and the formation of adducts between cisplatin and modified DNA could contribute to displace Ixr1 from regulated promoters to the regions of DNA lesions in order to block DNA repair and promote cell dead as previously established (McA´Nulty & Lippard, 1996). To test this hypothesis we compared the binding of Ixr1 to DNA promoters in absence and presence of cisplatin by ChIP-on-chip analyses. However, what we see is the opposite effect, the enrichment in promoter-targeting in presence of cisplatin treatment, increasing from 46.3% to 69.1% of promoter-targeting binding regions of Ixr1 when comparing controls and cisplatin treatments, respectively. Up to 37% of the Ixr1 promoter binding peaks contain a consensus sequence closely related to the specific binding region of the Rsc30 protein (figure 7). *Saccharomyces cerevisiae* RSC (*Remodels the Structure of Chromatin*) is an essential chromatin-remodeling enzyme consisting of 17 subunits (Sth1, Rsc1, Rsc2, Rsc3, Rsc4, Rsc6, Rsc7/Npl6, Rsc8, Rsc9, Sfh1, Arp7, Arp9, Rsc30, Htl1, Rtt102, Rsc58 and Ldb7).

(Cairns *et al.*, 1996). RSC is an ATP-dependent remodeling complex that modulates the access to chromatin, and therefore DNA metabolism, including DNA replication, transcription, recombination, and repair. A large number of different remodeling activities can be performed by these complexes, including exchange or incorporation of core histones or histone variants, eviction of histones or nucleosomes and repositioning or sliding of nucleosomes (Clapier & Cairns, 2009). Interestingly, it has been reported the role of RSC complexes in the transcription of genes related to the ribosome biogenesis (Angus-Hill *et al.*, 2001) and also to DNA DSB (*Double Strand Breaks*) repair, where strains lacking Rsc1 or Rsc2 protein components are hypersensitive to a variety of DNA damaging agents (Chai *et al.*, 2005; Shim *et al.*, 2005; Kent *et al.*, 2007; Chambers *et al.*, 2012).

The cytotoxic effect of cisplatin in humans is attributed to diverse mechanisms, among them a decrease in ribosome biogenesis. It has been reported that ribosomal RNA transcription *in vitro* using a reconstituted system, is

specifically inhibited in the presence of cisplatin-DNA adducts (Zhai *et al.*, 1998). The transcription factor hUBF is a HMG-box protein that stimulates ribosomal RNA synthesis (Treiber *et al*., 1994) but also has a high affinity for cisplatin-DNA adducts; in the reconstituted system a ratio of adducts/promoter binding of approximately 4:1 completely abolished the transcription activated by hUBF (Zhai *et al*., 1998). The inhibition of ribosomal RNA synthesis by cisplatin was also demonstrated *in vivo* (Jordan & Carmo-Fonseca, 1998) and this inhibition was accompanied by a redistribution of UBF and the transcriptional machinery associated to RNA polymerase I towards the nucleolus (Jordan & Carmo-Fonseca, 1998).

In *Saccharomyces cerevisiae* we have previously shown that cisplatin represses the transcription of genes related to rRNA and ribosomal proteins synthesis (Rodríguez Lombardero *et al*., 2014), and therefore this mechanism of cisplatin-induced cytotoxicity is also functional in yeast. In this work, we see that deletion of *IXR1* causes a decrease in this repression, which could contribute to the observed increase of resistance towards cisplatin observed in the *Δixr1 S. cerevisiae* strain (McA´Nulty & Lippard, 1996; Huang *et al*., 2005). Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules (Merz *et al*., 2008). Besides, Hmo1 interacts with TFIID and participates in start site selection by RNA polymerase II (Kasahara *et al*., 2008) as well as with Fhl1, a positive transcriptional regulator of RP genes together with the co-activator Ifh1, and therefore may affect the transcription of ribosomal proteins. Deletion of Hmo1 diminishes cisplatin resistance (SGD) indicating that not only is necessary for rRNA and RP transcription but it is also a necessary target of signals that promote cisplatin resistance. Our data indicate that transcription of the genes *HTB2, HTA2* and *HTA1* decreases in the *Δixr1 S. cerevisiae* strain *versus* wild type after cisplatin treatment. Thus, decrease of histones A and B availability, as

consequence of *IXR1* deletion, could favour Hmo1 binding to specific targets and therefore, increase cisplatin resistance.

Other open question is how cisplatin could signal the decay in TOR1C signaling that is necessary for the decrease of ribosome biogenesis. The known effects of cisplatin, DNA damage and stimulation of a response to oxidative stress (Marullo *et al.*, 2013) are apparently more related to TORC2 than to TORC1 signaling. TORC2-dependent regulation of the actin cytoskeleton is required to maintain the polarized nature of cell growth in budding yeast and is required for endocytosis as well as genome stability in response to DNA damage (deHart *et al.*, 2003; Shimada *et al.*, 2013). TORC2 and its downstream kinase Ypk1 regulate actin polarization by controlling reactive oxygen species (ROS) accumulation; both by vacuole-related ROS, controlled by the phospholipid flippase kinase Fpk1 and sphingolipids, and by mitochondria-mediated ROS, controlled by the PKA subunit Tpk3 (Niles & Powers, 2014). Besides, several connections between TORC1 and TORC2 are arising and it has been demonstrated that an interaction of TORC2 with ribosome is conserved from yeast to mammals and necessary for TORC2 activity (Zinzalla *et al.*, 2011); Future studies are needed to address this important question. RSC complex is required for the transcriptional induction of autophagy, activating the *ATG7*, *ATG8* and *ATG17* genes, and participates in the inactivation of the TORC1 pathway, enhancing the Rho1-Kog1 binding (Yu *et al.*, 2015). Furthermore, TORC1 pathway plays a crucial role in the induction of autophagy (Mizushima *et al.*, 2011) and, in spite of the fact that nutrient starvation inhibits ability of TOR kinase to induce autophagy in both yeast and mammalian cells, little is known about the mechanisms regulating the transcriptional activation of these genes and the role of the RSC complex in this process (Xie *et al.*, 2008).

Other point derived from our study is that the previously observed up-regulation of genes from the sulphur assimilation pathways and the biosynthesis of cysteine and methionine pathways, after cisplatin treatment is even enlarged by

IXR1 deletion. This could also favour cisplatin resistance in the null mutant by increasing chelating groups to immobilise the Pt compound or even promote glutathione biosynthesis to favour anti-oxidant reactions or for cisplatin extrusion out of the cell through the formation of cisplatin-glutathione complexes. The stimulation of a pathway of amino acid biosynthesis is paradoxical in a situation in which other protein biosynthesis pathways are down regulated, principally the biosynthesis of ribosomal proteins. However, specific sulphur-containing amino acids may be necessary to cope with oxidative stress. Besides, ribosomal proteins are among those with low cysteine content (Miseta & Csutora, 2000) and therefore, re-utilization of amino acids released from protein degradation may not support adequately the necessary supply in sulphur-containing amino acids.

4.- REFERENCES

Abdulrehman D., P.T. Monteiro, M.C. Teixeira, N.P. Mira, A.B. Lourenço, S.C. dos Santos, T.R. Cabrito, A.P. Francisco, S.C. Madeira, R.S. Aires, A.L. Oliveira, I.Sá-Correia and A.T. Freitas (2011). "YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface." Nucleic Acids Research **39** (Database issue): D136-D140.

Akache B., S. MacPherson, M.A. Sylvain and B. Turcotte (2004). "Complex interplay among regulators of drug resistance genes in Saccharomyces cerevisiae." Journal of Biological Chemistry **279**(27):27855-60.

Angus-Hill M.L., A. Schlichter, D. Roberts, H. Erdjument-Bromage, P. Tempst and B.R. Cairns (2001). "A Rsc3/Rsc30 zinc cluster dimer reveals novel roles for the chromatin remodeler RSC in gene expression and cell cycle control." Mollecular Cell **7**(4):741-51

Angus-Hill M.L., K.M. Elbert, J. Hidalgo, M.R. Capecchi (2011). "T-cell factor 4 functions as a tumor suppressor whose disruption modulates colon cell proliferation and tumorigenesis." Proceedings of the National Academy of Sciences of the United States of America **108**(12):4914-9.

Arenhart R.A., Y. Bai, L.F. de Oliveira, L.B. Neto, M. Schunemann, Fdos S. Maraschin, J. Mariath, A. Silverio, G. Sachetto-Martins, R. Margis, Z.Y. Wang, M. Margis-Pinheiro (2014). "New insights into aluminum tolerance in rice: the ASR5 protein binds the STAR1 promoter and other aluminum-responsive genes." Molecular Plant **7**(4):709-21.

Ashe M.P., S.K. De Long and A.B. Sachs (2000). "Glucose depletion rapidly inhibits translation initiation in yeast." Molecular Biology of Cell **11**(3):833-48.

Bailey T.L. and C. Elkin (1994). "Proceedings of the Second International Conference on Intelligent Systems" for Molecular Biology, AAAI Press, Menlo Park, CA, pp. 28–36.

Bar-Joseph Z., G.K. Gerber, T.I. Lee, N.J. Rinaldi, J.Y. Yoo, F. Robert, D.B. Gordon, E. Fraenkel,T.S. Jaakkola, R.A. Young and D.K. Gifford (2003). "Computational discovery of gene modules and regulatory networks." Nature Biotechnology **21**(11): 1337–1342.

Beinoraviciūte-Kellner R., G. Lipps, G. Krauss (2005). "*In vitro* selection of DNA binding sites for *ABF1* protein from *Saccharomyces cerevisiae*." FEBS Letters **579**(20):4535-40.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society **57**, 289–300.

Bertram P.G., J.H. Choi, J. Carvalho, T.F. Chan, W. Ai and X.F. Zheng (2002). "Convergence of TOR-nitrogen and Snf1-glucose signaling pathways onto Gln3." Molecular Cell Biology **22**(4):1246-52."

Bianco S., M. Brunelle, M. Jangal, L. Magnani, N. Gévry (2014). "LRH-1 governs vital transcriptional programs in endocrine-sensitive and -resistant breast cancer cells." Cancer Research **74**(7):2015–2025

Birrell G.W., J.A. Brown, H.I. Wu, G. Giaever, A.M. Chu, R.W. Davis, J.M. Brown (2002). "Transcriptional response of *Saccharomyces cerevisiae* to DNA-damaging agents does not identify the genes that protect against these agents." Proceedings of the National Academy of Sciences USA **99**(13):8778-83.

Bolduc N., A. Yilmaz, M.K. Mejia-Guerra, K. Morohashi, D. O'Connor, E. Grotewold, S. Hake (2012). "Unraveling the KNOTTED1 regulatory network in maize meristems." Genes Development **26**(15):1685-90.

Bourdineaud J.P., G. De Sampaïo and G.J. Lauquin (2000). "A Rox1-independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the *SRP1/TIR1* gene." Molecular Microbiology **38**(4):879-90.

Bricmont P.A., J.R. Daugherty, T.G. Cooper (1991). "The *DAL81* gene product is required for induced expression of two differently regulated nitrogen catabolic genes in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **11**(2):1161-6.

Brückner S., S. Kern, R. Birke, I. Saugar, H.D. Ulrich, H.U. Mösch (1990). "The TEA transcription factor Tec1 links TOR and MAPK pathways to coordinate yeast development." Genetics **189**(2):479-94.

Buchman A.R. & R.D. Kornberg (1990). "A yeast ARS-binding protein activates transcription synergistically in combination with other weak activating factors." Molecular and Cellular Biology **10**(3):887-97.

Busnelli S., F. Tripodi, R. Nicastro, C. Cirulli, G. Tedeschi, R. Pagliarin, L. Alberghina and P. Coccetti (2013). "Snf1/AMPK promotes SBF and MBF-dependent transcription in budding yeast." Biochimical Biophysics Acta **1833**(12):3254-64.

Cairns B.R., Y. Lorch, Y. Li, M. Zhang, L. Lacomis, H. Erdjument-Bromage, P. Tempst, J. Du, B. Laurent, R.D. Kornberg (1996). "RSC, an essential, abundant chromatin-remodeling complex." Cell **87**(7):1249-60.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, I. Carneiro, I. González-Siso and M.E. Cerdán (2009). "Regulatory factors controlling transcription of Saccharomyces cerevisiae IXR1 by oxygen levels: a model of transcriptional adaptation from aerobiosis to hypoxia implicating ROX1 and IXR1 cross-regulation." Biochemical Journal **425**(1):235-43.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso and M.E. Cerdán (2010b). "Ixr1p regulates oxygen-dependent HEM13 transcription." FEMS Yeast Research **10**(3):309-21.

Chai B., J. Huang, B.R. Cairns, B.C. Laurent (2005). "Distinct roles for the RSC and Swi/Snf ATP-dependent chromatin remodelers in DNA double-strand break repair." Genes Development **19**(14):1656-61.

Chambers A., C. Stanway, J.S. Tsang, Y. Henry, A.J. Kingsman, S.M. Kingsman (1990). "ARS binding factor 1 binds adjacent to *RAP1* at the UASs of the yeast glycolytic genes *PGK* and *PYK1*." Nucleic Acids Research **18**(18):5393-9.

Chambers A.L., P.M. Brownlee, S.C. Durley, T. Beacham, N.A. Kent, J.A. Downs (2012). "The two different isoforms of the RSC chromatin remodeling complex play distinct roles in DNA damage responses." PLoS One **7**(2):e32016.

Chow C.S., J.P. Whitehead, S.J. Lippard (1994) "HMG domain proteins induce sharp bends in cisplatin-modified DNA" Biochemistry **33**(50):15124-30.

Cipollina C., L. Alberghina, D. Porro, M. Vai. (2005) "SFP1 is involved in cell size modulation in respiro-fermentative growth conditions". Yeast. **22**(5):385-99.

Clapier C.R., B.R Cairns (2009). "The biology of chromatin remodeling complexes." Annual Review of Biochemistry **78**:273-304.

Cosma M. P., S. Panizza, and K. Nasmyth (2001). "Cdk1 triggers association of RNA polymerase to cell cycle promoters only after recruitment of the mediator by SBF." Molecular Cell **7**(6):1213–1220.

Costanzo M., O. Schub and B. Andrews (2003). "G1 transcription factors are differentially regulated in Saccharomyces cerevisiae by the Swi6-binding protein Stb1." Molecular of Cell Biology **23**(14):5064-77.

Dastidar R.G., J. Hooda, A. Shah, T.M. Cao, R.M. Henke, L. Zhang (2012). "The nuclear localization of SWI/SNF proteins is subjected to oxygen regulation." Cell & Bioscience **2**(1):30.

deHart A.K., J.D. Schnell, D.A. Allen, J.Y. Tsai, L. Hicke (2003). "Receptor internalization in yeast requires the Tor2-Rho1 signaling pathway." Molecular Biology of the Cell **14**(11):4676-84.

Della Seta F., I. Treich, J.M. Buhler, A. Sentenac (1990). "*ABF1* binding sites in yeast RNA polymerase genes." Journal of Biological Chemistry **265**(25):15168-75.

Dias P.J., M.L. Seret, A. Goffeau, I.S. Correia, P.V. Baret (2010). "Evolution of the 12-spanner drug:H$^+$ antiporter DHA1 family in hemiascomycetous yeasts." OMICS **14**(6):701-10.

Elkhaimi M., M.R. Kaadige, D. Kamath, J.C. Jackson, H. Biliran, J.M. Lopes (2000).

"Combinatorial regulation of phospholipid biosynthetic gene expression by the *UME6*, *SIN3* and *RPD3* genes." Nucleic Acids Research **28**(16):3160-7.

Fingerman I., V. Nagaraj, D. Norris, A.K. Vershon (2003). "Sfp1 plays a key role in yeast ribosome biogenesis." Eukaryotic Cell **2**(5):1061-1068.

French S. L., Y. N. Osheim, F. Cioci, M. Nomura, and A. L. Beyer (2003). "In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes." Molecular of Cell Biology **23**(5): 1558–1568.

Gailus-Durner V., J. Xie, C. Chintamaneni, A.K. Vershon (1996). "Participation of the yeast activator Abf1 in meiosis-specific expression of the *HOP1* gene." Molecular and Cellular Biology **16**(6):2777-86.

Geistlinger L., G. Csaba, S. Dirmeier, R. Küffner and R.A. Zimmer (2013) "A comprehensive gene regulatory network for the diauxic shift in Saccharomyces cerevisiae." Nucleic Acids Research **41**(18):8452-63.

Georis I., J.J. Tate, T.G. Cooper, E. Dubois (2008). "Tor pathway control of the nitrogen-responsive *DAL5* gene bifurcates at the level of Gln3 and Gat1 regulation in *Saccharomyces cerevisiae*." Journal of Biological Chemistry **283**(14):8919-29.

Gerlitz G., R. Hock, T. Ueda, M. Bustin (2009). "The dynamics of HMG protein-chromatin interactions in living cells." Biochemistry and Cell Biology **87**(1):127-37.

Ghaemmaghami S, W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'Shea and J.S. Weissman (2003). "Global analysis of protein expression in yeast." Nature **425**(6959):737-41.

Gietz R. and A. Sugino (1988). "New yeast–Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites." Gene **74**(2): 527–534.

Gimeno C.J. and G.R. Fink (1994). "Induction of pseudohyphal growth by overexpression of PHD1, a Saccharomyces cerevisiae gene related to transcriptional regulators of fungal development." Molecular of Cell Biology **14**(3):2100-12.

González A., A. Ruiz, A. Casamayor, J. Ariño (2009). "Normal function of the yeast TOR pathway requires the type 2C protein phosphatase Ptc1." Molecular and Cellular Biology **29**(10):2876-88.

Gorski J.J., K.I. Savage, J.M. Mulligan, S.S. McDade, J.K. Blayney, Z. Ge, D.P. Harkin (2011). "Profiling of the BRCA1 transcriptome through microarray and ChIP-chip analysis." Nucleic Acids Research **39**(22):9536-48.

Granneman S., E. Petfalski, and D. Tollervey (2011). "A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease." EMBO Journal **30**(19): 4006– 4019.

Harrington L. A. and B. J. Andrews (1996). "Binding to the yeast SwI4,6- dependent cell cycle box, CACGAAA, is cell cycle regulated in vivo." Nucleic Acids Research **24**(4):558–565.

Harris M.R., D. Lee, S. Farmer, N.F Lowndes and R.A. de Bruin (2013). "Binding specificity of the G_1/S transcriptional regulators in budding yeast." PLoS One **8**(4):e61059.

Heinz S., C. Benner, N. Spann, E. Bertolino E, Y.C. Lin, P. Laslo, J.X. Cheng, C. Murre, H. Singh and C.K. Glass (2010). "Simple Combinations of Lineage-Determining Transcription Factors Prime *cis*-Regulatory Elements Required for Macrophage and B Cell Identities." Molecular Cell **38**(4):576-589.

Hermann-Le Denmat, S., M. Werner, A. Sentenac and P. Thuriaux (1994). "Suppression of yeast RNA polymerase III mutations by FHL1, a gene coding for a fork head protein involved in rRNA processing." Molecular Cell Biology **14**(5): 2905–2913.

Ho Y., M. Costanzo, L. Moore, R. Kobayashi, B.J. Andrews (1999). "Regulation of transcription at the Saccharomyces cerevisiae start transition by Stb1, a Swi6-binding protein." Molecular Cell Biology **19**(8):5267-78.

Ho Y., A. Gruhler, A. Heilbut, G.D. Bader, L. Moore, S.L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A.R. Willems, H. Sassi, P.A. Nielsen, K.J. Rasmussen, J.R. Andersen, L.E. Johansen, L.H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B.D. Sørensen, J. Matthiesen, R.C. Hendrickson, F. Gleeson, T. Pawson, M.F Moran, D. Durocher, M. Mann, C.W. Hogue, D. Figeys and M. Tyers. (2002). "Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry." Nature **415**(6868):180–183.

Hontz R.D., S.L. French, M.L. Oakes, P. Tongaonkar, M. Nomura, A.L. Beyer, J.S. Smith (2008). "Transcription of multiple yeast ribosomal DNA genes requires targeting of UAF to the promoter by Uaf30." Molecular and Cellular Biology **28**(21):6709-19.

Hughes T.R., C.G. de Boer (2013). "Mapping yeast transcriptional networks. Genetics **195**(1):9–36.

Ito T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori and Y. Sakaki (2001). "A comprehensive two-hybrid analysis to explore the yeast protein interactome." Proceedings of the National Academy of Sciences USA **98**(8): 4569–4574.

Iyer V.R., C.E. Horak, C.S Scafe, D. Botstein, M. Snyder and P.O. Brown (2001). "Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF." Nature **409**(6819): 533–538.

Jin, Y. H., P.E. Dunlap, S.J. McBride, H. Al-Refai, P.R. Bushel and J.H. Freedman (2008). "Global transcriptome and deletome profiles of yeast exposed to transition metals." PLoS Genetics **4**(4): e1000053.

Jordan P. and M. Carmo-Fonseca (1998). "Cisplatin inhibits synthesis of ribosomal RNA in vivo." Nucleic Acids Research **26**(12): 2831–2836.

Jorgensen P., I. Rupes, J.R. Sharom, L. Schneper, J.R. Broach, M. Tyers (2004). "A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size." Genes Development **18**(20):2491-505.

Kasahara K., S. Ki, K. Aoyama, H. Takahashi, T. Kokubo (2008). "*Saccharomyces cerevisiae* HMO1 interacts with TFIID and participates in start site selection by RNA polymerase II." Nucleic Acids Research **36**(4):1343-57.

Kent N.A., A.L. Chambers, J.A. Downs (2007). "Dual chromatin remodeling roles

for RSC during DNA double strand break induction and repair at the yeast MAT locus." Journal of Biological Chemistry **282**(38):27693-701.

Kim J.H., J. Polish and M. Johnston (2003). "Specificity and regulation of DNA binding by the yeast glucose transporter gene repressor Rgt1." Molecular Cell Biology **23**(15):5208-16.

Koch C., T. Moll, M. Neuberg, H. Ahorn, and K. Nasmyth (1993). "A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase." Science **261**(5128):1551–1557.

Kos M. and D. Tollervey (2010). "Yeast pre-rRNA processing and modification occur cotranscriptionally." Molecular Cell **37**(6): 809–820.

Kuchin S., I. Treich and M.A. Carlson (2000). "A regulatory shortcut between the Snf1 protein kinase and RNA polymerase II holoenzyme." Proceedings of the National Academy of Sciences U S A **97**(14):7916-20.

Lambert JR, V.W. Bilanchone, M.G. Cumsky (1994) "The ORD1 gene encodes a transcription factor involved in oxygen regulation and is identical to *IXR1*, a gene that confers cisplatin sensitivity to *Saccharomyces cerevisiae*". Proceedings of the National Academy of Sciences U S A **91**(15):7345-9.

Larochelle M., S. Drouin S, F. Robert and B. Turcotte (2006). "Oxidative stress-activated zinc cluster protein Stb5 has dual activator/repressor functions required for pentose phosphate pathway regulation and NADPH production". Molecular Cell Biology **26**(17):6690-701.

Lee T. I., N.J. Rinaldi, F. Robert, D.T. Odom, Z. Bar-Joseph, G.K. Gerber, N. M. Hannett, C.T. Harbison, C.M. Thompson, I. Simon, J. Zeitlinger, E.G. Jennings, H.L. Murray, D.B. Gordon, B. Ren, J.J. Wyrick, J.B. Tagne, T.L. Volkert, E. Fraenkel, D.K. Gifford and R.A. Young (2002). "Transcriptional regulatory networks in *Saccharomyces cerevisiae*." Science **298**(5594):799–804.

Lee W., R.P. St Onge, M. Proctor, P. Flaherty, M.I. Jordan, A.P. Arkin, R.W. Davis, C. Nislow, G. Giaever (2005). "Genome-wide requirements for resistance to functionally distinct DNA-damaging agents." PLoS Genetics **1**(2):e24.

Lempiäinen H., A. Uotila, J. Urban, I. Dohnal, G. Ammerer, R. Loewith, D. Shore (2009). "Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling." Molecular Cell **33**(6):704-716.

Levy S., J. Ihmels, M. Carmi, A. Weinberger, G. Friedlander, N. Barkai (2007). "Strategy of transcription regulation in the budding yeast." PLoS One **2**(2):250

Li H., C.K. Tsang, M. Watkins, P.G. Bertram, X.F. Zheng (2006). "Nutrient regulates Tor1 nuclear localization and association with rDNA promoter." Nature. **442**(7106):1058-61.

Li S., W. Sun, H. Wang, D. Zuo, Y. Hua, Z. Cai (2015). "Research progress on the multidrug resistance mechanisms of osteosarcoma chemotherapy and reversal." Tumour Biology **36**(3):1329-38.

Liao C., B. Hu, M.J. Arno and B. Panaretou (2007) "Genomic screening in vivo reveals the role played by vacuolar H+ ATPase and cytosolic acidification in sensitivity to DNA-damaging agents such as cisplatin." Molecular Pharmacology **71**(2):416–425.

Liko D., M.G. Slattery, W. Heideman (2007). "Stb3 binds to ribosomal RNA processing element motifs that control transcriptional responses to growth in *Saccharomyces cerevisiae*." Journal of Biological Chemistry **282**(36):26623-8.

Lo W.S., L. Duggan, N.C. Emre, R. Belotserkovskya, W.S. Lane, R. Shiekhattar and S.L. Berger (2001). "Snf1: a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription." Science **293**(5532):1142-6."

Lucau-Danila A., T. Delaveau, G. Lelandais, F. Devaux and C. Jacq (2003). "Competitive promoter occupancy by two yeast paralogous transcription factors controlling the multidrug resistance phenomenon." Journal of Biological Chemistry **278**(52):52641-50.

Malcher M., S. Schladebeck, H.U. Mösch (2011). "The Yak1 protein kinase lies at the center of a regulatory cascade affecting adhesive growth and stress resistance in *Saccharomyces cerevisiae*." Genetics **187**(3):717-30.

Marion R.M., A. Regev, E. Segal, Y. Barash, D. Koller, N. Friedman, E.K. O'Shea (2004). "Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression." Proceedings of the National Academy of Sciences U S A **101**(40):14315-22.

Martin D.E., A. Soulard and M.N. Hall (2004). "TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1." Cell **119**(7):969-79.

Martin D.E., M.N. Hall (2005). "The expanding TOR signaling network. Current Opinion in Cell Biology **17**(2):158-66.

Martins, N. M., N.A. Santos, C. Curti, M.L. Bianchi & A.C. Santos (2008) "Cisplatin induces mitochondrial oxidative stress with resultant energetic metabolism impairment, membrane rigidification and apoptosis in rat liver" Journal of Applied Toxicology **28**(3): 337–344.

Marullo R., E. Werner, N. Degtyareva, B. Moore, G. Altavilla, S.S. Ramalingam, P.W. Doetsch (2013). "Cisplatin induces a mitochondrial-ROS response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions." PLoS One **8**(11):e81162.

Mayordomo I., F. Estruch and P. Sanz (2002). "Convergence of the target of rapamycin and the Snf1 protein kinase pathways in the regulation of the subcellular localization of Msn2, a transcriptional activator of STRE (Stress Response Element)-regulated genes." Journal of Biological Chemistry **277**(38):35650-6.

McA'Nulty M.M. and S.J. Lippard (1996). "The HMG-domain protein Ixr1 blocks excision repair of cisplatin-DNA adducts in yeast." Mutat Research **362**(1):75-86.

McA'Nulty M.M., J.P. Whitehead and S.J. Lippard (1996). "Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts in vitro and in vivo." Biochemistry. **35**(19):6089-99.

McLeay R.C. and L.T. Bailey (2010). "Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data" BMC Bioinformatics **11**:165.

Medina I., J. Carbonell, L. Pulido, S.C. Madeira, S. Goetz, A. Conesa, J. Tarraga, A. Pascual-Montano, R. Nogales-Cadenas, J. Santoyo, F. García, M. Marbà, D.

Montaner D, J. Dopazo (2010). "Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling." Nucleic Acids Research **38** (Web Server): W210–W213.

Merz K., M. Hondele, H. Goetze, K. Gmelch, U. Stoeckl, J. Griesenbeck (2008). "Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules." Genes Development **22**(9):1190-204.

Mi H., A. Muruganujan, P.D. Thomas (2013). "PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees." Nucleic Acids Research **41**(Database issue):D377-86.

Miseta A., P. Csutora (2000). "Relationship between the occurrence of cysteine in proteins and the complexity of organisms." Molecular Biology and Evolution **17**(8):1232-9.

Mizushima N., T. Yoshimori, Y. Ohsumi (2011). "The role of Atg proteins in autophagosome formation." Annual Review of Cell and Developmental Biology **27**:107-32.

Mönke G., M. Seifert, J. Keilwagen, M. Mohr, I. Grosse, U. Hähnel, A. Junker, B. Weisshaar, U. Conrad, H. Bäumlein, L. Altschmied (2012). "Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon." Nucleic Acids Research **40**(17):8240-54.

Moriya H., M. Johnston (2004). "Glucose sensing and signaling in *Saccharomyces cerevisiae* through the Rgt2 glucose sensor and casein kinase I." Proceedings of the National Academy of Sciences **101**(6):1572-1577.

Morrow B.E., S.P. Johnson, J.R. Warner (1989). "Proteins that bind to the yeast rDNA enhancer." Journal of Biological Chemistry **264**(15):9061-8.

Nath N., R.R. McCartney and M.C. Schmidt (2003). "Yeast Pak1 kinase associates with and activates Snf1." Molecular Cell Biology **23**(11):3909-17.

Niles B.J., T. Powers (2014). "TOR complex 2-Ypk1 signaling regulates actin polarization via reactive oxygen species." Molecular Biology of the Cell **25**(24):3962-72.

Ozcan S., T. Leong and M. Johnston (1996). "Rgt1p of *Saccharomyces cerevisiae*, a key regulator of glucose-induced genes, is both an activator and a repressor of transcription." Molecular Cell Biology **16**(11):6419-26.

Ozsarac N., M.J. Straffon, H.E. Dalton, I.W. Dawes (1997). "Regulation of gene expression during meiosis in *Saccharomyces cerevisiae: SPR3* is controlled by both *ABF1* and a new sporulation control element." Molecular and Cellular Biology **17**(3):1152-9.

Pan X. and J.Heitman J. (2000). "Sok2 regulates yeast pseudohyphal differentiation via a transcription factor cascade that regulates cell-cell adhesion." Molecular Cell Biology **20**(22):8364-72.

Para A., Y. Li, A. Marshall-Colón, K. Varala, N.J. Francoeur, T.M. Moran, M.B Edwards, C. Hackley, B.O. Bargmann, K.D. Birnbaum, W.R. McCombie, G. Krouk, G.M. Coruzzi (2014). "Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in *Arabidopsis*." Proceedings of the National Academy of Sciences

111(28):10371-6.

Planta R.J. (1997) "Regulation of ribosome synthesis in yeast." Yeast **13**(16):1505-18.

Polish J.A., J.H. Kim, M. Johnston (2005). "How the Rgt1 transcription factor of *Saccharomyces cerevisiae* is regulated by glucose." Genetics **169**(2):583-594.

Pratibha R, R. Sameer, P.V. Rataboli, D.A. Bhiwgade, C.Y. Dhume (2006) "Enzymatic studies of cisplatin induced oxidative stress in hepatic tissue of rats" Eur J Pharmacol. **532**(3):290-3.

Monteiro P.T., N. Mendes, M. C. Teixeira, S. d'Orey, S. Tenreiro, N. Mira, H. Pais, A. P. Francisco, A.M. Carvalho, A. Lourenço and I. SáCorreia, A.L. Oliveira and A.T. Freitas (2008). "YEASTRACT-DISCOVERER: new tools to improve the analysis of transcriptional regulatory associations in *Saccharomyces cerevisiae*." Nucleic Acids Research **36**(Database): D132-D136.

Rahner A., A.Schöler, E .Martens, B. Gollwitzer and H.J. Schüller (1996). "Dual influence of the yeast Cat1p (Snf1p) protein kinase on carbon source-dependent transcriptional activation of gluconeogenic genes by the regulatory gene CAT8." Nucleic Acids Research **24**(12):2331-7.

Raithatha S.A. and D.T. Stuart (2005). "Meiosis-specific regulation of the Saccharomyces cerevisiae S-phase cyclin CLB5 is dependent on MluI cell cycle box (MCB) elements in its promoter but is independent of MCB-binding factor activity." Genetics **169**(3):1329-42.

Randez-Gil F., N. Bojunga, M. Proft and K.D. Entian (1997). "Glucose derepression of gluconeogenic enzymes in Saccharomyces cerevisiae correlates with phosphorylation of the gene activator Cat8p." Molecular Cell Biology **17**(5):2502-10.

Reed S.H., M. Akiyama, B. Stillman, E.C. Friedberg (1999). "Yeast autonomously replicating sequence binding factor is involved in nucleotide excision repair." Genes Development **13**(23):3052-8.

Régnacq M., P. Alimardani P, B. El Moudni and T. Bergès (2001). "SUT1p interaction with Cyc8p(Ssn6p) relieves hypoxic genes from Cyc8p-Tup1p repression in *Saccharomyces cerevisiae*." Molecular Microbiology **40**(5):1085-96.

Reid J.L., V.R. Iyer, P.O. Brown, K. Struhl (2000). "Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase." Molecular Cell **6**(6):1297-1307.

Rhode P.R., K.S. Sweder, K.F. Oegema, J.L. Campbell (1989). "The gene encoding ARS-binding factor I is essential for the viability of yeast." Genes Development **3**(12A):1926-39.

Robinson MD, J. Grigull, N. Mohammad, T.R. Hughes. (2002) "FunSpec: a web-based cluster interpreter for yeast". BMC Bioinformatics **3**:35.

Rodriguez Lombardero S., A. Vizoso Vazquez, E. Rodriguez Belmonte, M.I. Gonza lez Siso and M.E. Cerdan (2012). "SKY1 and IXR1 interactions, their effects on cisplatin and spermine resistance in Saccharomyces cerevisiae." Canadian Journal of Microbiology **58**(2): 184–188.

Rohde J.R., M.E. Cardenas (2003). "The tor pathway regulates gene expression by linking nutrient sensing to histone acetylation." Molecular and Cellular Biology **23**(2):629-635.

Rolland F., J.H. De Winde, K. Lemaire, E. Boles, J.M. Thevelein, J. Winderickx (2000). "Glucose-induced cAMP signaling in yeast requires both a G-protein coupled receptor system for extracellular glucose detection and a separable hexose kinase-dependent sensing process." Molecular Microbiology **38**(2):348-358.

Rolland F., V. Wanke, L. Cauwenberg, P. Ma, E. Boles, M. Vanoni, J.H. de Winde, J.M. Thevelein, J. Winderickx (2001). "The role of hexose transport and phosphorylation in cAMP signaling in the yeast *Saccharomyces cerevisiae*." FEMS Yeast Research **1**(1):33-45.

Rouillon A., Y. Surdin-Kerjan and D. Thomas (1999). "Transport of sulfonium compounds. Characterization of the s-adenosylmethionine and s-methylmethionine permeases from the yeast Saccharomyces cerevisiae." Journal of Biological Chemistry **274**(40):28096-105.

Rudra D., Y. Zhao, J.R. Warner (2005). "Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins." EMBO Journal **24**(3):533-542.

Rützler M., A. Reissaus, M. Budzowska and W. Bandlow (2004). "SUT2 is a novel multicopysuppressor of low activity of the cAMP/protein kinase A pathway in yeast." European Journal of Biochemistry **271(**7):1284-91.

Santangelo G.M. (2006). "Glucose signaling in *Saccharomyces cerevisiae*." Microbiol Microbiology and Molecular Biology Reviews **70**(1):253-282.

Schaffner W. (1988). "Gene regulation. A hit-and-run mechanism for transcriptional activation?" Nature **336**(6198):427–428.

Schawalder S.B., M. Kabani, I. Howald, U. Choudhury, M. Werner, D. Shore (2004). "Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1." Nature **432**(7020):1058-10614.

Schmelzle T., T. Beck, D.E. Martin, M.N. Hall (2004). "Activation of the RAS/cyclic AMP pathway suppresses a TOR deficiency in yeast." Molecular and Cellular Biology **24**(1):338-351.

Schmidt M.C., R.R. McCartney, X. Zhang, T.S. Tillman, H. Solimeo, S. Wölfl, C. Almonte, S.C. Watkins (1999). "Std1 and Mth1 proteins interact with the glucose sensors to control glucose-regulated gene expression in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **19**(7):4561-4571.

Schöler A. and H.J. Schüller (1994). "A carbon source-responsive promoter element necessary for activation of the isocitrate lyase gene ICL1 is common to genes of the gluconeogenic pathway in the yeast Saccharomyces cerevisiae." Molecular Cell Biology **14**(6):3613-22.

Schroeder A., O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow and T. Ragg (2006). "The RIN: an RNA integrity number for assigning integrity values to RNA measurements." BMC Molecular Biology 7:3

Schüller H.J. and K.D. Entian (1991). "Extragenic suppressors of yeast glucose derepression mutants leading to constitutive synthesis of several glucose-repressible enzymes." Journal of Bacteriology **173**(6):2045-52.

Shefet-Carasso L., I. Benhar (2001). "Antibody-targeted drugs and drug resistance-challenges and solutions." Drug Resistance Updates **18**:36-46.

Shen C., C.S. Lancaster, B. Shi, H. Guo, P. Thimmaiah, M.A. Bjornsti (2007). "TOR signaling is a determinant of cell survival in response to DNA damage." Molecular and Cellular Biology **27**(20):7007-17.

Shenhar G. & Y. Kassir (2001). "A positive regulator of mitosis, Sok2, functions as a negative regulator of meiosis in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **21**(5):1603-12.

Shevchenko A., A. Roguev, D. Schaft, L. Buchanan, B. Habermann, D. Sakalar, H. Thomas, N.J. Krogan, A. Shevchenko A, A.F Stewart (2008). "Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment." Genome Biology **9**(11):R167.

Shim E.Y., J.L. Ma, J.H. Oum, Y. Yanez, S.E. Lee (2005). "The yeast chromatin remodeler RSC complex facilitates end joining repair of DNA double-strand breaks." Molecular and Cellular Biology **25**(10):3934-44.

Shimada K., I. Filipuzzi, M. Stahl, S.B. Helliwell, C. Studer, D. Hoepfner, A. Seeber, R. Loewith, N.R. Movva, S.M. Gasser (2013). "TORC2 signaling pathway guarantees genome stability in the face of DNA strand breaks." Molecular Cell **51**(6):829-39.

Sidorova JM, L.L. Breeden (1997) "Rad53-dependent phosphorylation of Swi6 and down-regulation of CLN1 and CLN2 transcription occur in response to DNA damage in *Saccharomyces cerevisiae*". Genes Development **11**(22):3032-45.

Silverstein R.A., K. Ekwall (2005). "Sin3: a flexible regulator of global gene expression and genome stability." Current Genetics **47**(1):1-17

Simon I., J. Barnett, N. Hannett, C. T. Harbison, N. J. Rinaldi, T. L. Volkert, J. J. Wyrick, J. Zeitlinger, D. K. Gifford, T. S. Jaakkola, and R. A. Young (2001). "Serial regulation of transcriptional regulators in the yeast cell cycle." Cell **106**(6):697–708.

Singh S. K., P. Gurha, and R. Gupta (2008). "Dynamic guide-target interactions contribute to sequential 29-O-methylation by a unique archaeal dual guide box C/D sRNP." RNA **14**(7): 1411–1423.

Slattery M.G., D. Liko, W. Heideman (2008). "Protein kinase A, TOR, and glucose transport control the response to nutrient repletion in *Saccharomyces cerevisiae*." Eukaryotic Cell **7**(2):358-367.

Smets B., P. De Snijder, K. Engelen, E. Joossens, R. Ghillebert, K. Thevissen, K. Marchal, J. Winderickx (2008). "Genome-wide expression analysis reveals TORC1-dependent and -independent functions of Sch9." FEMS Yeast Research **8**(8):1276-88.

Smoot M.E., K. Ono, J. Ruscheinski, P.L. Wang, T. Ideker (2011). "Cytoscape 2.8: new features for data integration and network visualization." Bioinformatics **27**(3):431-2.

Smyth G. K. (2005). "LIMMA: linear models for microarray data In Bioinformatics and Computational Biology Solutions using R and Bioconductor" pp. 397–420.

Edited by R. Gentleman, V. Carey, S. Dudoit, R. Irizarry & W. Huber. New York: Springer.

Spellman P. T., G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher (1998). "Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization." Molecular Biology Cell **9**(12):3273–3297.

Steber C.M. & R.E Esposito (1995). "*UME6* is a central component of a developmental regulatory switch controlling meiosis-specific gene expression." Proceedings of the National Academy of Sciences U S A **92**(26):12490-4.

Stillman DJ. (2013). "Dancing the cell cycle two-step: regulation of yeast G1-cell-cycle genes by chromatin structure." Trends in Biochemical Sciences **38**(9):467-75.

Stros M. (2010). "HMGB proteins: interactions with DNA and chromatin." Biochimica et Biophysical Acta. **1799**(1-2):101-13.

Tabuchi M., A. Audhya, A.B. Parsons, C. Boone, S.D. Emr (2006). "The phosphatidylinositol 4,5-biphosphate and TORC2 binding proteins Slm1 and Slm2 function in sphingolipid regulation." Molecular and Cellular Biology **26**(15):5861-75.

Tanwar J., S. Das, Z. Fatima, S. Hameed (2014). "Multidrug resistance: an emerging crisis." Interdisciplinary Perspectives on Infectious Diseases **2014**:541340.

Teixeira M.C., P. Monteiro, P. Jain, S. Tenreiro, A.R. Fernandes, N.P. Mira, M. Alenquer, A.T. Freitas, A.L. Oliveira and I. Sá-Correia (2006). "The YEASTRACT database: a tool for the analysis of transcription regulatory associations in Saccharomyces cerevisiae." Nucleic Acids Research 34(Database issue):D446–D451.

Teixeira M.C., P. T. Monteiro, J.F. Guerreiro, J.P. Gonçalves, N.P. Mira, S.C. dos Santos, T.R. Cabrito, M. Palma, C.Costa, A.P. Francisco, S.C. Madeira, A.L. Oliveira, A.T. Freitas, I. Sá-Correia (2014). "The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae." Nucleic Acids Research. 42(Database): D161-D166.

Tenreiro S., R.C. Vargas, M.C. Teixeira, C. Magnani, I. Sá-Correia (2005). "The yeast multidrug transporter Qdr3 (Ybr043c): localization and role as a determinant of resistance to quinidine, barban, cisplatin, and bleomycin." Biochemical and Biophysical Research Communications **327**(3):952-9.

Thomas P.D., A. Kejariwal, N. Guo, H. Mi, M.J. Campbell, A. Muruganujan, B. Lazareva-Ulitsky (2006). "Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools." Nucleic Acids Research **34**(Web Server issue):W645-50.

Tizón B., A.M. Rodríguez-Torres and M.E. Cerdán (1999). "Disruption of six novel Saccharomyces cerevisiae genes reveals that YGL129c is necessary for growth in non-fermentable carbon sources, YGL128c for growth at low or high temperatures and YGL125w is implicated in the biosynthesis of methionine." Yeast 15(2): 145–154.

Tkach J.M., A.Yimit, A.Y. Lee, M. Riffle, M. Costanzo, D. Jaschob, J.A. Hendry , J. Ou, J. Moffat, C. Boone, T.N. Davis, C. Nislow C and G.W. Brown (2012) "Dissecting DNA damage response pathways by analysing protein localization and

abundance changes during DNA replication stress." Nature Cell Biology **14**(9):966-76.

Toda T., I. Uno, T. Ishikawa, S. Powers, T. Kataoka, D. Broek, S. Cameron, J. Broach, K. Matsumoto, M. Wigler (1985). "In yeast, RAS proteins are controlling elements of adenylate cyclase." Cell **40**(1):27-36.

Treiber D.K., X. Zhai, H.M. Jantzen, J.M. Essigmann (1994). "Cisplatin-DNA adducts are molecular decoys for the ribosomal RNA transcription factor hUBF (human upstream binding factor)." Proceedings of the National Academy of Sciences of the United States of America **91**(12):5672-6.

Tsaponina O., E. Barsoum, S.U. Aström and A. Chabes (2011). "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools." PLoS Genetics **7**(5):e1002061.

Vallier L.G. and M. Carlson (1994). "Synergistic release from glucose repression by mig1 and ssn mutations in Saccharomyces cerevisiae." Genetics **137**(1):49-54.

van Helden J. (2003). "Regulatory sequence analysis tools." Nucleic Acids Research 31(13): 3593–3596.

van Roermund C.W., E.H. Hettema, M. van den Berg, H.F. Tabak, R.J. Wanders (1999). "Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria in *Saccharomyces cerevisiae* and identification of a plasma membrane carnitine transporter, Agp2p." EMBO Journal **18**(21):5843-52.

Verma R, J. Smiley, B. Andrews, J.L. Campbell (1992) "Regulation of the yeast DNAreplication genes through the Mlu I cell cycle box is dependent on SWI6" Proceedings of the National Academy of Sciences U S A **89**(20):9479-83.

Vincent O. and M.Carlson (1998). "Sip4, a Snf1 kinase-dependent transcriptional activator, binds to the carbon source-responsive element of gluconeogenic genes." EMBO Journal **17**(23):7002-8.

Vinod P.K., N. Sengupta, B.J. Bhat, K.V. Venkatesh (2008). "Integration of global signaling pathways, cAMP-PKA, MAPK and TOR in the regulation of *FLO11*." PLoS One **3**(2):e1663.

Volkov K.V., A.Y. Aksenova, M.J. Soom, K.V. Osipov, A.V. Svitin, C. Kurischko, I.S. Shkundina, M.D. Ter-Avanesyan, S.G. Inge-Vechtomov, L.N. Mironova. (2002) "Novel non-Mendelian determinant involved in the control of translation accuracy in *Saccharomyces cerevisiae*". Genetics **160**(1):25-36.

Wade J.T., D.B. Hall, K. Struhl (2004). "The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes." Nature **432**(7020):1054-1058.

Warner J.R. (1999). "The economics of ribosome biosynthesis in yeast." Trends Biochemical Sciences **24**(11):437-440.

Wedaman K.P., A. Reinke, S. Anderson, J. Yates, J.M. McCaffery, T. Powers (2003). "Tor kinases are in distinct membrane-associated protein complexes in *Saccharomyces cerevisiae*." Molecular Biology of the Cell **14**(3):1204-2.

Williams R.M., M. Primig, B.K. Washburn, E.A. Winzeler, M. Bellis, C. Sarrauste de Menthiere, R.W. Davis, R.E Esposito (2002). "The Ume6 regulon coordinates metabolic and meiotic gene expression in yeast." Proceedings of the National

Academy of Sciences USA **99**(21):13431-6.

Woolford J.L. Jr, S.J. Baserga (2013). "Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*." Genetics **195**(3):643-81.

Xia L., L. Jaafar, A. Cashikar, H. Flores-Rozas (2007). "Identification of genes required for protection from doxorubicin by a genome-wide screen in *Saccharomyces cerevisiae*." Cancer Research **67**(23):11411-8.

Xu Z., D. Norris (1998). "The *SFP1* gene product of *Saccharomyces cerevisiae* regulates G_2/M transitions during the mitotic cell cycle and DNA-damage response." Genetics **150**(4):1419-28.

Yu F., Y. Imamura, M. Ueno, S.W. Suzuki, Y. Ohsumi, M. Yukawa, E. Tsuchiya (2015). "The yeast chromatin remodeler Rsc1-RSC complex is required for transcriptional activation of autophagy-related genes and inhibition of the TORC1 pathway in response to nitrogen starvation." Biochemical and Biophysical Research Communications **464**(4):1248-53.

Zaman S., S.I. Lippman, L. Schneper, N. Slonim, J.R. Broach (2009). "Glucose regulates transcription in yeast through a network of signaling pathways." Molecular Systems Biology **5**:245.

Zhai X., H. Beckmann, H.M. Jantzen, J.M. Essigmann (1998). "Cisplatin-DNA adducts inhibit ribosomal RNA synthesis by hijacking the transcription factor human upstream binding factor." Biochemistry **37**(46):16307-15.

Zhao Y., K.B. McIntosh, D. Rudra, S. Schawalder, D. Shore, J.R. Warner (2006). "Fine-structure analysis of ribosomal protein gene transcription." Molecular and Cellular Biology **26**(13):4853-62.

Zhu C., K.J. Byers, R.P. McCord, Z. Shi, M.F. Berger, D.E. Newburger, K. Saulrieta, Z. Smith, M.V. Shah, M. Radhakrishnan, A.A. Philippakis, Y. Hu, F. De Masi, M. Pacek, A. Rolfs, T. Murthy, J. Labaer, M.L. Bulyk. (2009) "High-resolution DNA-binding specificity analysis of yeast transcription factors". Genome Research **19**(4):556-66.

Zinzalla V., D. Stracka, W. Oppliger, M.N. Hall (2011). "Activation of mTORC2 by association with the ribosome." Cell **144**(5):757-68.

Zitomer R. S. and B.D. Hall (1976). "Yeast cytochrome c messenger RNA. In vitro translation and specific immunoprecipitation of the CYC1 gene product." Journal of Biological Chemistry **251**(20): 6320–6326.

Chapter 4

Characterization of the DNA-binding properties and molecular dynamics of the two in tandem HMG-boxes of Ixr1

SUMMARY

Ixr1 from *Saccharomyces cerevisiae* is a protein that contains two DNA-binding HMG-box domains, A and B, linked by only three amino acids and flanked by several polyglutamine regions. In the present work, we demonstrate biochemical and folding differences between both HMG-box domains. The results suggest that HMG-box A, more stable and with higher binding affinities, is necessary to promote the optimal conformation of HMG-box B, thus favouring binding to linear DNA in a positive cooperativity fashion. Furthermore, our results show differences in the binding mode of each HMG-box domain to cisplatin-modified DNA.

1.- INTRODUCTION

In eukaryotes, the number of regulated genes far exceeds the number of regulatory proteins in the cell. To overcome this numerical barrier, these proteins associate and they form intricate networks of regulatory combinatorial units. Furthermore, the DNA binding domains present in the regulatory proteins are flexible and versatile regions that can recognize regulatory elements in the DNA in a wide variety of ways. There is a certain degree of relaxation in the specificity of the DNA sequence recognition by the regulatory protein, which is allowed by the flexibility caused by the conformational adaptation during the interaction process between the two molecules (Wright & Dyson, 1999; Li et al., 2003).

The high mobility group (HMG) proteins are abundant non-histone nuclear proteins that associate with chromatin and contain a good and interesting example of a versatile eukaryotic DNA binding domain, the HMG-box. Their associations with DNA are not confined to specific sites but they have a highly dynamic behavior, in a "hit and run" fashion, scanning the potential chromatin binding sites and moving from one chromatin site to another (Schaffner, 1988; Reeves et al.,

2010; Malarkey and Churchill, 2012).

The structures of various HMG-boxes are well studied and it is known that their folding is far more conserved than the amino acid sequences. In general, the HMG-box domain contains 65-85 amino acids and has a characteristic L-shaped fold formed by three α-helices with an angle of ≈80° between the two arms. The long arm or minor wing is composed by the extended N-terminal strand and third α-helix, while first and second α-helix form the short arm, or major wing. Its small size, unique and simple fold pattern, together with its structural conservation in a huge number of proteins from different species, make the HMG-box domain a good model for general studies of protein folding and stability (Thomas & Travers, 2001; Stros, 2010).

There are two broad subfamilies of HMG-box containing proteins, based on structural and phylogenetic studies. One class includes those that bind to distorted DNA with low or without sequence specificity (_Non Sequence Specificity_, NSS, HMG-box domains) (Grosschedl _et al._, 1994; Bustin and Reeves, 1996) and have, in general, two or more in tandem arranged HMG-box domains. Examples of proteins without sequence specificity are the mammalian HMGB1-4 and UBF proteins, HMGD from _Drosophila_ or Nhp6 from _Saccharomyces cerevisiae_. Their role is related to chromatin modification, participating in many different functions such as co-activation or silencing of transcription and V(D)J junction recombination. A second class of HMG-box containing proteins bind to DNA by recognizing a specific DNA sequence (_Sequence Specificity_, SS, HMG-box domains) (Grosschedl _et al_., 1994; Bustin and Reeves, 1996) and they usually contain a single HMG-box domain. They generally function as transcription factors, only expressed in a few cell types, and they also contain other regulatory associated domain. The determinants for DNA sequence specificity lie mainly in the minor wing of the HMG-box. Examples of this kind of HMGB proteins are the mammalian lymphoid enhancer factor (Lef-1), the sex determining factor (Sry) and the Sry-related HMG-

box (SOX) family; or the hypoxic gene repressor (Rox1) from *S. cerevisiae*.

Despite these differences, both subfamilies of HMG-box proteins are able to bind to B-form DNA through the minor groove with high affinity and they induce a large DNA bending, ultimately forming complexes of rather similar structure. They use their concave surface to intercalate one or two (SS and NSS HMG-box domains, respectively) bulky hydrophobic amino acids between base-pairs in the minor groove. Other extensive protein-DNA contacts, not sequence-specific as well as hydrogen bonds, are made,with the phosphate backbone. As a result of these interactions, the DNA is bent and under-wound with positive roll angles, widening the minor groove and compressing the major groove (Thomas & Travers, 2001; Stros, 2010).

The *IXR1* gene of *S. cerevisiae* encodes for a protein of 67.2 kDa that contains three poly-glutamine regions and two HMG-box domains that bind to DNA (Lambert *et al*., 1994). Among the yeast HMGB proteins, Ixr1 protein is the least known, concerning structure and functions. Its named from Intra-strand cross-(X)-link Recognition, because it was first characterized during screenings looking for structure-specific recognition proteins (SSRPs) that bind to DNA containing intrastrand cross-links formed by the anticancer drug cisplatin. It has been suggested (McA'Nulty & Lippard, 1996) that Ixr1 protein would be involved in the cytotoxicity caused by the drug. In the proposed mechanism, the binding of Ixr1 protein to the cisplatin-DNA adduct through its HMG-box domains, could mask these adducts, thus preventing the action of the *Nucleotide Excision Repair* (NER) DNA repair system. Ixr1 protein also behaves as a transcription factor in response to other stress signals such as low oxygen levels (Vizoso Vázquez *et al*., 2012) or the presence of reactive oxygen species, ROS, (Castro-Prego *et al*., 2010a). Recently it has been reported that cisplatin resistance observed in *Δixr1* mutants depends on pre-activation of the control systems of genomic integrity through the Mec1-Rad53-Dun1 signalling pathway, which regulate transcription of the gene *RNR1*,

encoding the large subunit of ribonucleotide reductase, DNA damage inducible and necessary to maintain the dNTP pools for DNA synthesis and repair (Tsaponina *et al*., 2011; Tsaponina *et al*., 2013). Besides, Ixr1 protein deregulation and chromatin decoupling can induce a necrotic cell death response through the TORC1 signaling pathway as a consequence of nutrient stress (Chen *et al*., 2013).

In the present work, we apply *in vitro* assays to characterize the binding properties of the two HMG-boxes present in Ixr1 protein (P33417) at positions 360-430 (called HMG-box A) and 433-503 (called HMG-box B), and to determine their role in the function of the protein in yeast cells. We used isothermal titration calorimetry (ITC) and fluorescence anisotropy to characterize the thermodynamic aspects of Ixr1 protein binding to different sequence specific double-stranded DNAs, cisplatin-modified DNA and four-way junction cruciform DNA. Previous studies demonstrated that the existence of two HMG-box domains in tandem is crucial to determine the specific functions of proteins, such as HMGB1 (Müller *et al*., 2001) and TFAM (Rubio-Cosials *et al*., 2011) from human or Hmo1 from *S. cerevisiae* (Kamau *et al*., 2004). To understand the significance of the presence of two HMG-box domains in the Ixr1 protein, we have compared the interactions with DNA for the full-length protein, for the domain including the tandem of two HMG-boxes and for the single HMG-box domains A and B separately. To this purpose EMSA assays, fluorescence anisotropy (FA), Isothermal Titration Calorimetry (ITC) and Fluorescence Resonance Energy Transfer (FRET) techniques have been employed. In addition, nuclear magnetic resonance (NMR) and circular dichroism (CD) experiments were made trying to understand the protein dynamics of the Ixr1 HMG-box domains in presence or absence of DNA.

2.- MATERIAL AND METHODS

2.1.- Cloning, expression and protein purification

The constructs encoding full-length protein (residues 1-597) and tandem-AB HMG-box didomain (residues 338-510) were produced by PCR amplification and subsequent ligation. DNA oligonucleotides were purchased from Isogen Lifes Sciences, Inc. The primers used were as follows in table 1.

Table 1. Oligonucleotides used in this study to DNA clonning.

Name	Sequence[a,b]	Strand (W/C)[c]	Added site	Hybridization position[d]
ECV681ixr1f	cgga**gagctc**AACACCGGTATCTCGCCC	W	*SacI*	+4
ECV682r	cggc**ctcgag**TTATTCATTTTTTATGATCGAACC	C	*XhoI*	+1794
ECV683hmgf	cgga**gagctc**CCAGTGGTGAAGAAATTATCTTC	W	*SacI*	+1012
ECV785AV	cggc**ctcgag**GGTTGGGTTACCGTTTGG	C	*XhoI*	+1527
HMGAF	aagagaccctctggtcccTTTATTCAGTTCACCCAGG	W	-	+1312
HMGAR	gggaccagagggtctcttTTATGGGGGCAAAGTCTTTTCG	C	-	+1284
HMGBF	ggcgaattcgagctcaagtacagaGTTGTGAGAGATGCTTA	W	-	+1256
HMGBR	gatctgtacttgagctcgaaTTCGCCCTGGAAATAC3AGG	C	-	+1048

[a]Lower case, sequences added for corrected restriction enzyme digestions.
[b]Restriction sites and homology regions in bold.
[c]Numbering is considering +1 for the adenine in the first start codon.
[d]W: Watson strand; C: Crick strand.

PCR reactions were performed using Vent polymerase (*NewEngland Biolabs*) and genomic DNA from *Saccharomyces cerevisiae* BJ350 strain as template. The PCR products were gel-purified, and PCR ligated into the pKLSL150 plasmid (Mancheño *et al.*, 2005) between *SacI* and *XhoI* restriction sites. The

resulting plasmids were designated pKLSL150-Ixr1 and pKLSL150-tandemAB. The constructs encoding individual HMG-box domains A (residues 338-439) and B (residues 430-510) were produced by the one-step PCR-based method described by Qi and co-workers (Qi *et al.*, 2007) using the primers described in table 1. PCR reactions were performed using Vent polymerase (*NewEngland Biolabs*) and the pKLSL150-tandemAB as template. All constructions were verified by DNA sequencing.

The proteins were expressed in BL-21(DE3) cells (Novagen) transformed with the appropriate plasmid, except full-length protein, which was expressed in Rosetta[TM] 2(DE3)pLysS cells (Novagen). Production of overexpressed proteins was made cultivating the cells in 2xTY medium, supplemented with 33 μg/ml kanamycin, and expression was induced with 1 mg/mL IPTG during 3 hours at 37 °C and 200 rpm of shaking. After expression, cell pellets were collected and lysed by sonication in high salt lysis buffer (50 mM sodium phosphate buffer, pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 2X complete protease inhibitor cocktail, Roche). After clarification by centrifugation 30 minutes at 23000 x g, lysates were passed through Sepharose CL-6b resin (Sigma-Aldrich) packed into a Tricorn® column (GE Healthcare) equilibrated in wash buffer A (50 mM sodium phosphate buffer, pH 6.9, 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) and the aid of a peristaltic pump (GE Healthcare). Proteins were eluted in an AKTAprime plus system (GE Healthcare) by linear gradient from 0% to 100% of buffer elution A (50 mM potassium phosphate buffer, pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 300 mM lactose) and loaded on a HisTrap HP 5 mL column (GE Healthcare) equilibrated in wash buffer B (50 mM sodium phosphate buffer, pH 6.9, 200 mM NaCl, 2 mM dithiothreitol). After elution by linear gradient from 0% to 100% of buffer elution B (50 mM potassium phosphate buffer, pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 300 mM imidazole), the different polypeptides were dialyzed (50 mM sodium phosphate buffer, pH 6.9, 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) and simultaneously digested with TEV protease (Sigma-Aldrich) for 16 hours and 8 °C.

Next day, proteins digested were further purified using gel filtration chromatography using a Hi-load Superdex 200 16/60 column (GE Healthcare) pre-equilibrated with running buffer (50 mM potassium phosphate buffer, pH 6.9, 100 mM KCl, 2 mM dithiothreitol and 1 mM EDTA). The proteins were concentrated by ultrafiltration using Amicon® Ultra-15 Centrifugal Filters, 3 kDa (Merk-Millipore). The homogeinity of the purified protein sample was checked by SDS-PAGE (Laemmli, 1970) and protein concentrations were determined by absorbance at 280 nm using extinction coefficients calculated with Protparam (http://web.expasy.org/protparam/).

To obtain the ^{15}N labeled proteins used for the NMR experiments, *Escherichia coli* BL21(DE3), cells containing the different pKLSL150 constructions were grown with shaking at 200 rpm at 37 ^{o}C in MOPS minimal medium prepared as previously described (Neidhardt *et al.*, 1974) to an absorbance at 600 nm of 0.6, and expression was induced with 1 mM IPTG overnight. For the preparation of ^{15}N-labeled protein, 100 mM ^{15}NH$_4$Cl was used as the sole nitrogen source as appropriate. After expression, cell pellets were collected and lysed by passing four times through a French press in high salt lysis buffer (50 mM sodium phosphate buffer, pH 6.8, 1 M NaCl, 2 mM dithiothreitol and 2X complete protease inhibitor cocktail, Roche). After centrifugation 1 h at 23000 x g, lysates were passed through Sepharose CL-6b resin (Sigma-Aldrich) packed into a Tricorn® column (GE Healthcare) equilibrated in wash buffer A (50 mM sodium phosphate buffer, pH 6.8, 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) with a peristaltic pump (GE Healthcare). Proteins were eluted in an ÄKTA FPLC system (GE Healthcare) by linear gradient from 0% to 100% of buffer elution A (50 mM potassium phosphate buffer, pH 6.8, 1 M NaCl, 2 mM dithiothreitol and 300 mM lactose) and loaded on a HisTrap HP 5 mL column (GE Healthcare) equilibrated in wash buffer B (50 mM sodium phosphate buffer, pH 6.8, 200 mM NaCl, 2 mM dithiothreitol). After elution by linear gradient from 0% to 100% of buffer elution B (50 mM potassium

phosphate buffer, pH 6.8, 1 M NaCl, 2 mM dithiothreitol and 300 mM imidazole), the different polypeptides were dialyzed (50 mM sodium phosphate buffer, pH 6.8, 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) and digested with TEV protease (Sigma-Aldrich) at the same time for 16 hours and 8 ^{o}C. Next day, proteins digested were further purified using a Heparin HP 5 mL column (GE Healthcare) pre-equilibrated with wash buffer C (20 mM sodium phosphate buffer, pH 6.8, 50 mM KCl, 2 mM dithiothreitol and 1 mM EDTA) and extensively dialyzed (10 mM sodium phosphate buffer, pH 6.5, 50 mM KCl, 2 mM dithiothreitol and 1 mM EDTA). The proteins were concentrated by ultrafiltration using Amicon$^®$ Ultra-15 Centrifugal Filters, 3 kDa (Merk-Millipore). The homogeneity of the purified protein sample was checked by SDS-PAGE (Laemmli, 1970) and protein concentrations were determined by absorbance at 280 nm using extinction coefficients calculated with Protparam (http://web.expasy.org/protparam/).

2.2.- Oligonucleotide annealing

DNA oligonucleotides were purchased from Isogen Lifes Sciences, Inc. and sequences are summarizes in table 2. Concentrations of single strands and duplexes were determined from the A_{260} of the nucleotides. For determination of the concentration of labeled single strands the contribution of FAM and TAMRA absorption at 260 nm (28,000 M^{-1} cm^{-1} and 29,000 M^{-1} cm^{-1}, respectively) were taken into account. DNA duplexes were prepared by mixing the complementary oligonucleotides in equimolar amounts, heating to 95 ^{o}C for 5 minutes and cooling slowly to room temperature in darkness. Solutions of DNA for the experiments were prepared by extensive dialysis against the solvent, as required.

DNA platination was made following the protocol described by Cohen and col. (Cohen *et al.*, 2000). Activated cisplatin was obtained by mixing two equivalents of silver nitrate with one equivalent of cisplatin in water overnight. The mixture was protected from light and centrifuged to remove precipitated silver

chloride. To obtain the platinated DNA duplex, AVV219 oligonucleotide was first de-protected with ammonium hydroxide by incubating the crude reaction mixtures at 65 $^{\circ}$C for 1 h and the excess of reaction was eliminated by several washes through a 3 kDa Amicon concentrator (Millipore). After that, AVV219 oligonucleotide with isolated GG sites was platinated with 2.5 equivalents of activated cisplatin in water solution and was incubated at 37 $^{\circ}$C overnight in darkness. Excess of cisplatin was eliminated extensively by several washes through a 3 kDa Amicon concentrator (Millipore).

Table 2. Oligonucleotides used in this study as DNA ligand.

Name	Sequence[a]	Strand (W/C)[b]	Added fluorophore	Gene
AVV190	AGGGCCTATTGTTGCTGCCT	W	5' FAM	ROX1
AVV191	AGGCAGCAACAATAGGCCCT	C	-	ROX1
AVV209	AGGGCCTATTGTTGCTGCCT	W	-	ROX1
AVV210	AGGGCCT**TGCGCA**GCTGCCT	W	5' FAM	ROX1
AVV211	AGGCAGC**TGCGCA**AGGCCCT	C	-	ROX1
AVV212	GTGTTGAACGGTTCACAGCG	W	5' FAM	HEM13
AVV213	CGCTGTGAACCGTTCAACAC	C	-	HEM13
AVV214	AATTTCAATTGTTTAGAAAG	W	5' FAM	HEM13
AVV215	CTTTCTAAACAATTGAAATT	C	-	HEM13
AVV218	TTAGTCTAGGCCTTCTATT	W	5' FAM	-
AVV219	AATAGAAGGCCTAGACTAA	C	-	-
AVV228	AGGCAGCAACAATAGGCCCT	C	5' TAMRA	ROX1
AVV234	GAAGAGCGATATCGCGTGCG	W	5' FAM	-
AVV235	CGCACGCGATATCGCTCTTC	C	-	-
AVV249	CGCAATCCTGAGCACG	-	5' FAM	-
AVV250	CGTGCTCACCGAATGC	-	-	-
AVV251	GCATTCGGACTATGGC	-	-	-
AVV252	GCCATAGTGGATTGCG	-	-	-
AVV210NO	AGGGCCT**TGCGCA**GCTGCCT	W	-	ROX1
AVV212NO	GTGTTGAACGGTTCACAGCG	W	-	HEM13
AVV214NO	AATTTCAATTGTTTAGAAAG	W	-	HEM13
AVV218NO	TTAGTCTAGGCCTTCTATT	W	-	-
AVV234NO	GAAGAGCGATATCGCGTGCG	W	-	-
AVV249NO	CGCAATCCTGAGCACG	-	-	-

[a]*ROX1* promoter region mutated in bold.
[b]W: Watson strand; C: Crick strand.

2.3.- Far-UV circular dichroism

Circular dichroic measurements were performed on a JASCO spectropolarimeter (J-815) with a thermostatically controlled cell holder attached to a Peltier PTC-423S system. All proteins were extensively dialyzed (10 mM K_2HPO_4 pH 6.8, 50 mM NaF, 2 mM DTT, 1 mM EDTA). Spectra were collected at 5 ^{o}C, in continuous scanning mode, with a scanning rate of 50 nm/s and a response time of 2 s. Each spectrum was the average of ten scans. Far-UV CD spectra were taken in the range of 190-260 nm with 1 nm steps in a cell of 0.1 cm path length. From raw data, collected spectra were buffer-subtracted and converted from millidegrees (Θ_{obs}, mdeg) to molar ellipticity ([Θ], degrees cm^2 $dmol^{-1}$ $residue^{-1}$) using the equation $[\Theta] = (\Theta_{obs} \times M) / (10 \times l \times C)$, where M is the protein mean residue molecular weight, l is the optical path length of the cuvette in centimeters, and C is the concentration of the protein in mg mL^{-1}. The percent secondary structure content for alpha class proteins was calculated by K2D2 method (Perez-Iratxeta et al., 2008) for α-helix class proteins.

The temperature dependence of the circular dichroism spectra of the different proteins were determined upon continuous heating, with a rate of 1 K min^{-1} in the temperature range 5-95 ^{o}C at 222 nm wavelength using a 0.1 cm path length sealed cell. Ellipticity signals were plotted as a function of temperature and a non-linear Boltzmann fit was performed, and melting temperature was calculated as the maximum of the first derivative of this curve using GraphPad Prism 6.0 (GraphPad software).

2.4.- Gel Mobility Shift assays

Reaction mixtures (15 µl) containing 10 nM DNA 5' labeled with FAM were made in 20 mM Potassium phosphate buffer, pH 6.8, 100 mM KCl, 1 mM EDTA, 2 mM DTT, 5% (v/v) glycerol, 500 µg/ml bovine serum albumin, and protein at various concentrations, as indicated. The mixtures were kept on ice for 30 minutes,

and then analyzed in 5% polyacrylamide gels (in TBE buffer, 100 mM Tris base, 100 mM boric acid, 2 mM EDTA) that had been pre-run at 90 V for 30 minutes at 4 $^{\circ}$C. Electrophoresis was carried out at 4 $^{\circ}$C and 90 V in a MiniPROTEAN Tetracell system (Bio-Rad), and gels were scanned for fluorescence in a Typhoon^{TM} FLA 7000 biomolecular imager v.1.2 (GE Healthcare) to detect the FAM fluorophore, using 473 nm laser excitation and a filter Y520. Band density quantification was performed in ImageQuant TL v8.1 softwate (GE Healthcare). The fraction of DNA bound in each reaction was plotted versus the protein concentration. The data were fit with the following binding equation using Prism software to perform non-linear regression and obtain a value for K_d.

$$Y = B_{max} \cdot X^h/(K_d{}^h + X^h)$$

where Y are the fluorescence polarization values, B_{max} is the maximum value of Y, X are the concentration values of ligand (in our case, different DNAs), K_d is the dissociation constant (in the same units that X), and h is the hill slope.

2.5.- Fluorescence Anisotropy titrations

Fluorescein-labeled (FAM) dsDNA oligonucleotides (see table 2: AVV190-AVV191, AVV210-AVV211, AVV212+AVV213, AVV214-AVV215, AVV218-AVV219 and AVV249-AVV250-AVV251-AVV252) were extensively dialyzed (10 mM K$_2$HPO$_4$ pH 6.8, 100 mM KCl, 2 mM DTT, 1 mM EDTA, 500 µg/ml bovine serum albumin). Fluorescence anisotropy titrations were performed at 25 $^{\circ}$C on a Multi-modal Synergy^{TM} H1 plate reader (Biotek$^{®}$) using 384-well Well Low Volume Black Round Bottom Polystyrene NBS™ Microplate (Corning) with 15 µl per well. The excitation and detection wavelengths were 485 and 528 nm, respectively, with dichroic mirror (510 nm) and polarizer filter assembled. Tumbling rates or changes in the rotational times of the small labelled-DNAs when are tightly bound to large proteins were used to calculate fluorescence anisotropy values. In each titration

the fluorescence anisotropy of a solution of 50 nM fluorescein-tagged duplex DNA was measured and represented in a porcentaje of ligand bound as a function of the added protein concentration. For each competition experiment, the polarization signal was followed over time. As a result, a 30 minutes incubation period was selected as an adequate time to reach equilibrium (data not shown). Binding data were fitted to a simple one site saturation-binding model by nonlinear least squares regression using GraphPad Prism 6.0 (GraphPad software). Each titration was performed three times, and the final affinity was taken as the mean of these measurements.

Gibbs energies of association were derived from the association constants obtained by fluorescence anisotropy following the equation:

$$\Delta G = -RTLn(K_a)$$

where R is the gas constant (R = 8.314472 J L^{-1} mol^{-1}) and T in temperature (T= 288.15 K).

Salt-titration experiments were carried out by adding concentrated KCl solution into protein/DNA solution. Volumen displacements were taken into account, as well as the slight dependence of the anisotropy of free DNA on salt concentration, measured in independent experiments. For fitting, the following equation was used:

$$log(K_a)= log(K_a^{nel}) - Z\Psi log[KCl]$$

where K_a^{nel} is the association constant at the standard state in 1 M KCl, Z is the number of DNA phosphate groups that interact with the peptide and Ψ (0.64 for short DNA duplexes) is the number of cations associated with phosphate groups. K_a^{nel} and ZΨ were treated as fitting parameters (Olmsted *et al*., 1995; Dragan *et al*., 2003; Dragan *et al*., 2004; Privalov *et al*., 2007). This permits splitting the Gibbs energies of binding into two components the non-electrostatic and the

electrostatic components:

$$\Delta G_a = \Delta G_a^{nel} + \Delta G_a^{el}$$

2.7.- Isothermal Titration Calorimetry

ITC experiments were conducted with a MicroCal ITC-200 machine (GE Healthcare). Protein and DNA samples were extensively dialyzed against 3 L of buffer (10 mM K_2HPO_4, pH 6.8, 100 mM KCl, 0.5 mM TCEP) and degassed before each measurement. Different DNA samples at suitable concentrations (ranging from 100 µM to 500 µM) were titrated in 2.5 µL injections to the cell containing the protein at different concentrations (ranging from 20 µM to 40 µM). The raw heats of injection were measured while the cell was stirred at 1000 rpm at 25 ^{o}C. As a control, DNA was titrated into buffer, and the heat evolved was subtracted from the heats for the protein/DNA injections. The heat of injection measured from the first titration point was discarded. Raw heats of injection were baseline corrected, integrated with respect to time, and fit to a binding isotherm to calculate the dissociation constants and energies of enthalpy of the reaction using Origin® v7 SR4 scientific plotting software (OriginLab Corporation). Macroscopic cooperative constant (ρ) was obtained from the equation:

$$\rho = 4\beta_2 / \beta_1^2$$

Electrostatic (el) and non electrostatic (nel) components of association where then calculated with the equations:

$$\Delta G_a^{nel} = \Delta H_a - TAS_a^{nel}$$

$$\Delta G_a^{el} = - TAS_a^{el}$$

2.8.- Fluorescence Resonance Energy Transfer assays

Experiments were conducted with DNA^{ROX1} duplex template, that was 3´-labeled in one strand with TAMRA (AVV285, acceptor) and with FAM (AVV190,

donor) in the other one, and they were extensively dialyzed (10 mM K_2HPO_4 pH 6.8, 100 mM KCl, 2 mM DTT, 1 mM EDTA, 500 µg/ml bovine serum albumin). Fluorescence anisotropy titrations were performed at 25 $^{\circ}$C on a Multi-modal Synergy^{TM} H1 plate reader (Biotek$^{®}$) using 384-well Well Low Volume Black Round Bottom Polystyrene NBS™ Microplate (Corning) with 15 µl per well.

FRET effect (FE) is calculated from the following equation:

$$FE = I_{490}/I_{560}$$

where I_{490} is the fluorescence emission of the extracted TAMRA signal at 580 nm when excited at 490 nm, and I_{560} is the fluorescence emission of TAMRA at 580 nm when excited at 560 nm where FAM does not absorb. The extracted TAMRA signal is obtained by fitting the spectrum of the appropriate 3′-FAM only labeled DNA excited at 490 nm to the donor region of the of the dual labeled DNA spectra exited at 490 nm and subtracting the fluorescence intensity of the 3′-FAM only labeled DNA from the dual labeled DNA. FE can be used to measure the distance between the FAM donor and the TAMRA acceptor since the distance between donor and acceptor varies as a function of the sixth power of the efficiency of FRET (E) (Clegg, 1992; Stuhmeier *et al*., 2000), as follows:

$$FE = E[\varepsilon FAM_{490}/\varepsilon TAMRA_{560}] + \varepsilon TAMRA_{490}/\varepsilon TAMRA_{560}$$

The values of $\varepsilon FAM_{490}/\varepsilon TAMRA_{560}$ and $\varepsilon TAMRA_{490}/\varepsilon TAMRA_{560}$ are constants and were calculated to be 0.149 and 0.071 for DNA^{ROX1} duplex, respectively. Distance between the FRET donor and acceptor (D_{da}) can be calculated from the following equation:

$$D_{da} = ((1/E)-1)^{1/6} \cdot D_r$$

where D_r is the Förster radius for the two fluorophores used in this experiment. D_r value calculated for DNA^{ROX1}, as previously described (Clegg, 1992; Dragan *et al*.,

2003), corresponds to 49.8 Å. In the absence of protein, the calculated D^o_{da} value for DNA^{ROX1} was 75 Å. If one assumes that the upon protein binding, the DNA is bent at the middle of the DNA sequence, an estimate of the bend angle can be made using the following equation:

$$Bend\ angle = 2cos^{-1}(D_{da}/D^o_{da})$$

where D_{da} is the end-to-end distance in the presence of a given concentration of protein, and D^o_{da} is the end to end distance in the absence of protein.

2.9.- NMR spectroscopy

NMR measurements were made on samples containing 0.5 mM ^{15}N-labeled protein or ^{15}N-labeled protein-DNA complex, 10% ^2H_2O in 10 mM sodium phosphate (pH 6.8), 0.1 mM EDTA and 1 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) (Sigma-Aldrich). 1D ^1H and ^1H-^{15}N HSQC spectra experiments were carried out at 298 K on a Bruker Avance III AV600 spectrometer equipped with a quad-resonance HCNF probe head and actively shielded z-gradients. Data were processed using the AZARA suite of programs (v. 2.8, © 1993-2013; Wayne Boucher and Department of Biochemistry, University of Cambridge, unpublished).

2.12. Bioinformatics analysis and 3-D modeling

The Uniprot (http://www.uniprot.org) accession numbers of proteins used in Figure 1 are P33417 (Ixr1 protein from *Saccharomyces cerevisiae*), Q00059 (Tfam from *Homo sapiens*), Q02486 (Abf2 from *Saccharomyces cerevisiae*), Q24537 (Dsp1 from *Drosophila melanogaster*), P09429 (HMGB1 from *Homo sapiens*), P26583 (HMGB2 from *Homo sapiens*), O15347 (HMGB3 form *Homo sapiens*) and Q8WW32 (HMGB4 from *Homo sapiens*). Alignment was done by ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) using the PAM matrix and edited by ESPript 3 (http://espript.ibcp.fr/ESPript/ESPript/index.php). Phylogenetic tree was

constructed using Neighbor-Joining method. 3-D Homology modeling of Ixr1 HMG-box domains were done by Phyre2 server (Kelley *et al.*, 2015) (http://www.sbg.bio.ic.ac.uk/phyre2/html/ page.cgi?id=index) based on protein templates human HMGB1 (PDB 2YRQ), Tox2 protein from Mus musculus (PDB 2CO9) and human HMGB (PDB 2E6O) in Protein Data Bank (http://www.rcsb.org). Protein model pictures were made by PyMol package (v1.7) (www.pymol.org). All figures were created or edited by Adobe Illustrator CS6 v16.0.0 (Adobe systems).

3.- RESULTS

3.1.- Contributions of the HMG-box domains A and B of Ixr1 to formation of DNA complexes

3.1.1.- Comparative alignment of the HMG-boxes of Ixr1 with those from other proteins with two HMG-boxes

A multiple alignment of the two Ixr1 HMG-box domains with other proteins that contain also two HMG-box domains in tandem was made, including the paralog ARS-binding factor 2 (Abf2) in *Saccharomyces cerevisiae*, the dorsal repressor DSP1 of *Drosophila melanogaster* and human HMGB1-4 proteins (figure 1). In the Ixr1 protein, the first box (HMG-box A) extends between amino acids 360 and 430 and the second (HMG-box B) between 433 and 503, sharing only 24.63% identity and 44.92% similarity between them. Both Ixr1 and Abf2 proteins distinguish from the others in the alignment because they present a short linker region (only three and four amino acids, respectively) connecting both HMG-box domains (figure 1a, green box).

The HMG-box domains of both sequence-specific (SS) and nonsequence-specific (NSS) proteins usually have a non-polar intercalating residue at the beginning of first α-helix (figure 1, represented in magenta). In addition, as opposed to SS HMG-box, NSS HMG-box domains have a second extra intercalating

non-polar residue that is exposed at the amino-terminal side of the second α-helix (figure 1, represented in green), meanwhile at the same position SS HMG-box domains contain a polar residue that forms extra base-specific contacts with the DNA by hydrogen bonds. Alignment of Ixr1 protein with the other selected HMGB proteins shows two hydrophobic amino acids, putative DNA-intercalating residues, in each HMG-box (F369, in purple, and V388, in green, for HMG-box A; I442, in purple, and L461, in green, for HMG-box B of Ixr1 protein).

It is interesting to note that, although Ixr1 protein preferably binds to cisplatin-modified DNA, thus producing cellular sensitization to the drug (Brown *et al.*, 1993; McA'Nulty and Lippard, 1996; McA'Nulty *et al.*, 1996; Rodríguez Lombardero *et al.*, 2012), it lacks the phenylalanine present in hHMGB1 (at position 37) (Ohndorf *et al.*, 1999) or Nhp6A (at position 48) (Wong *et al.*, 2002), whose aromatic side chain located at the N-terminus of second α-helix intercalates into the 1,2-intrastrand DNA adduct formed by cisplatin (figure 1, represented in yellow). The phenyl ring of the F37 residue in hMGB1 acts as a wedge, by stacking onto the solvent exposed face of the DNA adduct, and therefore a F37A mutant is severely impaired in its ability to bind cisplatin-modified DNA (Ohndorf *et al.*, 1999). Instead of an aromatic amino acid, both HMG-box domains of Ixr1 possess in the second intercalation site an hydrophobic residue with aliphatic side chains (V388 and L461) that were previously demonstrated to be involved in DNA intercalation, as described in HMG1 (Toney *et al.*, 1989; Pil & Lippard, 1992; Brown *et al.*, 1993; Chow *et al.*, 1994; Lambert *et al.*, 1994) SSRP1 (Chow *et al.*, 1994), mtTFA (Chow *et al.*, 1994), and hUBF (Treiber *et al.*, 1994), but nothing is known about the binding mechanism with cisplatin-modified DNA .

Summarizing, the comparative alignment reflects that both HMG-box domains of Ixr1 protein show some typical characteristics of NSS HMG-box domains, although the absence of perfect fitting in the intercalating residues rises the interest to look for new features that could permit clarify their origin and

function.

3.1.2.- *In vitro* characterization of the interactions of the HMG-boxes from Ixr1 with different forms of DNA

In order to characterize the role of each of the HMG-box domains of the Ixr1 protein for DNA binding, in terms of thermodynamic properties and specificity, the region containing the HMG-box domains in tandem (called tandem-AB) and the two single HMG-box domains A and B were expressed and purified (for detailed composition and properties of the Ixr1 regions analyzed see figure 2a and table 3). The solubility for the purified HMG-box domain A was low, once the protein was separated from the lectin fused moiety used as a tag for purification (see material and methods section); this is probably consequence of a bad or partial folding of the incomplete protein. This problem was resolved adding the TLPPKRPSG sequence (431-439 amino acids) to the amino terminal end in the construct. Figure 2b shows the purified bands of all the proteins selected for this study, all reaching high purity (more than 95%).

Table 3. Summary table of the different polypeptides used in the present work

Protein	Position length[a]	Molecular Weight (kDa)	pI	Concentration of the purified polypeptide (M)
Tandem-AB	338-510	20.42	9.59	333×10^{-6}
HMG-box A	338-439	12.13	10.18	309×10^{-6}
HMG-box B	430-510	9.47	8.03	1473×10^{-6}

[a]P33417 protein reference from Uniprot database.

(a)

(b)

⬅   **Figure 1.** (a) Alignments of double HMG-box domains arranged in tandem. Including Ixr1 protein from *S. cerevisiae* (Uniprot code P33417), Tfam from *Homo sapiens* (Uniprot code Q00059), Abf2 from *S. cerevisiae* (Uniprot code Q02486), Dsp1 from *D. melanogaster* (Uniprot code Q24537), HMGB1 from *H. sapiens* (Uniprot code P09429), HMGB2 from *H. sapiens* (Uniprot code P26583), HMGB3 form *H. sapiens* (Uniprot code O15347) and HMGB4 from *H. sapiens* (Uniprot code Q8WW32). Primary site (purple font) intercalating residues and secondary site (green font) intercalating/hydrogen bonding residues are indicated. Phenylalanine intercalating residues are shown in yellow font. Conserved residues are shown in red font. First HMG-box domain (HMG-box domain A) is highlighted with a green box and second HMG-box domain (HMG-box domain B) is highlighted as an orange box. Linker region falls into the grey box. Residue numbering follows the original positions of full-length Ixr1 protein (upper-side numbers) and the rest of proteins used in the alignment (right-side numbers). (b) Cartoon representation of the L-shaped domain structure of HMG-box A (green) and HMG-box B (orange), showing the amino acids that could act as the DNA intercalators F369 and I442 in purple sticks, and V388 and L461 in green sticks. HMG-box A and HMG-box B were modeled by Phyre2 server using Tox2 protein from Mus musculus (PDB 2CO9) and HMGB1 from human (PDB 2E6O) as templates.

The DNA sequences used to study the interaction with the Ixr1 protein and its HMG-box domains are presented in figure 3. They have been selected to represent different DNA forms, which have been previously associated to HMGB protein interactions specific for recognition of DNA sequences or DNA structures. These are: i) B-form DNA duplexes from sequences present in the promoter regions of the *ROX1* (Castro-Prego *et al.*, 2010a) and *HEM13* (Castro-Prego *et al.*, 2010b), which are genes regulated by Ixr1; ii) DNA^{AT} duplex used to characterize the structure of the complex of HMG-D D74 from *Drosophila melanogaster*; iii) a DNA without specific sequence recognition (Dragan *et al.*, 2003); iv) cisplatin-modified DNA from a previous study done with the Ixr1 protein (McA'Nulty *et al.*, 1996); v) a cruciform DNA that plays an important role in the regulation of natural processes, including replication, regulation of gene expression, nucleosome structure and recombination, as well as it is implicated in the evolution and development of diseases including cancer (Taudte *et al.*, 2001).

(a)

Tandem AB

GVVKKLSSTQSRIERRKQLKKQGPKRPSSAYFLFSMSIRNELLQQFPEAKVPELSKLASARWKELTDDQKKPFYEEFR

TNWEKYRVVRDAYEKTLPPKRPSGPFIQFTQEIRPTVVKENPDKGLIEITKIIGERWRELDPAKKAEYTETYKKRLKE

WESCYPDENDPNGNPTG

HMG box A

GVVKKLSSTQSRIERRKQLKKQGPKRPSSAYFLFSMSIRNELLQQFPEAKVPELSKLASARWKELTDDQKKPFYEEFR

TNWEKYRVVRDAYEKTLPPKRPSG

HMG box B

GKTLPPKRPSGPFIQFTQEIRPTVVKENPDKGLIEITKIIGERWRELDPAKKAEYTETYKKRLKEWESCYPDENDPNG
NPTG

(b)
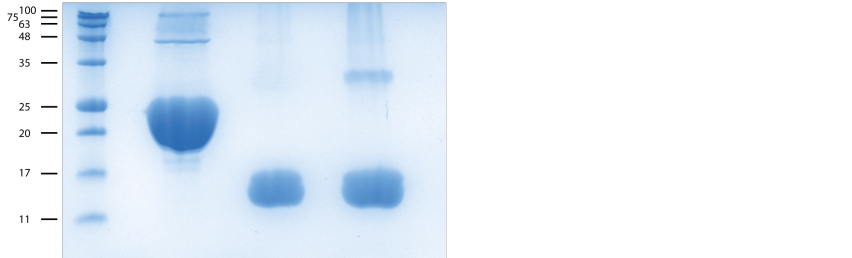
Figure 2. (a) Sequences of the Ixr1 protein regions, which were object of our study; Ixr1 full-length protein sequence was omitted. Residues that form the globular domain of the protein are enclosed in green boxes to distinguish them from the rest of the Ixr1 protein region included. Positively charged residues are shown in blue and negatively charged residues in red. (b) SDS-PAGE of tandem-AB (lane1), HMG-box A (lane 2) and HMG-box B (lane 3) proteins purified and used in the present study.
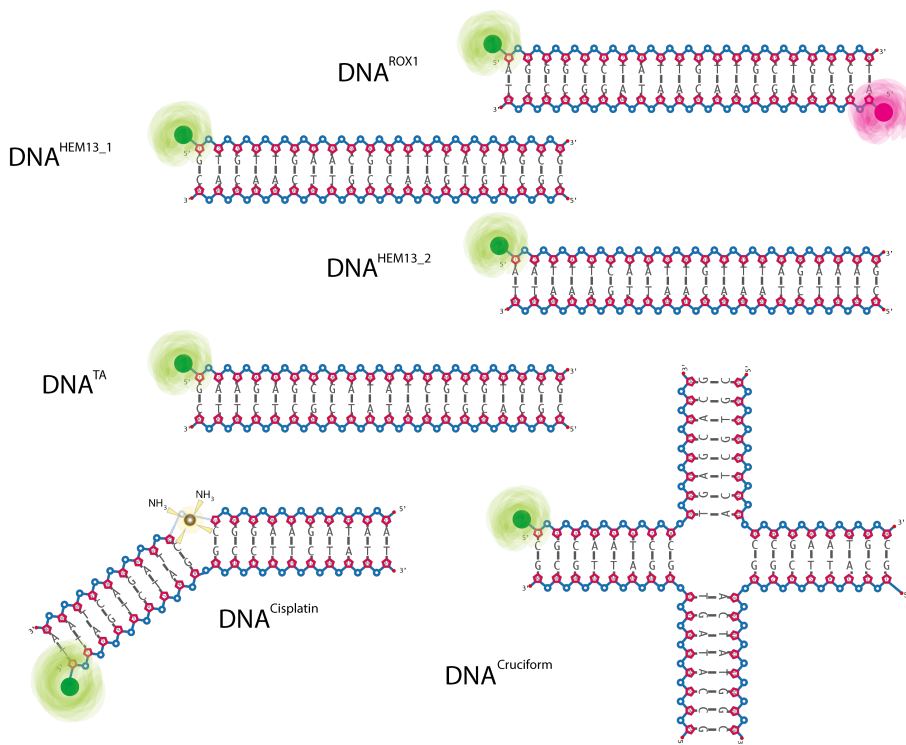
Figure 3. DNA templates used in this study, including lineal duplexes (DNA[ROX1], DNA[HEM13_1], DNA[HEM13_2] and DNA[TA]), platinated DNA (DNA[Cisplatin]) and four-way junctions (DNA[Cruciform]). Unlabeled DNAs were used in calorimetric, circular dichroism, NMR, SAXs and fluorescence anisotropy competition experiments. 5' FAM-labeled (green) DNAs were used in fluorescence anisotropy experiments and double 5' FAM-TAMRA-labeled (green-pink) DNA[ROX1] was used in FRET experiments.

3.1.2.1. CD analysis of protein folding and changes associated to DNA[ROX1] binding

To check the state and correct folding of all the Ixr1 protein-regions purified, they were analyzed by circular dichroism (figure 4). Far UV-CD was used to determine the secondary structure content of proteins, using the K2D2 software to deconvolute the CD profiles (table 4).

As expected, all the purified HMG-box domains have a predominant α-helix structured profile, with two negative major peaks at 208 and 222 nm. Spectra

comparison suggests that HMG-box A and HMG-box B have the ability to fold independently. The lower content in α-helix of HMG-box A may be due to the amino terminal tail included in the protein purified (table 4). When folding was analyzed after adding a linear DNA containing a specific sequence recognized by Ixr1, and corresponding to the promoter region of the *ROX1* gene (Castro-Prego *et al*., 2010a), only the tandem-AB protein showed a slight increase in α-helical content (7%) (figure 4a), meanwhile individual HMG-box domains did not change their circular dichroism profiles (figures 4b and 4c). This result indicates that both HMG-box are simultaneously needed to allow a dynamic transition from the unbound to the bound state with this DNA and this transition stabilizes the tandem folding.

Table 4. Secondary structure content obtained by CD deconvolution analysis.

Protein	α-helix	β-strand	Random coil
Tandem-AB	69	2	29
Tandem-AB + DNA[ROX1] complex	76	2	22
HMG-box A	48	8	44
HMG-box A + DNA[ROX1] complex	48	8	44
HMG-box B	69	2	29
HMG-box B + DNA[ROX1] complex	69	2	29

3.1.2.2.- Thermodynamics of DNA binding by Ixr1 protein

The characterization of the energetic components that take part in the formation of the protein-nucleic acid complexes is important for a better understanding of the nature of the different molecular interactions that govern transcriptional regulation and modulation of chromatin states. Ixr1, like other proteins containing HMG-box domains (Thomas & Travers, 2001), interacts specifically with the minor groove of DNA, inducing structural effects such as bending and kinking, as a consequence of the intercalation of bulky hydrophobic amino acid residues of the HMG-boxes between successive base-pairs within the

DNA minor groove, which are accompanied with partial unwinding, widening of the minor groove and bending towards the major grove (Thomas & Travers, 2001). Intercalating residues of the HMG-box are flanked by conserved basic residues that bind to the phosphodiester bond of DNA and contribute to stabilize the complex (Travers, 2000). The DNA bending/binding is further modulated by the N- and C-terminal flanking sequences of the HMG-box (Štros, 1998; Travers, 2000; Thomas & Travers, 2001).

(a)

(b)

(c)

Figure 4. Far-UV circular dichroism spectra of Tandem-AB (a), HMG-box A (b) or HMG-box B (c) in absence (continuous lines) or presence of DNA[ROX1] (dashed lines).

This DNA bending by HMGBs provides a mechanism by which the proteins stimulate the activity of various gene promoters by enhancement of binding of transcription factors or bringing distant regulatory sequences into close proximity (Štros, 2010).

To obtain a complete thermodynamic description of the Ixr1-DNA complex formation process from the free components, we made measurements of the association constant and obtained the Gibbs energy, with its enthalpic and entropic components, as reported bellow.

3.1.2.2.1.- Thermal stability determination by CD of the free proteins

First, the HMG-box domains of Ixr1 were characterized in terms of thermal stability by Circular Dichroism (CD). Figure 5 shows the temperature dependence of the molar ellipticity at 222 nm of individual Ixr1 HMG-box domains and linked in tandem. Results show a decrease in the negative values of molar ellipticity at 222 nm over the range from 20 ^{o}C to 80 ^{o}C up to a value close to zero, corresponding to a completely unfolded state. Besides the major transition corresponding to unfolding by denaturation, the ellipticity also changes even at lower temperatures (5 ^{o}C to 20 ^{o}C). Typically, temperature unfolding of free HMG-boxes show that their L-shaped structure does not represent a single fold-unit but it is composed of sub-units which differ in stability and some units have high lability (Dragan *et al.*, 2004). In this sense, figure 5 (continouos lines) shows that the ellipticity change below 20 ^{o}C is especially pronounced in the case of HMG-box B and tandem-AB, indicating that HMG-box B is specially labile in solution and their binding to DNA[ROX1] could stabilize a conformation.

The first derivative of the circular dichroism melting curves (figure 5 insets) show that melting temperatures of HMG-box A and B of Ixr1 are very close among them (table 5). In the case of the tandem-AB, a single transition was

observed, suggesting that either HMG-box A and HMG-box B unfold at similar temperatures, or unfold in a cooperative manner (figure 5a). When duplex DNA^{ROX1} was added, thermal stability improvement was only observed in the case of HMG-box B (4 ^{o}C higher) (figure 5c, dashed line), meanwhile transition midpoints of tandem-AB and HMG-box A remain unchanged (figures 5a and 5b, dashed lines).

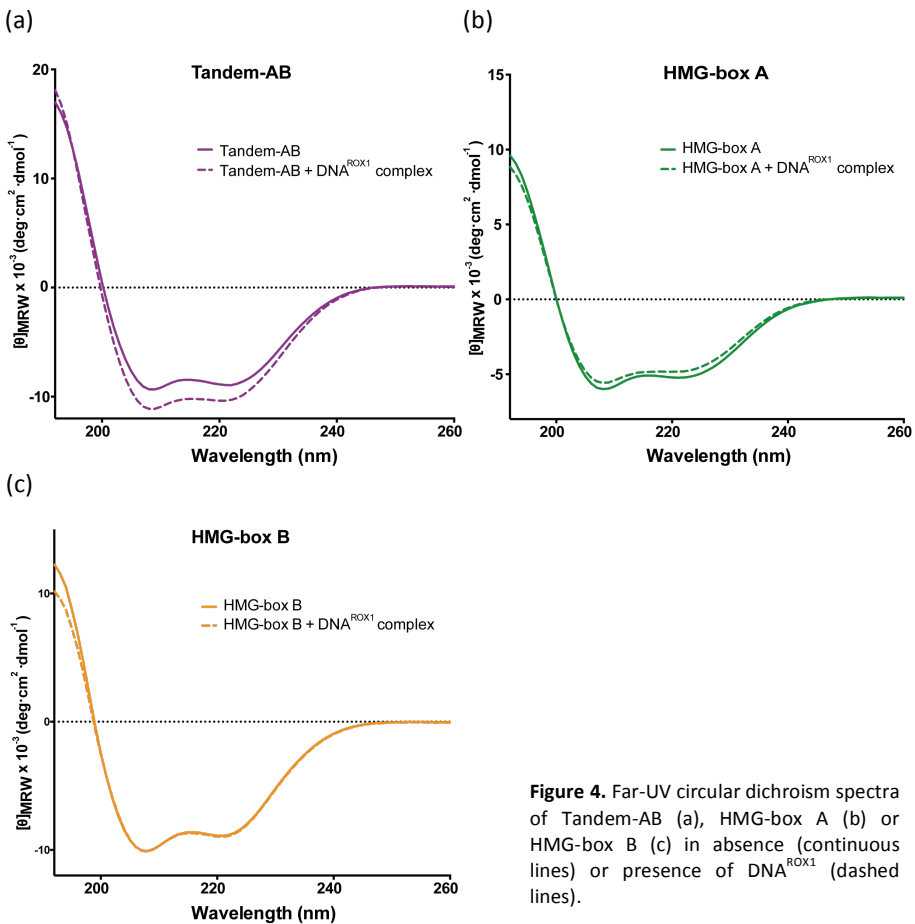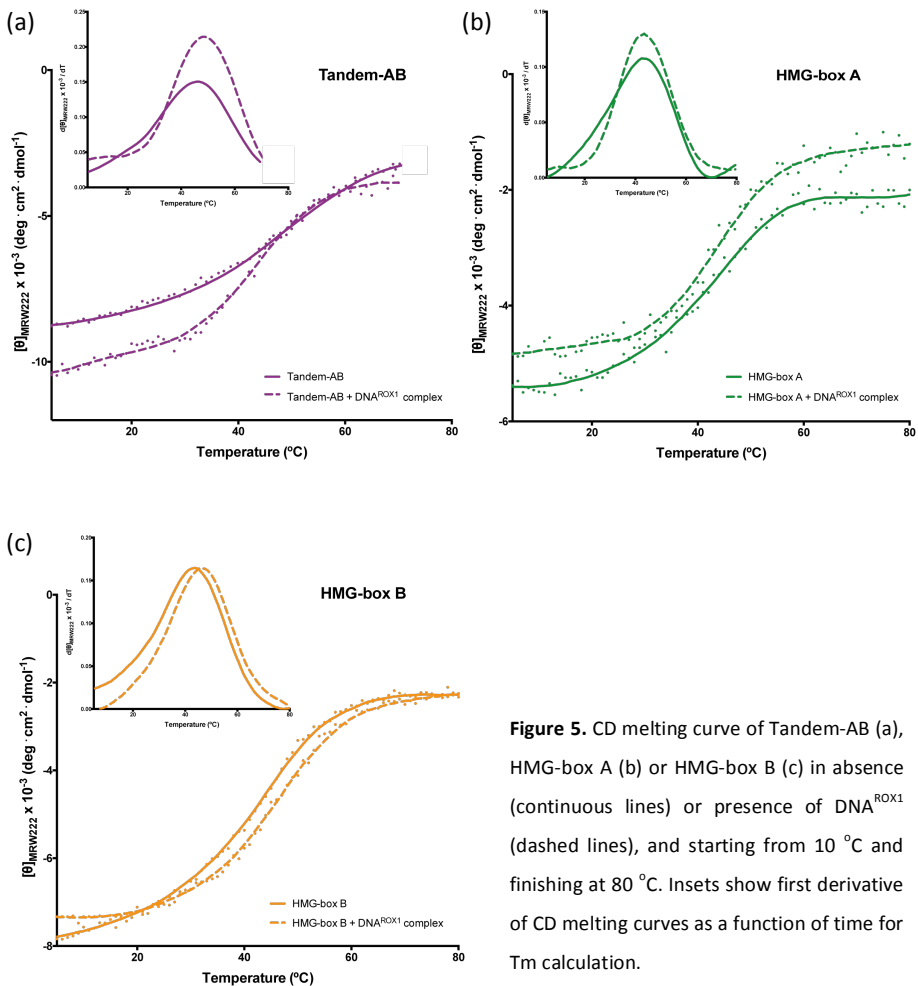**Figure 5.** CD melting curve of Tandem-AB (a), HMG-box A (b) or HMG-box B (c) in absence (continuous lines) or presence of DNA^{ROX1} (dashed lines), and starting from 10 ^{o}C and finishing at 80 ^{o}C. Insets show first derivative of CD melting curves as a function of time for Tm calculation.

238

Table 5. Melting temperatures obtained by ellipticity at 222 nm.

Protein	T_m (^oC)
Tandem-AB	46.5
Tandem-AB + DNA^{ROX1} complex	47.5
HMG-box A	43
HMG-box A + DNA^{ROX1} complex	43.5
HMG-box B	43.5
HMG-box B + DNA^{ROX1} complex	47.5

3.1.2.2.2. Thermodynamic DNA-binding constants HMG-box domains of Ixr1 by fluorescence anisotropy and ITC

To elucidate the contribution of the different HMG-box domains to DNA binding, equilibrium-binding assays were made comparing the different polypeptides purified. Since the association constants of the HMG-box domains are usually high (nanomolar range), and it was necessary to calculate them under different solvent conditions, they were only measured by optical methods that are more efficient. Moreover, isothermal calorimetry was used to measure the binding enthalpies of association. Using both approaches, all thermodynamic terms of each protein-DNA union were obtained with high accuracy (see Materials and methods).

In a first approach, DNA binding to the polypeptides under study was analyzed by EMSA (figure 6). After verifying the interaction, the apparent dissociation constants (K_d) of the complexes formed between the HMGs constructs and the DNA^{ROX1} forms was calculated as previously described (see section 2.4). In this sense, the dissociation constants obtained with DNA^{ROX1} were K_d = 3.45 x 10^{-7} for tandem-AB, K_d = 4.05 x 10^{-7} for HMG-box A and K_d = 8.95 x 10^{-6} for HMG-box B. Similar values were obtained for DNA$^{HEM13\_1}$, DNA$^{HEM13\_2}$, DNA^{TA} and DNA^{Cisplatin}, with K_d values in the range of 10^{-7} for tandem-AB and HMG-box A and 10^{-6} for HMG-box B (Figures S1, S2, S3 and S4). High values of dissociation constants for individual HMG-box domains were previously observed in other HMG-box proteins

with tandem domains such as HMO1, HMGB1 or TFAM proteins (Müller *et al.*, 2001; Kamau *et al.*, 2004; Rubio-Cosials *et al.*, 2011).
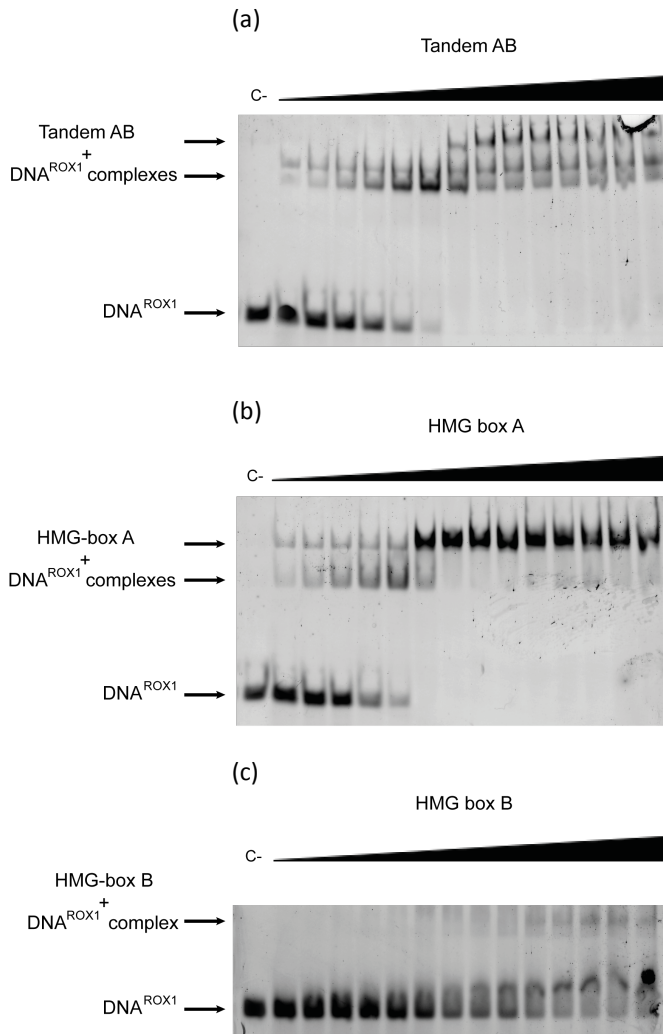
Figure 6. Binding of tandem-AB (a), HMG-box A (b) and HMG-box B (c) to DNA^{ROX1}. Lane C- contains free duplex (50 nM). Lanes 2-15 contains growing protein concentrations, ranging from 10 nM to 81.92 µM (tandem-AB), from 9 nM to 73.73 µM (HMG-box A) and from 44 nM to 360.45 µM (HMG-box B), doubling the concentration of protein in each lane from the previous one.

Fluorescence anisotropy titrations were then carried out to obtain more accurate association isotherms of Ixr1 HMG-box domains to different types of DNA. By fluorescence excitation of the DNA duplex labeled with FAM, it is possible to observe the anisotropy of its fluorescence, which reflects the change in DNA tumbling rate caused by protein binding, thereby yielding the binding isotherm (see Materials and Methods).

Tables 6, 7 and 8 summarize the thermodynamic components of association at 30 ^{o}C of the different Ixr1 protein regions. Tandem- AB+DNA complexes exhibit positive cooperative effects in the binding to both sequence specific and structure specific DNA samples, with Hill coefficients between 1.5 and 2 (less with DNA^{Cruciform}). Affinities (expressed as association constants, K_a) were calculated by fluorescence anisotropy and results show that the tandem-AB binds to DNA^{ROX1} 1.5-2 folds more tightly than HMG-box A, and up to 7 folds higher than HMG-box B. This increases to 60 fold higher when comparing the tandem-AB+DNA^{Cisplatin} with the HMG-box B+DNA^{Cisplatin} complexes. In general, all protein constructs analyzed show higher affinities with modified DNA structures than with classical B-form DNA duplexes. For the tandem-AB+DNA^{ROX1} (K_a = 1.69 x 10^6 M^{-1}), tandem-AB+DNA$^{HEM13\_1}$ (K_a = 1.96 x 10^6 M^{-1}), tandem-AB+DNA$^{HEM13\_2}$ (K_a = 3.23 x 10^6 M^{-1}) and tandem-AB+DNA^{TA} (K_a = 3.57 x 10^6 M^{-1}) complexes, association constants are up to five folds lower than for tandem-AB+DNA^{Cisplatin} (K_a = 8.33 x 10^6 M^{-1}) or tandem-AB+DNA^{Cruciform} (K_a = 1.0 x 10^7 M^{-1}).

For the HMG-box A+DNA^{ROX1} (K_a = 1.30 x 10^6 M^{-1}), HMG-box A+DNA$^{HEM13\_1}$ (K_a = 1.19 x 10^6 M^{-1}), HMG-box A+DNA$^{HEM13\_2}$ (K_a = 1.59 x 10^6 M^{-1}) and HMG-box A+DNA^{TA} (K_a = 1.67 x 10^6 M^{-1}) complexes, affinity constants are up to 2.75 folds lower than for HMG-box A+DNA^{Cisplatin} (K_a = 4.17 x 10^6 M^{-1}) or HMG-box A+DNA^{Cruciform} (K_a = 3.57 x 10^6 M^{-1}). In the case of HMG-box B, binding constants obtained differ from tandem-AB and HMG-box A by one order of magnitude. Thus, association constants of HMG-box B differs from HMG-box A ranging from 5 folds

lower for binding to DNA^{ROX1}, up to 30 folds lower for complexes with DNA^{Cisplatin}.

Table 6. Thermodynamic characteristics of tandem-AB binding to DNA

DNA	K_a (M^{-1})	Hill slope	ΔH_1 (kJ/mol)	ΔH_2 (kJ/mol)	ΔG_1 (kJ/mol)	ΔG_2 (kJ/mol)	$T\Delta S_1$ (kJ/mol)	$T\Delta S_2$ (kJ/mol)
DNA^{ROX1}	1.69 x 10^6 ± 0.03 (FA)	2.03	--	--	--	--	--	--
site1	2.05 x 10^7 ± 0.92 (ITC)	--	7.60	--	-40.34	--	47.94	--
site2	2.91 x 10^5 ± 0.14 (ITC)	--	--	89.45	--	-30.14	--	119.59
DNA$^{HEM13\_1}$	1.96 x 10^6 ± 0.04 (FA)	1.55	--	--	-34.70	-34.70	--	--
DNA$^{HEM13\_2}$	3.23 x 10^6 ± 0.02 (FA)	1.54	--	--	-35.90	-35.90	--	--
DNA^{TA}	3.57 x 10^6 ± 0.02 (FA)	1.63	--	--	-36.10	-36.10	--	--
DNA^{Cisplatin}	8.33 x 10^6 ± 0.01 (FA)	1.84	--	--	--	--	--	26.76
site1	3.95 x 10^6 ± 2.07 (ITC)	--	-60.69	--	-36.4	--	24.29	--
site2	3.57 x 10^3 ± 1.34 (ITC)	--	--	772.78	--	-19.6	--	792.38
DNA^{Cruciform}	1.00 x 10^7 ± 0.01 (FA)	1.05	--	--	--	--	--	--
site1	4.97 x 10^6 ± 0.31 (ITC)	--	-5.55	--	-36.94	--	31.39	--
site2	3.86 x 10^4 ± 0.22 (ITC)	--	--	125.71	--	-25.30	--	151.01

K_a: Affinity constant
FA = Fluorescence Anisotropy
ITC = Isothermal Titration Calorimetry; sites 1 and 2 refer to first and second binding events obtained in ligand titrations by ITC
Enthalpies (ΔH) were obtained by ITC
Gibbs energy (ΔG) were obtained by isothermal calorimetry with the equation
$\Delta G = -RTLn(K_a)$, where R= 8.314472 J L^{-1} mol^{-1}; and T= 288.15 K
Enthropy (ΔS) were obtained from the equation: $T\Delta S = \Delta H - \Delta G$

Table 7. Thermodynamic characteristics of HMG-box A binding to DNA

DNA	K_a (M^{-1})	ΔH (kJ/mol)	ΔG (kJ/mol)	$T\Delta S$ (kJ/mol)
DNA^{ROX1}	$1.30 \times 10^6 \pm 0.04$ (FA)	--	-33.70	67.49
	$5.29 \times 10^5 \pm 0.27$ (ITC)	30.79	--	--
DNA$^{HEM13\_1}$	$1.19 \times 10^6 \pm 0.02$ (FA)	--	-33.50	--
DNA$^{HEM13\_2}$	$1.59 \times 10^6 \pm 0.05$ (FA)	--	-34.20	--
DNA^{TA}	$1.67 \times 10^6 \pm 0.06$ (FA)	--	-34.30	--
DNA^{Cisplatin}	$4.17 \times 10^6 \pm 0.03$ (FA)	--	-36.50	-3.43
	$3.40 \times 10^6 \pm 0.60$ (ITC)	-39.93	--	--
DNA^{Cruciform}	$3.57 \times 10^6 \pm 0.03$ (FA)	--	-36.10	-8.82
	$5.49 \times 10^6 \pm 0.36$ (ITC)	-44.92	--	--

K_a: Affinity constant
FA = Fluorescence Anisotropy
ITC = Isothermal Titration Calorimetry
Enthalpies (ΔH) were obtained by ITC
Gibbs energy (ΔG) were obtained by fluorescence anisotropy with the equation
$\Delta G = -RTLn(K_a)$, where R= 8.314472 J L^{-1} mol^{-1}; and T= 288.15 K
Enthropy (ΔS) were obtained from the equation: $T\Delta S = \Delta H - \Delta G$

Table 8. Thermodynamic characteristics of HMG-box B binding to DNA

DNA	K_d (M)	ΔH (kJ/mol)	ΔG (kJ/mol)	$T\Delta S$ (kJ/mol)
DNA^{ROX1}	$2.30 \times 10^5 \pm 0.46$ (FA)	--	-29.60	61.82
	$2.40 \times 10^4 \pm 3.78$ (ITC)	32.22	--	--
DNA$^{HEM13\_1}$	$1.19 \times 10^5 \pm 1.86$ (FA)	--	-28.01	--
DNA$^{HEM13\_2}$	$1.18 \times 10^5 \pm 1.50$ (FA)	--	-27.98	--
DNA^{TA}	$1.68 \times 10^5 \pm 1.58$ (FA)	--	-28.82	--
DNA^{Cisplatin}	$1.37 \times 10^5 \pm 1.52$ (FA)	--	-28.30	66.68
	$7.25 \times 10^4 \pm 0.46$ (ITC)	38.38	--	--
DNA^{Cruciform}	$1.67 \times 10^5 \pm 0.99$ (FA)	--	--	--

K_a: Affinity constant
FA = Fluorescence Anisotropy
ITC = Isothermal Titration Calorimetry
Enthalpies (ΔH) were obtained by ITC
Gibbs energy (ΔG) were obtained by fluorescence anisotropy with the equation
$\Delta G = -RTLn(K_a)$, where R= 8.314472 J L^{-1} mol^{-1}; and T= 288.15 K
Enthropy (ΔS) were obtained from the equation: $T\Delta S = \Delta H - \Delta G$

(a)

(b)

(c)
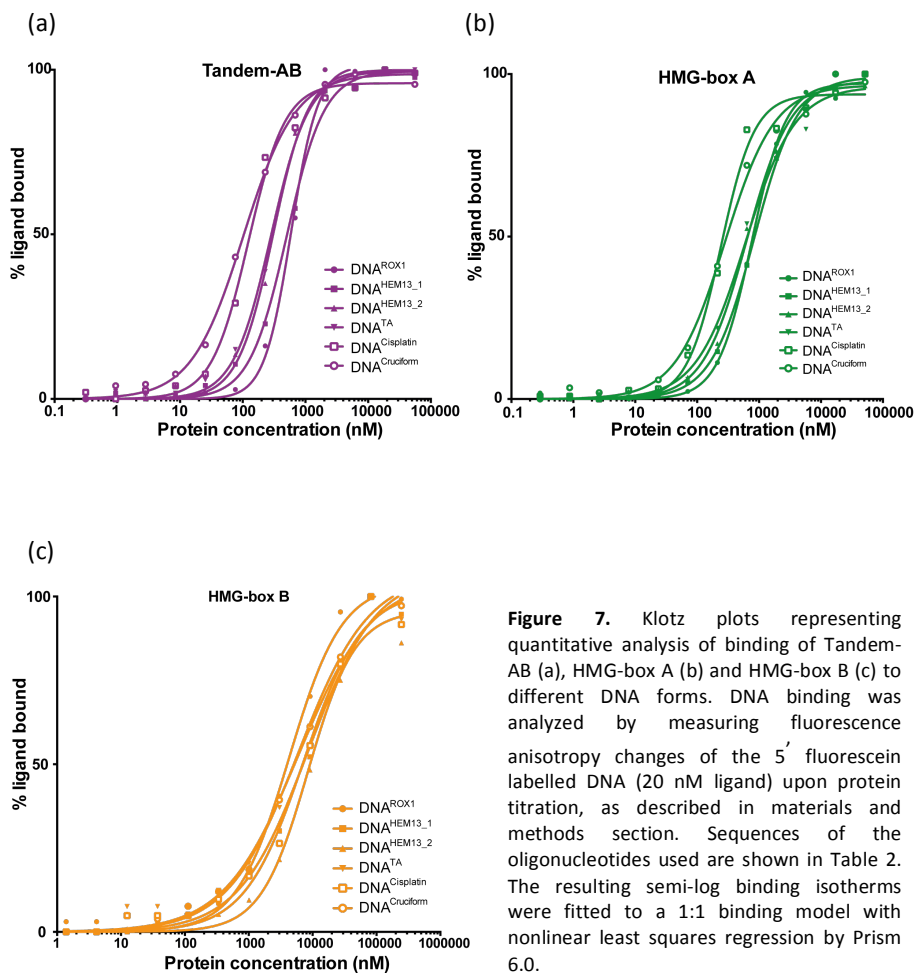
Figure 7. Klotz plots representing quantitative analysis of binding of Tandem-AB (a), HMG-box A (b) and HMG-box B (c) to different DNA forms. DNA binding was analyzed by measuring fluorescence anisotropy changes of the 5′ fluorescein labelled DNA (20 nM ligand) upon protein titration, as described in materials and methods section. Sequences of the oligonucleotides used are shown in Table 2. The resulting semi-log binding isotherms were fitted to a 1:1 binding model with nonlinear least squares regression by Prism 6.0.

Because of the low difference among association constants obtained of HMG-box constructions bound to duplex DNA with different sequences, competition experiments were done with the aim of distinguish more accurately affinity preferences to different specific DNA sequences. For that, unlabeled DNA titrations were performed to the protein-DNA^{ROX1}-labeled complex and inhibition constants (Ki) were calculated. They were used the B-form DNA duplexes, including DNA^{ROX1}, DNA$^{HEM13\_1}$, DNA$^{HEM13\_2}$, DNA^{TA} and mutation form of DNA^{ROX1} called DNA$^{ROX1\_mut}$, in which the core sequence 5′ *AGGGCCT**ATTGTT**GCTGCCT* 3′ was

substituted by 5' *AGGGCCT**TGCGCA**GCTGCCT* 3'. The results and data fitting obtained are summarized in figure 8 and table 9. HMGA-box A domain offers the highest sequence specificity to *ROX1* promoter site, showing lower K_i to DNA^{ROX1} than Tandem-AB and HMG-boxB, and larger K_i change when sequence is mutated in DNA$^{ROX1\_mut}$. For the two *HEM13* promoter binding sites described by Castro-Prego and co-workers (Castro-Prego et al., 2010b), table 9 shows that HMG-box A has higher sequence preference to DNA$^{HEM13\_1}$ and HMG-box B to DNA$^{HEM13\_2}$.

Table 9. Inhibition constants obtained to different lineal DNA duplexes in competition assays with FAM-labelled DNA^{ROX1} in complex with tandem-AB, HMG-box A and HMG-box B.

Protein	DNA^{ROX1} (K_i, nM)	DNA$^{ROX1\_mut}$ (K_i, nM)	DNA$^{HEM13\_1}$ (K_i, nM)	DNA$^{HEM13\_2}$ (K_i, nM)	DNA^{TA} (K_i, nM)
Tandem-AB	3948	4910	3660	4865	4547
HMG-box A	2486	3795	1980	2292	2366
HMG-box B	4812	4247	2103	2271	3330

Then, isothermal calorimetry experiments were performed, in which DNA ligand was titrated into a solution of HMG-box A or HMG-box B, respectively. The ITC thermogram exhibits a typical calorimetric reaction upon the addition of DNA aliquots, showing an initial strong heat uptake that decreases when the binding sites on DNA become saturated. In the last injections of each titration, only dilution heat effects were observed. Figure 9 shows the heat effects of subsequent injections of DNA^{ROX1} (figures 9d and 9g), DNA^{Cisplatin} (figures 9e and 9h) and DNA^{Cruciform} (figure 9f) into solutions of box HMG-box A and HMG-box B, respectively. No suitable thermogram was obtained for HMG-box B binding to DNA^{Cruciform}. The plot of heat evolved per injection *(ΔQ_i)* versus molar ratio is showed in the inset panel of Figure 9. A nonlinear least-squares fit of the binding curves to a model with one (type) binding site was successfully applied, obtaining

association constant and enthalpy values (see tables 7 and 8) with an integer stoichiometry of ligand:protein close to 1:1 (n ≈ 1). In the case of HMG-box A, association constants are similar to those obtained by fluorescence anisotropy, meanwhile large differences were observed for HMG-box B association, probably because of a poor fitting of the model with the recorded data. It is interesting to note that the binding of HMG-box A to DNA^{Cisplatin} and DNA^{Cruciform} is an enthalpy driven process (with enthalpy values of -39.93 kJ/mol and -44.92 kJ/mol, respectively), meanwhile the binding to linear DNA^{ROX1} is an entropy driven process (ΔH = 30.79 kJ/mol). In contrast, binding of HMG-box B to both DNA^{ROX1} and DNA^{Cisplatin} describes unfavorable enthalpy binding values, indicating different binding mechanisms between both individual HMG-box domains to platinated DNA. In binding reactions there is a phenomenon known as the entropy/enthalpy compensation, in which protein folding, ligand binding and solvent and salt effects cause very large changes in entropy and enthalpy that mostly cancel out each other to give small changes in free energy. A favourable exothermic enthalpy driven association is one with strong and directional hydrogen bonds that leads to a more rigidly held complex, with less residual disorder and unfavorable entropy in the form of intermolecular motion. Alternatively, the association can have a favorable entropy, and thus be rather mobile in its geometry, with consequently weak hydrogen bonds. The optimum position of this balance varies from one ligand to another, and hence the balance between enthalpy and entropy varies. Then, more conformationally rigid complexes have favourable enthalpy, whereas more mobile ones have favourable entropy. Nevertheless, entropy and enthalpy are highly influenced by the large number of water molecules released during binding reactions, and arrangements in protein and ligand folding (Calderon & Williams, 2001; Williamson, 2012).
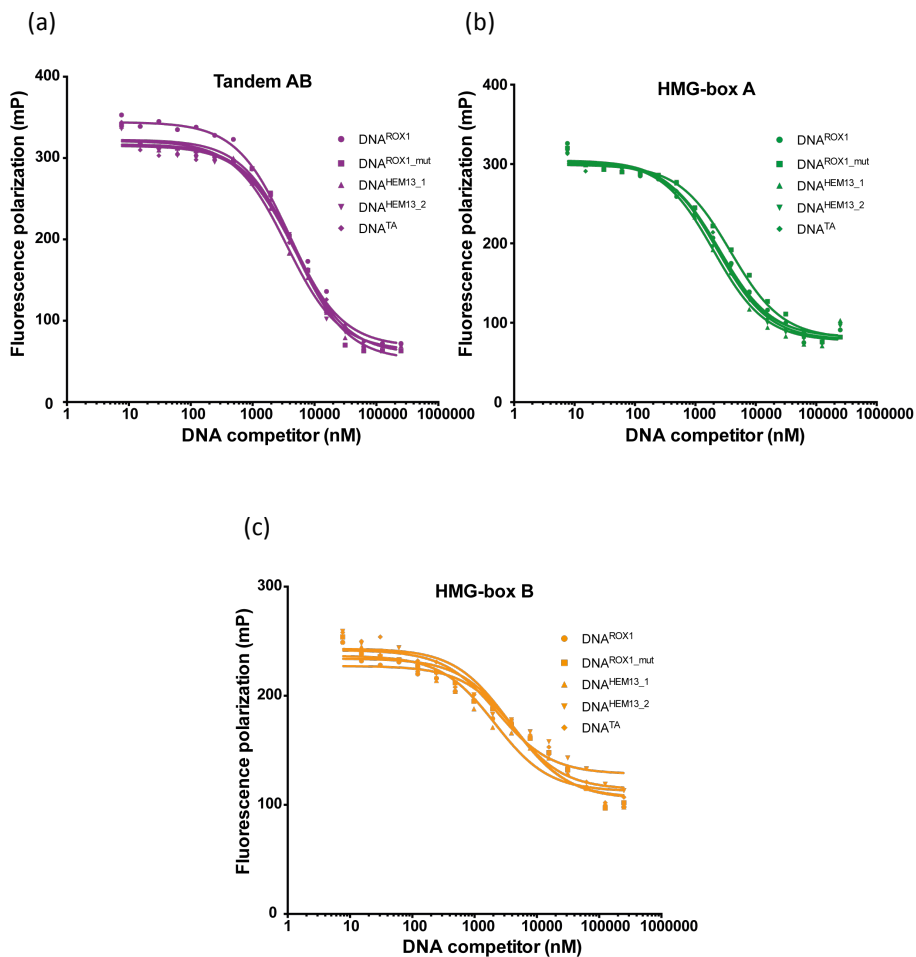
Figure 8. Competition assays of Tandem-AB (a), HMG-box A (b) and HMG-box B (c) in complex with DNA^{ROX1}, with different DNA forms. DNA binding was analyzed by measuring fluorescence anisotropy changes of the 5$^{'}$ fluorescein labelled DNA (50 nM ligand) upon unlabelled DNA titration, as described in materials and methods section. Sequences of the oligonucleotides used are shown in Table 2. The resulting semi-log binding isotherms were fitted to a 1:1 binding model with nonlinear least squares regression by Prism 6.0.
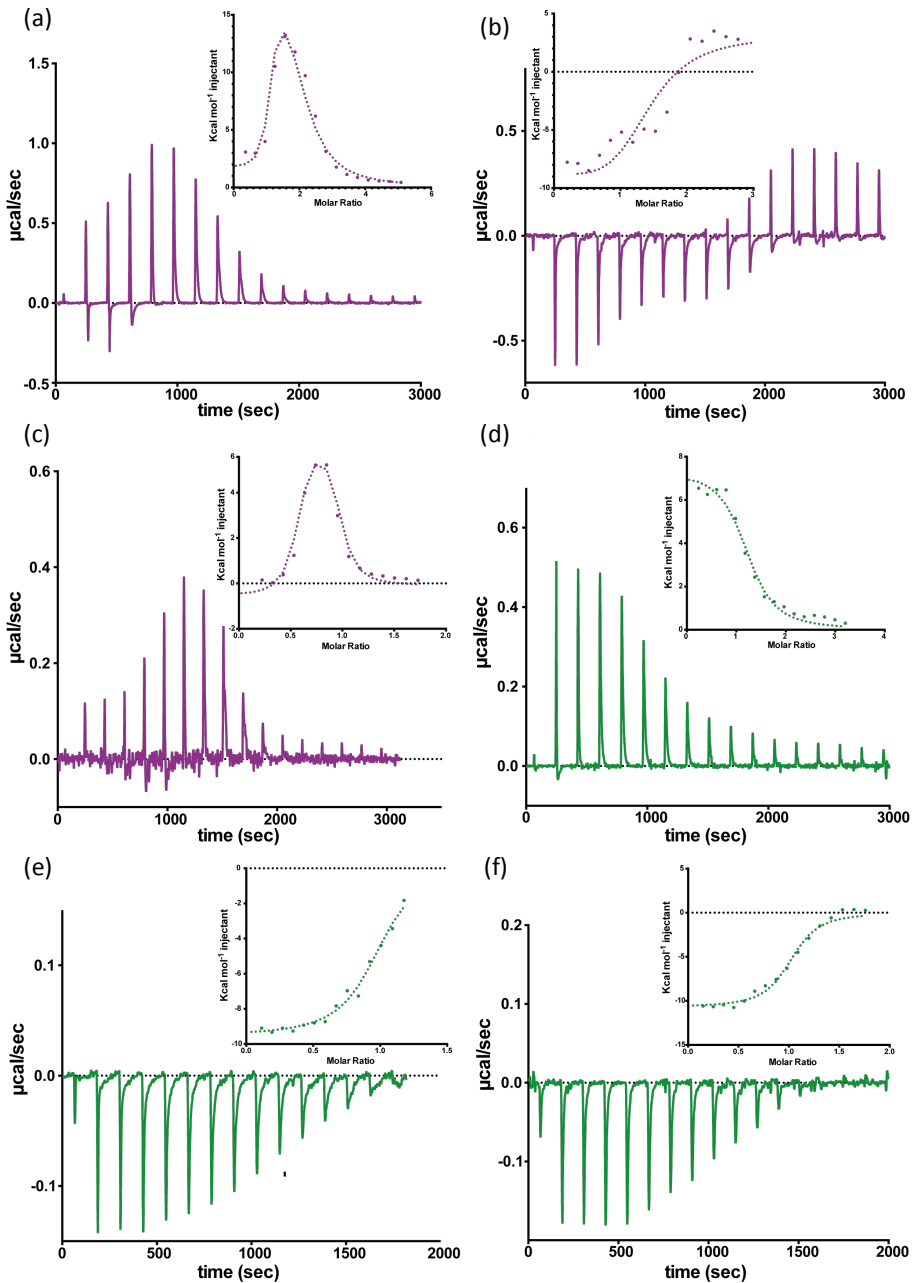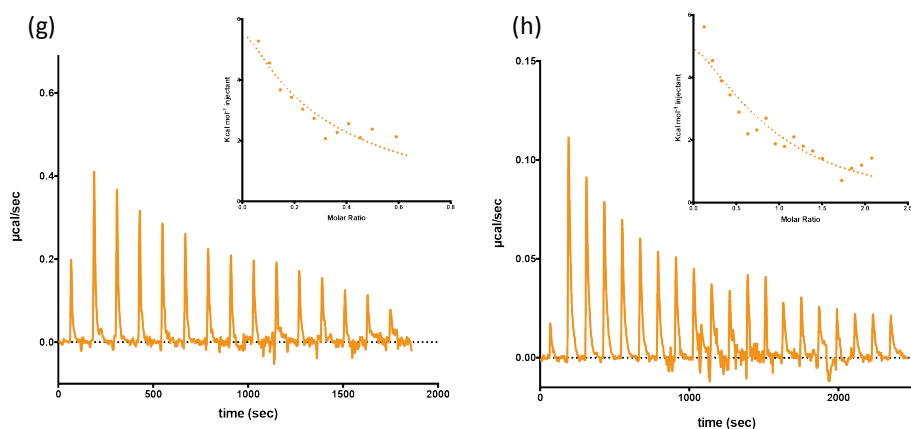
Figure 9. Thermodynamic analysis of tandem-AB (purple), HMG-box A (green) and HMG-box B (orange) to DNA[ROX1] (a, d, g), DNA[Cisplatin] (b, e, h) and DNA[Cruciform] (c, f). ITC thermograms show the raw heats absorbed from injecting target DNA (syringe) into protein solution (cell). Inset plots show the integrated absorbed heats with respect to time with the heat of mixing substracted and against protein:DNA molar ratio.

Previous studies of the thermodynamics of DNA binding for sequence-specific (Dragan *et al*., 2004; Malarkey *et al*.,2012) and structure specific HMG-box proteins (Müller *et al*., 2001; Dragan *et al*., 2003; Kamau *et al*., 2004; Dragan *et al*., 2004) showed differential profiles with unique thermodynamic signatures for both kind of DNA recognition mechanisms. The main difference between the two types of HMG-box+DNA complexes is that in sequence-specific interactions predominate extensive van der Waals contacts between apolar groups because of a more densely packed complex interface is formed, which decreases the overall positive enthalpy of binding up to close to zero and produces a negative heat capacity effect (enthalpy driven). In contrast, structure specific HMG-box+DNA complexes show positive enthalpy values around 40 kJ/mol, and a positive heat capacity effect (entropy driven). Based on this, figure 9 shows large positive enthalpy binding values to B-form lineal DNA for both HMG-box domains, indicative of low number of van der Waals and hydrogen bond contacts as a consequence of a poor interfacial complementarity between protein and DNA. These results indicate that,

when assayed separately, both HMG-box domains present a thermodynamic signature corresponding to the previously defined for the NSS family.

Combining the two HMG-boxes in the construct tandem-AB (residues 338-510) showed a number of very interesting thermodynamic binding features. Overall binding isotherms for the interaction of tandem-AB with DNA[ROX1] (figure 9a), DNA[Cisplatin] (figure 9b) and DNA[Cruciform] (figure 9c) present a 'biphasic' fashion with additional heat effects at high saturation, corresponding with the binding of both HMG-box domains. Since there is no information regarding the sequential DNA binding in Ixr1 protein, it is very important to analyze the ITC data carefully due to these featureless ITC thermograms. As shown in other HMG-box proteins (Muller *et al.*, 2001; Kamau *et al.*, 2004; Malarkey *et al.*, 2012), DNA binding in one HMG-box domain might affect the DNA binding to the other/s HMG-box domain/s. The equilibrium of a macromolecule with number (n) multiple ligand binding sites can be described by two different association constants: the macroscopic and microscopic association constants. The macroscopic association constant is model independent and describes the overall behavior of the n sites, whereas the microscopic association constant, k, takes into account how binding occurs at each site and is therefore model dependent. Macroscopic association constants are determined by ITC and can take the form of either an overall binding constant, β_j, or a stepwise binding constant, K_j, for ligation of the jth site.

Analyzing the tandem-AB binding to linear DNA[ROX1] using a model with two sequential binding sites (see materials and methods) yielded overall association values (β_j) and associated enthalpy changes of 2.05 x 10^7 M^{-1} (β_1) and 7.60 kJ/mol (ΔH_1), and 5.97 x 10^{12} M^{-1} (β_2) and 89.45 kJ/mol (ΔH_2) for the two HMG-box domains. Deduced by affinity constant correlations, HMG-box A corresponds with the first binding event and HMG-box B (lower affinity) corresponds with the second binding site. The macroscopic cooperative constant ρ was then determined to be 0.056 (see materials and methods), which indicates that the binding model

may be defined by either two independent binding sites or two identical binding sites with a negative cooperativity.

A model with two identical binding sites was discarded because of the binding of one of the HMG-box domain excluded the other one to the same ligand region. Using a model with two independent sites, fitting yields values of K_a and enthalpy of 2.13×10^7 M^{-1} (K_{a1}) and 11.72 kJ/mol (ΔH_1), and 2.95×10^5 M^{-1} (K_{a2}) and 63.68 kJ/mol (ΔH_2) for the two HMG-box domains respectively. Large and positive enthalpy values, specially in the second binding event, are indicative of a NSS binding event, with lower density packing protein-DNA interface, and positive entropy change as a result of releasing water from its ordered states. However, the use of the model with two independent sites is refutable because the values of heat released due to DNA binding to individual HMG-box domains alone (see tables 7 and 8) are between those of ΔH_1 and ΔH_2, suggesting that the model with independent binding is unlikely. Not all the cases are distinguishable experimentally and sometimes a macromolecule with two different binding sites exhibiting positive cooperative effect might resemble a macromolecule with two identical and independent binding sites, because both features, binding curve and thermodynamic information, could have compensating effects (Freire *et al.*, 2009).

Binding of tandem-AB to DNA^{Cruciform} and DNA^{Cisplatin} also occurs showing a 'biphasic' fashion. However, tandem-AB+DNA^{Cisplatin} complex formation presents an enthalpy favourable binding event, probably because of the recognition and binding of HMG-box A to the cisplatin DNA adduct, and a second enthalpy unfavourable binding event as a consequence of binding of HMG-box B to an adjacent DNA region.

3.1.3.- Ionic dependence of Ixr1p binding

A complete description of the interactions between DNA, a highly charged macromolecule, and the DNA binding domains present in the proteins requires

separating the overall binding energy into its electrostatic and non-electrostatic components. The counter-ion condensation concept is based on a linear dependence between the logarithmic values of the measured binding constants and the logarithmic values of the salt concentration in the media, which reflects the electrostatic interactions in this process. The effect of salt addition is a purely entropic process in which formation of ion pairs between the cationic amino acid residues of the protein and the poly-anionic DNA results in the release of counter-ions and an entropic increment (Dragan *et al.*, 2003; Dragan *et al.*, 2004; Privalov *et al.*, 2007).

In general, the electrostatic component (*el*) of the binding energy, which results from the ionic and polar interactions with the DNA phosphate groups, predominates over the non-electrostatic (*nel*) for both, sequence-specific and structure-specific, types HMG-boxes (Dragan *et al.*, 2003; Dragan *et al.*, 2004; Privalov *et al.*, 2007). Furthermore, the electrostatic component is in any case independent on the target DNA. Specificity depends uniquely on the non electrostatic component, and therefore this component varies with the DNA sequence, being higher for the more perfect target (Privalov *et al.*, 2011). This non-electrostatic component is influenced by protein refolding, a decrease in the translational/rotational degrees of freedom and changes in the hydration of the reaction components, including positive contributions by dehydration of contacting groups and negative contributions by the incorporation of water molecules at the newly formed interface. Fluorescence anisotropy titrations were carried at increasing salt concentrations, ranging from 100 mM to 400 mM of KCl. Figure 10 shows that binding to different linear B-form DNA sequences does not alter the slope of the *log(K_a) versus log[KCl]* plots, giving different but close extrapolated values at *log[KCl]* = 0 (see Table 10). In contrast, binding to non-conventional DNAs (DNA^{Cisplatin} and DNA^{Cruciform}) produce changing slopes, mainly in the case of the tandem-AB where the changes are abrupt, indicating different electrostatic relationship from one to another DNA type.
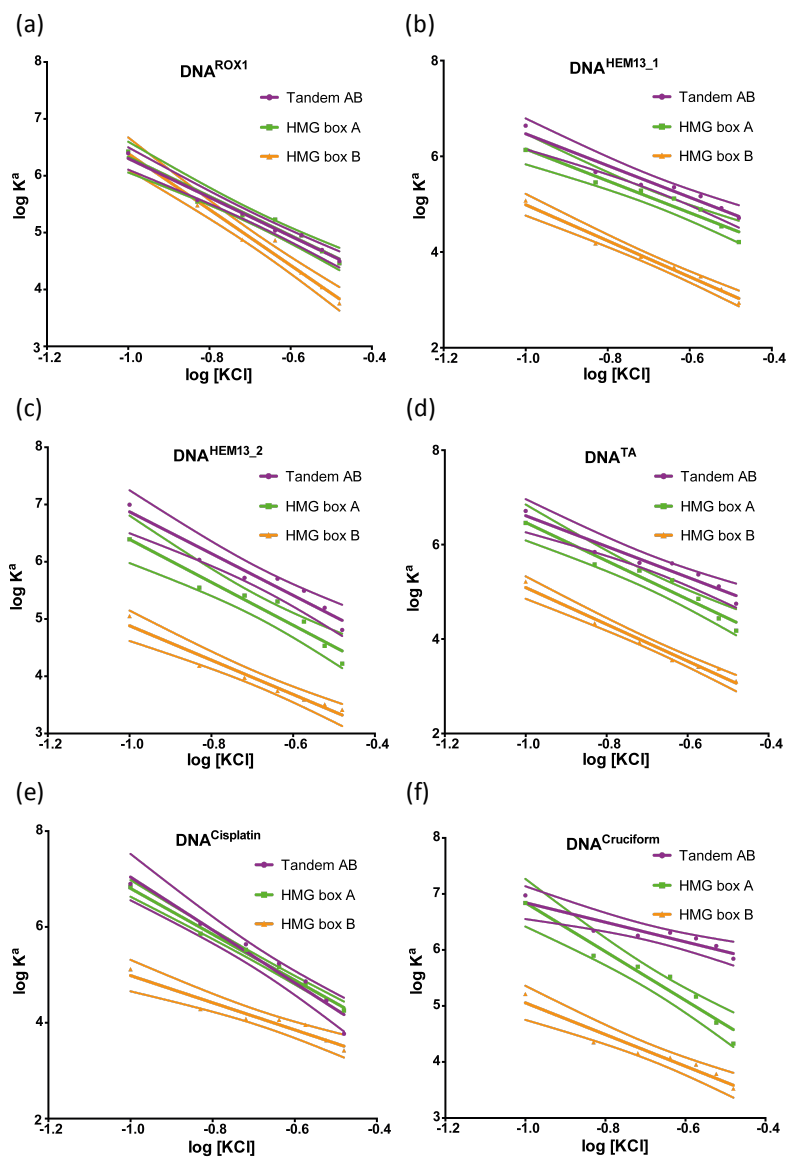
Figure 10. Ionic strength dependence of tandem-AB, HMG-box A and HMG-box B in complex formation with DNA^{ROX1} (a), DNA^{HEM13_1} (b), DNA^{HEM13_2} (c), DNA^{TA} (d), $DNA^{Cisplatin}$ (e) and $DNA^{Cruciform}$ (f) upon increasing KCl concentrations (100, 200, 300 and 400 mM KCl) in 10 mM potassium phosphate (pH6.8). Binding isotherms obtained (figure S5) were plotted in the logarithm of association constants against the logarithm of salt concentrations. Dashed lines indicate 95% confidence intervals.

Table 10. Ionic dependence of HMG-box domains to different target DNAs, showing the log of association constants in 100 mM KCl ($log(K_a)$), their extrapolated value at 1 M KCl ($log(K_a)_0$), the slopes of the $log(K_a)$ *versus* log[KCl] plots, and the number of ionic contacts formed between a HMG-box domain and the phosphate groups of DNA (Z).

Domain charges	Property	DNA^{ROX1}	DNA$^{HEM13\_x}$	DNA$^{HEM13\_z}$	DNA^{TA}	DNA^{Cisplatin}	DNA^{Cruciform}
Tandem-AB	$log(K_a)$	6.23±	6.29±	6.51±	6.55±	6.92±	7.00±
35$^+$-26$^-$ = 9$^+$		0.15	0.14	0.17	0.17	0.20	0.20
	$log(K_a)_0$	2.89±	3.15±	3.22±	3.36±	1.52±	5.10±
		0.15	0.24	0.28	0.36	0.36	0.22
	$-\delta\ log\ (K_a)/\ \delta log\ [KCl]$	3.41±	3.32±	3.65±	3.26±	5.52±	1.75±
		0.21	0.35	0.40	0.38	0.52	0.48
	Z^*	5.33±	5.19±	5.69±	5.09±	8.63±	2.73±
		0.33	0.55	0.64	0.59	0.81	0.23
HMG-box A	$log(K_a)$	6.11±	6.08±	6.20±	6.22±	6.63±	6.55±
23$^+$-12$^-$= 9$^+$		0.14	0.17	0.13	0.12	0.15	0.15
	$log(K_a)_0$	2.88±	2.82±	2.64±	2.41±	2.02±	2.48±
		0.20	0.24	0.31	0.29	0.13	0.32
	$-\delta\ log\ (K_a)/\ \delta log\ [KCl]$	3.45±	3.33±	3.75±	4.06±	4.79±	4.36±
		0.29	0.34	0.45	0.41	0.19	0.46
	Z^*	5.39±	5.20±	5.86±	6.34±	7.48±	6.81±
		0.45	0.53	0.70	0.64	0.29	0.72
HMG-box B	$log(K_a)$	5.37±	5.08±	5.07±	5.22±	5.14±	5.22±
15$^+$-13$^-$ = 2$^+$		0.34	0.27	0.18	0.20	0.18	0.01
	$log(K_a)_0$	1.48±	1.22±	1.88±	1.19±	2.15±	2.23±
		0.21	0.17	0.20	0.18	0.25	0.23
	$-\delta\ log\ (K_a)/\ \delta log\ [KCl]$	4.91±	3.77±	3.00±	3.89±	2.83±	2.83±
		0.31	0.24	0.28	0.26	0.35	0.33
	Z^*	7.67±	5.89±	4.69±	6.08±	4.42±	4.42±
		0.48	0.38	0.44	0.41	0.55	0.52

*Z values were obtained from the equation $log(K_a)= log(K_a^{nel}) - Z\Psi\ log[KCl]$, where Ψ is 0.64 for DNA duplexes (Olmsted *et al.*, 1995)

Water activity is less affected at these relatively low concentrations of salt (100-400 mM KCl), and the entropy effect is simply proportional to the number of counter-ions released from the DNA (Ha *et al.*, 1992; Manning, 2003). As a consequence, the logarithm of the association constant can be split in the expression $log(K_a)= log(K_a^{nel}) - Z\Psi log[KCl]$, where Z is the number of DNA phosphates that interact with the protein and Ψ is the number of cations (K^+) per

phosphate group released upon protein binding (Dragan *et al*., 2004; Privalov *et al*., 2011). The first part of the equation results form the non-electrostatic interactions between DNA and protein, and the second side corresponds to the release of counter-ions that produces electrostatic effects. From the slope of the *log(K$_a$)* function is possible to calculate number of phosphate groups which release their counter-ions (*Z*) upon protein binding to DNA, i.e. the number of ionic contacts between protein and DNA. These numbers, given in table 10, are in good agreement with previously data obtained for individual HMG-box domains without flanking tails (Dragan *et al*., 2004).

The fact that the *Z* numbers were closer to net charge rather that the total positive charges indicates the presence of a substantial number of internal salt links on the surface of the HMG-box domains, especially on helix 3, which is remote from the DNA (see figure 11) (Weir *et al*., 1993; Love *et al*., 1995; Murphy *et al*., 1999; Murphy *et al*., 2001; Dragan *et al*., 2003; Dragan *et al*., 2004; Privalov *et al*., 2011). It is interesting to note that the binding of the tandem-AB and HMG-box A to DNA[Cisplatin] produces the release of ≈9 and ≈8 counter-ions, respectively (see table 10), indicating extra salt links that could been stabilizing the union, increasing the affinity in comparison to linear DNA.

At higher salt contentrations (≈1 M KCl), the electrostatic part vanishes (*log[KCl]*= 0) and *ΔG$_a$ = -RTLn(K$_a$)* correspond nearly to the non-electrostatic part of the Gibbs energy of complex formation (*ΔG$_a^{nel}$*), allowing split the measured Gibbs energies of binding in the electrostatic and non electrostatic components (*ΔG$_a$= ΔG$_a^{nel}$ + ΔG$_a^{el}$*). It is known that the *ΔG$_a^{el}$* component is entirely entropic (Dragan *et al*., 2004). The entropy factor of the non-electrostatic component (*TΔS$_a^{nel}$*) can then be derived by substracting the non-electrostatic Gibbs energy (*ΔG$_a^{nel}$*) from the total enthalpy of association (*ΔH$_a$*) (*TΔS$_a^{nel}$= ΔH$_a$ - ΔG$_a^{nel}$*). Figure 12a shows that electrostatic and non-electrostatic components of Gibbs energy do not change substantially for both individual HMG-box domains in complex with different

specific (DNA^{ROX1}, DNA$^{HEM13\_1}$, DNA$^{HEM13\_2}$) in front of non-specific (DNA^{TA}) DNA targets.

However, HMG-box A shows a higher non-electrostatic component in comparison with HMG-box B, representing about 50% of total Gibbs energy of binding. This could indicate that HMG-box A participates more actively in the selectivity of promoter sequences during the gene regulation function of Ixr1 protein. Figure 12b shows similar non-electrostatic thermodynamic signatures for both individual HMG-box domains in the binding to DNA^{ROX1}, and slightly differences could be a consequence of refolding events in the complex formation.
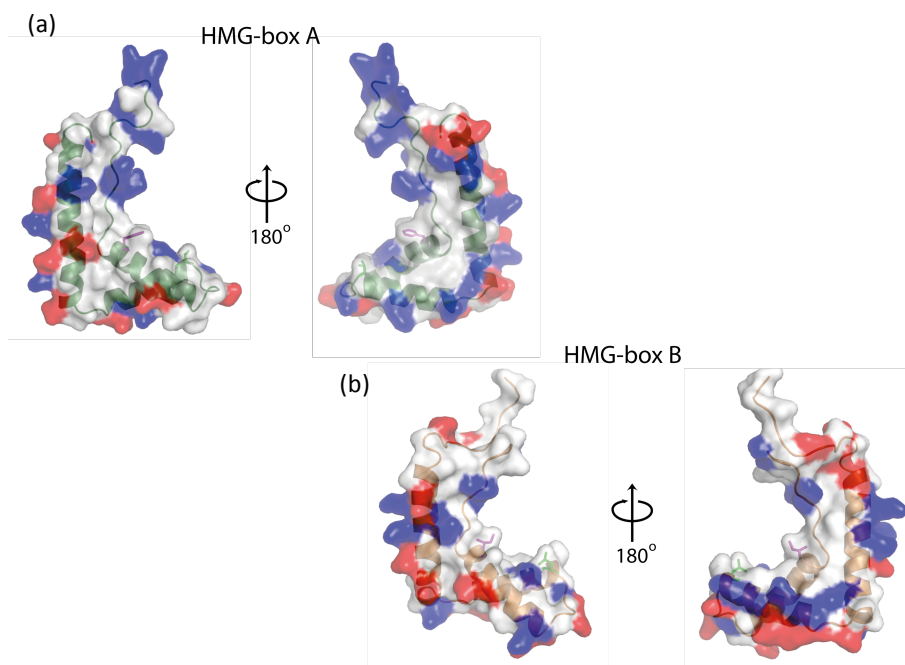
Figure 11. Charge distribution of HMG-box A (a) and HMG-box B (b), showing positive charges as red surfaces and negative charges as blue surfaces.
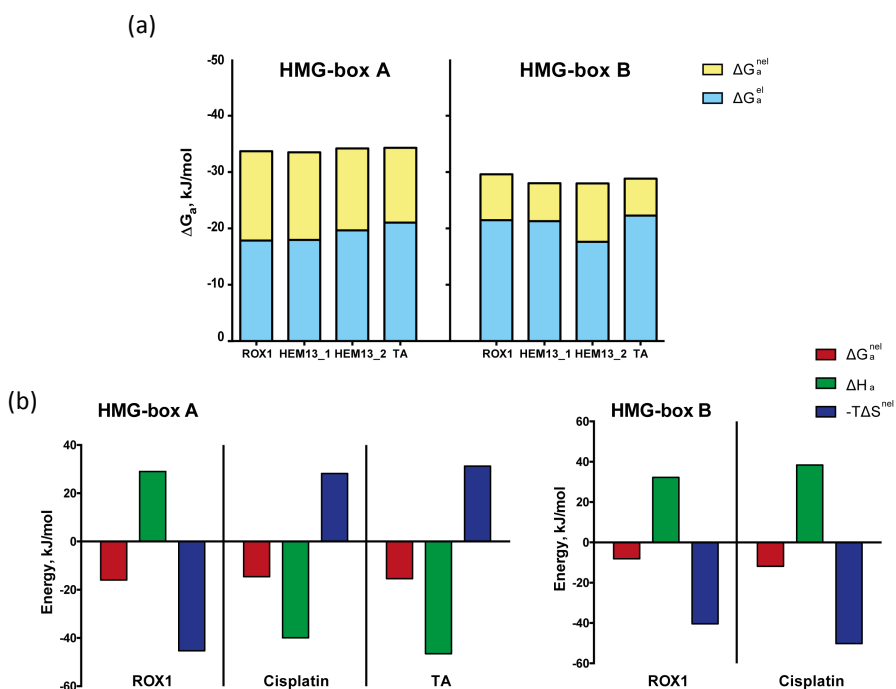
Figure 12. (a) The electrostatic (in blue) and non-electrostatic (in yellow) components of the total Gibbs free energy of binding of HMG-box A and HMG-box B with different linear DNA duplexes. (b) The enthalpic and entropic contributions to the non-electrostatic Gibbs energy of HMG-box A and HMG-box B with various DNA types. The non-electrostatic Gibbs energies of binding are shown by red bars, the enthalpies by green bars and the entropic factors by blue bars.

3.1.4.- Bending properties of both individual HMG-box domains and tandem

Since spectral overlap exists between FAM and TAMRA fluorophores, a decrease in the fluorescence emission of FAM at 520 nm and an increase in the fluorescence emission of TAMRA at 580 occur when the end-to-end distance of the DNA^{ROX1} and therefore the distance between the fluorophores decrease. DNA^{ROX1} bending by HMG-box domains decreases the distance between FAM and TAMRA, and results in an increased FRET effect (figure 13).

The data obtained for DNA^{ROX1} indicate that HMG-box A binding produces a bending of 61° in the linear DNA, and HMG-box B binding bends DNA in only 38°.

Sequential binding of tandem-AB produces a bend angle of 96°.

(a)

(b)

(c)
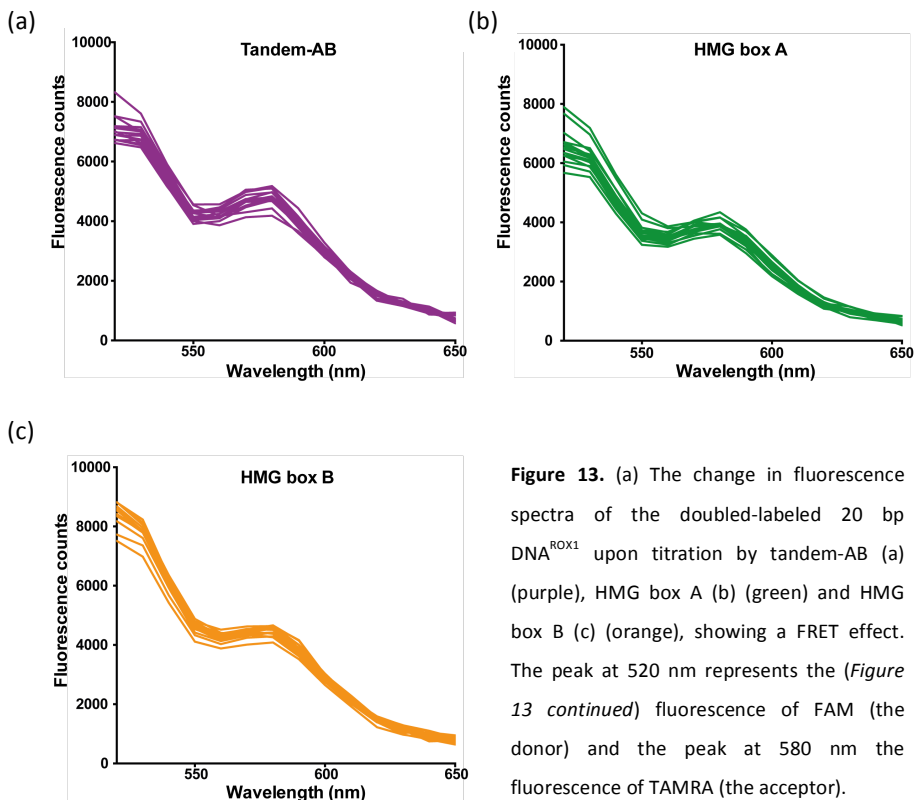
Figure 13. (a) The change in fluorescence spectra of the doubled-labeled 20 bp DNA^{ROX1} upon titration by tandem-AB (a) (purple), HMG box A (b) (green) and HMG box B (c) (orange), showing a FRET effect. The peak at 520 nm represents the (*Figure 13 continued*) fluorescence of FAM (the donor) and the peak at 580 nm the fluorescence of TAMRA (the acceptor).

3.2.- HMG-box B domain needs to be in tandem with A to reach a complete stable folding

A striking difference in the organization of HMG-box A and HMG-box B domains in Ixr1 in comparison to other HMGB proteins with two HMG-box domains is the short linker region between them. In order to determine whether the two tandem HMG-box domains of Ixr1 tandem-AB are independent in solution or depend each other to be correctly folded due to its proximity, we compared 2D NMR ^{1}H-^{15}N HSQC spectra of single HMG-box A and HMG-box B with those of free tandem-AB. The cross-peaks in the ^{1}H-^{15}N HSQC spectrum of tandem-AB (figure 14a) are broader than cross-peaks in the 2D spectrum of single HMG-box domains, indicating that the tumbling of the tandem-AB is slower than individual ones

(figures 15c and 15d). The addition of linear DNA to form tandem-AB+DNA[ROX1] complex does not improve the signal spectra (figure 14b). Although tandem-AB (magenta) peaks can generally be associated with corresponding HMG-box A (green) and HMG-box B (orange) peaks, there are some differences in chemical shift that could indicate that the environment is different in the tandem-AB and individual HMG-box A or HMG-box B domains (figure 14e). In the case of HMG-box A (figure 14c), 2D spectrum peaks fit well with the 2D spectrum of tandem-AB counterpart, indicating little conformational changes in the single domain. Nevertheless, in the case of HMG-box B domain (figure 14d), the ^{1}H-^{15}N HSQC spectrum shows extra peaks in the 8.0-8.5 ppm region that do not have magenta (tandem-AB domain) counterparts (figure 14e), implying that there is some unfolded protein present. Counting the non-amide peaks in ^{1}H-^{15}N HSQC spectrum of HMG-box B domain (figure 14d), it reaches more than one hundred peaks, which far exceeds the number of 71 expected peaks (number of peaks = [number of AAs] - [number of prolines] - [1 nitrogen terminus]). This supports the assertion that the HMG-box B domain folding implies more than one conformation and that it may be less stable due to the absence of the HMG-box A domain.

4.- DISCUSSION

The difference between sequence-specific (SS) and non-sequence-specific (NSS) HMG-box proteins depends on individual DNA-intercalating residues and the global features of the HMG-box domains, which determine the mode of DNA recognition. In this sense, considerable numbers of structural and thermodynamic studies have been carried out in recent years in order to characterize the binding of SS and NSS HMG-box DNA binding proteins (Müller *et al*., 2001; Dragan *et al*., 2003; Kamau *et al*., 2004; Dragan *et al*., 2004; Privalov *et al*., 2007; Privalov *et al*., 2011; Rubio-Cosials *et al*., 2011; Malarkey *et al*., 2012).
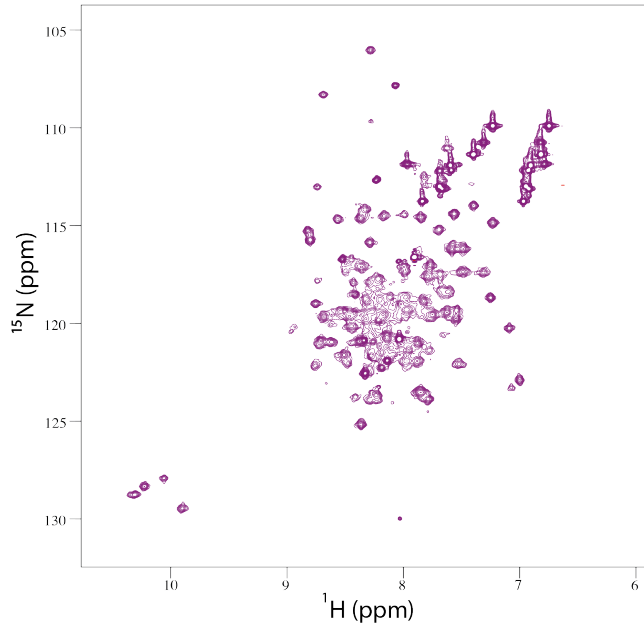
(a)

(b)
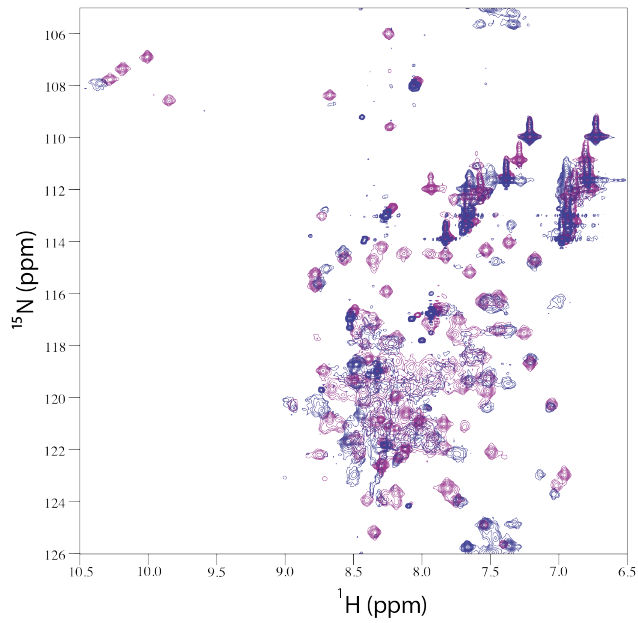
(c)

(d)

(e)
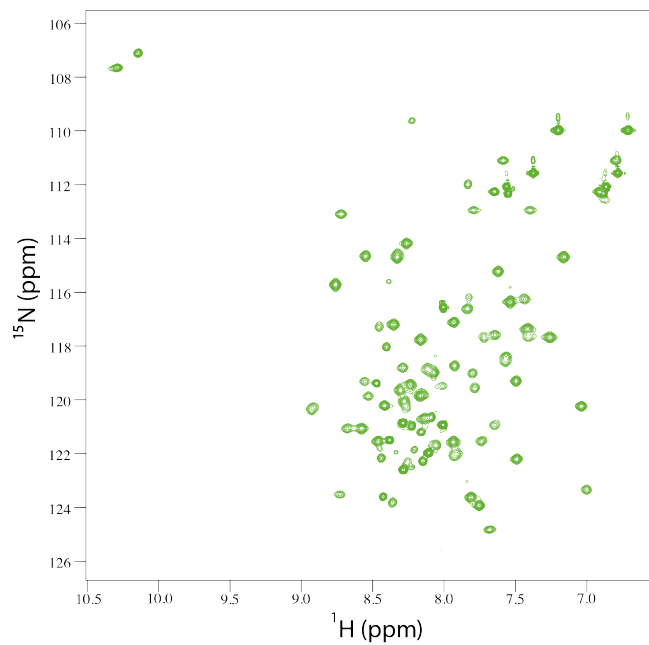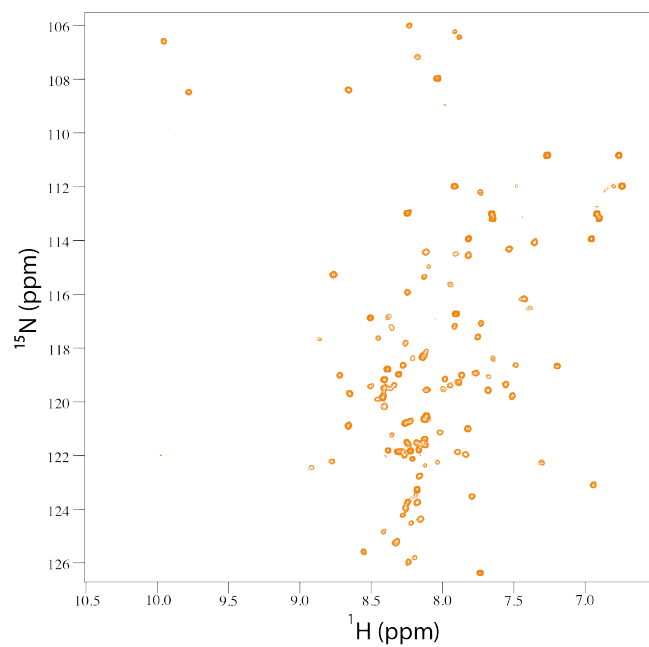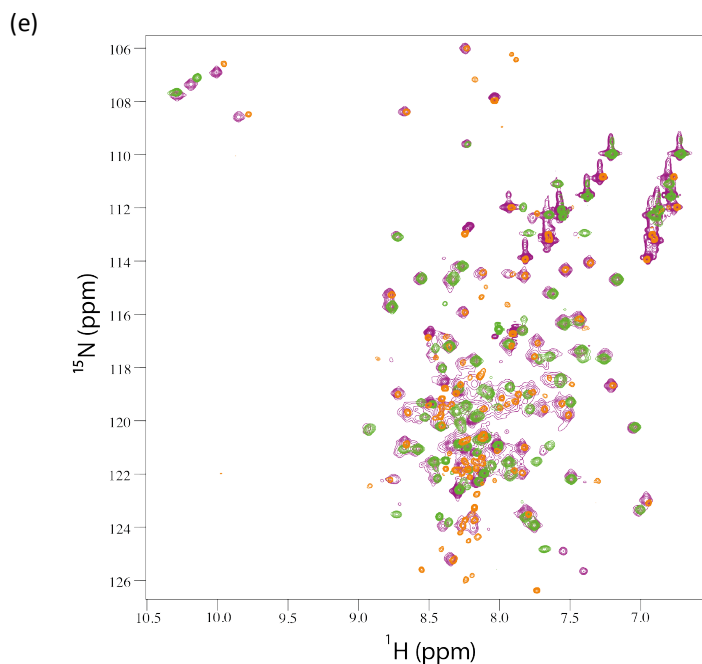
Figure 14. 2D ^{1}H-^{15}N HSQC spectra of free tandem-AB (a), tandem-AB in complex with DNA^{ROX1} (b) and individual HMG-box A (c) and HMG-box B (d). (e) shows the superposition of free tandem-AB (in purple), single HMG-box A (in green) and single HMG-box B (in orange).

In our study of the HMG-box domains of Ixr1, calculated binding affinities of single HMG-box domains to linear DNA (DNA^{ROX1}, DNA$^{HEM13\_1}$, DNA$^{HEM13\_2}$ and DNA^{TA}) obtained by fluorescence anisotropy and isothermal titration calorimetry experiments showed that they are in the range of 10^{-7} M and 10^{-6} M for HMG-box A and HMG-box B, respectively. Similar affinity values were obtained for single HMG-box proteins like human Lef1 or HMG-D from *D. melanogaster* (Dragan *et al.*, 2004) or other proteins with two HMG-box domains in tandem like human HMGB1 (Müller *et al.*, 2001) or Hmo1 from *S. cerevisiae* (Kamau *et al.*, 2004) in their binding to linear DNA. Also similar affinities have been observed for Ixr1 complexed with cisplatin-modified DNA (McA'Nulty *et al.*, 1996).

HMG-box proteins bind to the minor groove of DNA, and the interaction is

characterized by positive changes in enthalpy (Privalov *et al.*, 2009). Indeed, structural studies on Sox2 (Remenyi *et al.*, 2003; Williams *et al.*, 2004), SRY (Werner *et al.*, 1995), HMGD (Murphy *et al.*, 1999) and HMGB1 (Stott *et al.*, 2006) revealed extensive hydrophobic interactions in a highly bent and distorted minor groove of DNA. On the other hand, most proteins that recognize the major groove of DNA have a favorable binding enthalpy (Privalov *et al.*, 2009). The high enthalpy value associated with binding to the minor groove is likely due to the displacement of highly ordered water molecules lining the minor groove of AT-rich sites in the DNA, as well as the energy required for DNA bending, which is not compensated by favorable bond formation (Privalov *et al.*, 2007).

Even though Ixr1 protein was previously described as a transcription factor (Castro-Prego *et al.*, 2010a; Castro-Prego *et al.*, 2010b; Tsaponina *et al.*, 2011; Vizoso Vázquez *et al.*, 2012; Tsaponina *et al.*, 2013) and therefore a SS binding was expected, our results show that both individual HMG-box domains bind to sequence specific promoter sites (DNA^{ROX1}) in an entropy-driven manner, sharing large entropy values that compensate positive enthalpy values of around ≈ 40 kJ mol^{-1} (Tables 7 and 8), which is characteristic of NSS HMG-box domains (Dragan *et al.*, 2004). On the contrary, in sequence-specific interactions extensive van der Waals contacts between apolar groups predominate because of a more densely packed complex interface is formed, which decreases the overall positive enthalpy of binding up to close to zero. Nevertheless, enthalpy values summarized in tables 7 and 8 for binding to linear DNA do not allow a clear distinction of the two HMG boxes of Ixr1 between SS or NSS complexes. SRY, typical representative of SS HMG-box proteins, produce values of 27 kJ mol^{-1} in complex formation with their specific DNA sequence (Dragan *et al.*, 2004) that does not fit to those presented for Ixr1 HMGB boxes (tables 7 and 8).

Analyzing the electrostatic and non-electrostatic Gibbs energies of the binding of Ixr1 HMG-boxes to different linear DNAs (Figure 12), energy component

distribution shows that both HMG-box domains have similar electrostatic components and consequently, differences in affinities with the DNA have to be attributed to the non-electrostatic component. In this sense, competition experiments (Table 9) showed higher sequence preference of HMG-box A to DNA^{ROX1} and DNA$^{HEM13\_1}$ promoter binding sites, meanwhile HMG-box B shows larger sequence preference to DNA$^{HEM13\_2}$ promoter binding site, in correspondence with the non-electrostatic binding values represented in figure 12a.

Biological systems are complex networks that require careful regulation. Cooperativity is an effective mechanism of regulation and provides a way to transfer information, amplify or nullify a response to changes in local concentration and regulate the overall reaction pathway. Cooperativity is a thermodynamic parameter that reflects the influence of one bound protein on the binding affinity of a second, and is a result of protein-protein interactions and/ or short-range conformational changes that are induced in the protein upon binding of the DNA (ligand). Cooperative effects are either positive (synergistic) or negative (interfering), depending on whether the binding of the first ligand increases or decreases the affinity for subsequent ligands. Non-cooperative (additive) binding does not affect the affinity for subsequent ligands and the interaction sites can be considered independent. Previous studies with proteins containing two HMG-box domains arranged in tandem have demonstrated that they bind to DNA with a positive cooperative effect. Mitochondrial human Tfam protein has two HMG-box domains linked by a large flexible α-helical region that confers the unique characteristic of produce a DNA U-turn of 180° in which the binding and bending processes of the first HMG-box domain assist to the binding of the second one (Rubio-Cosials *et al.*, 2011). A positive cooperative relation between both HMG-boxes was also observed in human HMGB1 (Webb *et al.*, 2001) and Hmo1 (Murugesapillai *et al.*, 2014). Here we demonstrate by ITC that tandem-AB region of Ixr1 protein (amino acids 338-510) binds to linear DNA^{ROX1} in a positive

cooperative way. The biphasic ITC thermogram suggests that tandem-AB binds sequentially to DNA, and that HMG-box A binds first because of its higher association constants. Cooperativity is often ascribed to subtle conformational changes in macromolecular structure and protein motions. Proteins are dynamic ensembles of diverse conformations in which allosteric motions might occur even in the absence of ligand (Popovich *et al.*, 2006). Binding to the ligand merely shifts the dynamic equilibrium by preferentially stabilising a particular motion (Furnham *et al.*, 2006). Several experiments carried out in our study suggest that, in order to make possible that the HMG-box B domain could adopt an optimal conformation favourable for lineal DNA binding, a previous structural change on HMG-box A, upon the first step of DNA binding, is necessary. Circular dichroism spectra showed that tandem-AB in complex with DNA^{ROX1} results in an increment in the α-helical content of 7% with respect to the free protein. Furthermore, NMR experiments showed that HMG-box B is less stable in solution than HMG-box A and presents more co-existing conformations than when it is co-arranged with HMG-box A in tandem.

There are experimental situations that cannot be assigned to a particular model. For two or more binding sites, different binding models can give rise to mathematically equal binding equations. In those cases, the discrimination between models cannot be made on the basis of binding data alone and requires extrathermodynamic arguments. Consequently, unless a binding model has been validated, the binding polynomial should be the preferred starting point for the analysis of complex binding situations. In the present work we propose a possible model for tandem-AB binding to DNA (Figure 15). The analysis from the ITC data of tandem-AB determines the overall binding constant (β_j) or stepwise binding constant (K_j) for ligation of the jth site. On the other hand, the individual intrinsic binding constants for HMG-box A (K_{HMGA}) and for HMG-box B (K_{HMGB}) of tandem-AB were determined by using single HMG-box domains separately.
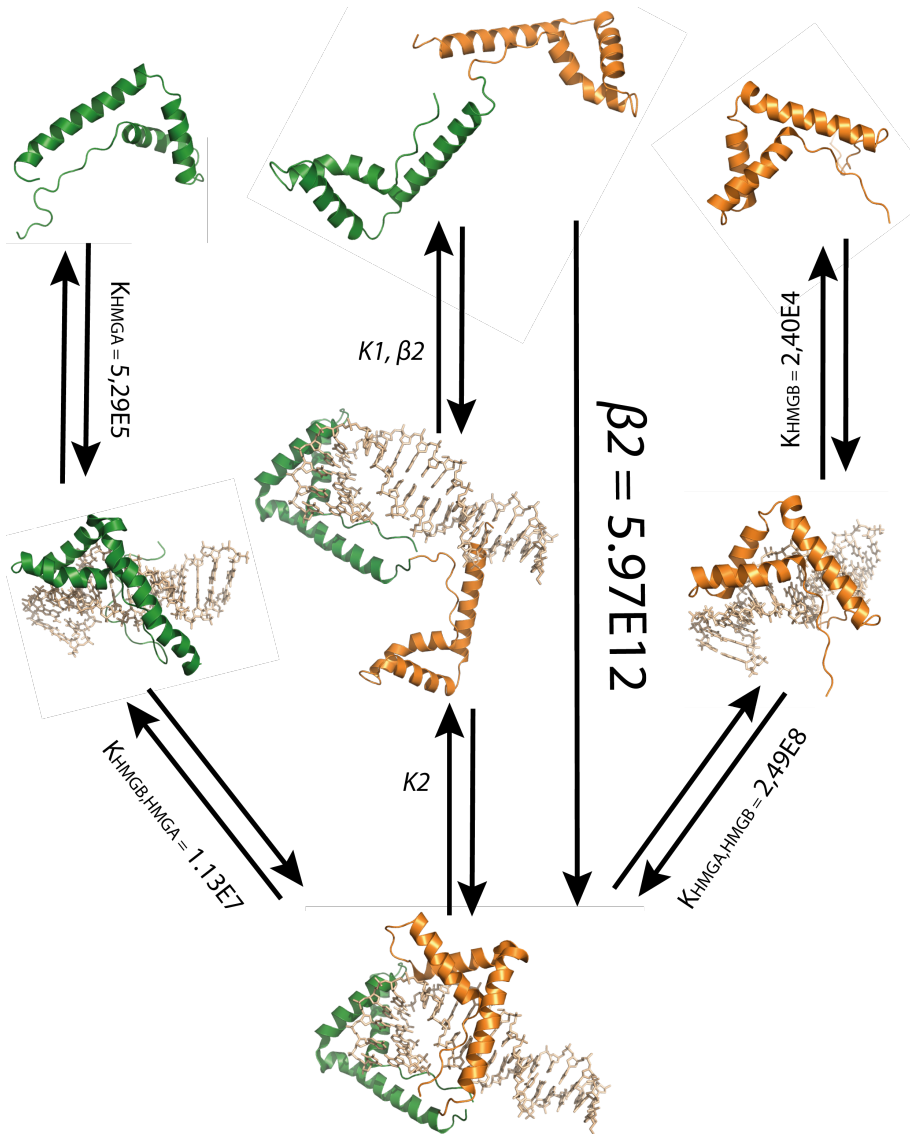
Figure 15. Model proposed for tandem-AB binding to DNA[ROX1]. A microscopic binding model is proposed based on the studies of individual HMG-box A and HMG-box B proteins purified, that determine their microscopic binding constants (K_{HMGA} and K_{HMGB} for HMG-box A and HMG-box B, respectively) and the studies of tandem-AB that determine their overall binding constants, β_1 and β_2. Due to the equilibrium among all states, the equation $\beta_2 = K_{HMGB} \times K_{HMGA,HMGB} = K_{HMGA} \times K_{HMGB,HMGA}$ was used to determine

(*Figure 15 continued*) intrinsic binding constants of HMG-box A in the presence of DNA[ROX1] bounded to HMG-box B ($K_{HMGA,HMGB}$) and of HMG-box B in the presence of DNAROX1 bounded to HMG-box A ($K_{HMGB,HMGA}$). The microscopic cooperativity constant, c_{12}, was determined by the equation: $c_{12} = K_{HMGB,HMGA} / K_{HMGB} = K_{HMGA,HMGB} / K_{HMGA}$.

Because the thermodynamic binding constants are independent of the reaction pathway, the overall binding constant to the two HMG-boxes present in Ixr1 is defined by the equation:

$$\beta_2 = K_{HMGB} \times K_{HMGA,HMGB} = K_{HMGA} \times K_{HMGB,HMGA}$$

where β_2 is the overall binding constant with two binding sites, $K_{HMGA,HMGB}$ is the intrinsic DNA binding constant for HMG-box A when HMG-box B is already bound to the DNA and $K_{HMGB,HMGA}$ is the intrinsic DNA binding constant for HMG-box B when HMG-box A is already bound to the DNA. Thus the microscopic cooperative constant (c_{12}) is determined by the definition of $K_{HMGA,HMGB}/K_{HMGA}$ or $K_{HMGB,HMGA}/K_{HMGB}$.

Our data, derived from the study with the complex formed between tandem-AB and DNA[ROX1], indicate that the DNA first binds to the HMG-box A, with higher DNA binding affinity, producing a bending angle of 61° in the linear DNA. The resulting binding causes a conformational change that propagates to HMG-box B that, along with higher contact probability to contact with DNA, enhance its DNA binding affinity, in which $c_{HMGA,HMGB}$ is estimated to be between ≈470, and causing an additional bend angle up to 96°.

In the cell there are a diversity of distorted DNA targets, including features of folded chromatin structures, sites of DNA damage, tight loops introduced during site-specific recombination and bends introduced during the assembly of nucleoprotein complexes involved in transcription initiation. Subtle preferences of particular proteins for particular targets, as well as relative abundances, may play a crucial role in orchestrating the contributions of HMG-box proteins to nuclear

events and repair. For this reason, additional studies were done to analyze the binding of HMG-box domains of Ixr1 protein to cisplatin-modified DNA (DNA^{Cisplatin}) and four-way junctions (DNA^{Cruciform}). A striking difference found between both HMG-box domains of Ixr1 is the enthalpy sign of the binding to DNA^{Cisplatin} and DNA^{Cruciform}.

In the particular case of DNA^{Cisplatin}, tandem-AB binding showed a large enthalpy changes for both HMG-box domains, with negative favourable values for HMG-box A and positive unfavorable enthalpy binding values for HMG-box B. In the same way, individual HMG-box A also shows a favorable enthalpy of -39.93 kJ/mol, a value close to -52.5 kJ/mol obtained experimentally (Park & Lippard, 2011) and -63 kJ/mol by metadynamic simulations (Nguyen *et al.*, 2014) for HMGB1 complex formation with platinated DNA. Meanwhile, HMG-box B shows an unfavorable enthalpy value of 38.38 kJ/mol and a association constant 30 times lower than HMG-box A. In the case of tandem-AB+DNA^{Cisplatin} complex formation, a feasible explanation is that the binding of HMG-box A to cisplatin-modified DNA adduct (higher affinity) forces HMG-box B to bind to adjacent DNA regions in a entropy driven process, with high degrees of freedom. More difficult to explain is the positive enthalpy value of individual HMG-box B to DNA^{Cisplatin}, and could be indicating large folding rearrangements in the protein or no binding to the DNA adduct, but to adjacent DNA regions. Nevertheless, these alternative possible explanations need to be explored in future experiments.

5.- REFERENCES

Brown A. (2009). "Analysis of cooperativity by isothermal titration calorimetry." International Journal of Molecular Sciences **10**(8):3457-77.

Brown S.J., P.J. Kellett and S.J. Lippard (1993). "Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin." Science **261**(5121):603-5.

Bustin, M. and R. Reeves (1996). "High-mobility-group chromosomal proteins. Architectural components that facilitate chromatin function." Progress in Nucleic Acid Research and Molecular Biology **54**: 35-100.

Calderone C.T. and D.H. Williams (2001). "An enthalpic component in cooperativity: the relationship between enthalpy, entropy, and noncovalent structure in weak associations." Journal of the American Chemical Society **123**(26):6262-7.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, I. Carneiro, I. González-Siso and M.E. Cerdán (2009). "Regulatory factors controlling transcription of Saccharomyces cerevisiae IXR1 by oxygen levels: a model of transcriptional adaptation from aerobiosis to hypoxia implicating ROX1 and IXR1 cross-regulation." Biochemical Journal **425**(1):235-43.

Chow C. S., J.P. Whitehead and S.J. Lippard (1994). "HMG domain proteins induce sharp bends in cisplatin-modified DNA." Biochemistry **33**(50):15124-15130.

Clegg R.M. (1992). "Fluorescence resonance energy transfer and nucleic acids." Methods in Enzymology **211**:353-388.

Cohen S.M., Y. Mikata, Q. He and S.J. Lippard (2000). "HMG-domain protein recognition of cisplatin 1,2-intrastrand d(GpG) cross-links in purine-rich sequence contexts." Biochemistry **39**(38):11771-6.

Freire E., A. Schön and A. Velazquez-Campoy (2009). "Isothermal titration calorimetry: general formalism using binding polynomials." Methods in Enzymology 455:127-55.

Grosschedl R., K. Giese and J. Pagel (1994). "HMG-domain proteins: architectural elements in the assembly of nucleoprotein structures." Trends in Genetics **10**(3):94-100.

Ha J. H., M.W. Capp, M.D. Hohenwalter, M. Baskerville and M.T Record Jr (1992). "Thermodynamic stoichiometries of participation of water, cations and anions in specific and non-specific binding of lac repressor to DNA. Possible thermodynamic origins of the glutamate effect on protein- DNA interactions." Journal of Molecular Biology **228**(1):252-64.

J.O. Thomas and A.A. Travers (2001). "HMG1 and 2, and related 'architectural' DNA-binding proteins." Trends in Biochemical Science **26**(3):167-174.

Kamau E., K.T. Bauerle and A. Grove (2004). "The Saccharomyces cerevisiae high mobility group box protein HMO1 contains two functional DNA binding domains." Journal of Biological Chemistry **279**(53):55234-40.

Lambert J. R., V.W. Bilanchone and M.G. Cumsky (1994). "The ORD1 gene encodes a transcription factor involved in oxygen regulation and is identical to IXR1, a gene that confers cisplatin sensitivity to Saccharomyces cerevisiae." Proceedings of the National Academy of Sciences U.S.A. **91**(15):7345-7349.

Laue T.M., B.D. Shah, T.M. Ridgewa and S.L. Pelletier (1992). "Computer-aided interpretation of analytical sedimentation data for proteins" in Analytical ultracentrifugation in biochemistry and polymer science (Harding, S.E., Rowe, A.J. and Horton, J.C., Eds.), pp. 90-125, The Royal Society of Chemistry, Cambridge, UK.

Li Z., S. Van Calcar, C. Qu, W.K. Cavenee, M.Q. Zhang and B. Ren (2003). "A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells." Proceedings of the National Academy of Sciences U.S.A. **100**(14): 8164-8169.

Louis-Jeune C., M.A. Andrade-Navarro and C. Perez-Iratxeta (2012). "Prediction of protein secondary structure from circular dichroism using theoretically derived spectra." Proteins **80**(2):374-81.

Love J. J., X. Li, D.A. Case, K. Giese, R. Grosschedl and P.E. Wright (1995). "Structural basis for DNA bending by the architectural transcription factor LEF-1." Nature **376**(6543):791-5.

Malarkey C.S., M.E. Churchill (2012). "The high mobility group box: the ultimate utility player of a cell." Trends in biochemical sciences **37**(12):553-562.

Malarkey C.S., M. Bestwick, J.E. Kuhlwilm, G.S. Shadel and M.E. Churchill (2012). "Transcriptional activation by mitochondrial transcription factor A involves preferential distortion of promoter DNA." Nucleic Acids Research **40**(2):614-24.

Mancheño J.M., H. Tateno, I.J. Goldstein, M. Martínez-Ripoll and J.A. Hermoso (2005). "Structural analysis of the Laetiporus sulphureus hemolytic pore-forming lectin in complex with sugars." Journal of Biological Chemistry **280**(17):17251-9.

Manning G. S. (2003). "Is a small number of charge neutralizations sufficient to bend nucleosome core DNA onto its superhelical ramp?" Journal of the American Chemical Society **125**(49):15087-92.

McA'Nulty M.M. and S.J. Lippard (1996). "The HMG-domain protein Ixr1 blocks excision repair of cisplatin-DNA adducts in yeast." Mutat Research **362**(1):75-86.

McA'Nulty M.M., J.P. Whitehead and S.J. Lippard (1996). "Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts in vitro and in vivo." Biochemistry 14;35(19):6089-99.

Müller S., M.E. Bianchi and S. Knapp (2001). "Thermodynamics of HMGB1 interaction with duplex DNA." Biochemistry **40**(34):10254-61.

Murphy E. C., V.B. Zhurkin, J.M. Louis, G. Cornilescu and G.M. Clore (2001). "Structural basis for SRY- dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation." Journal of Molecular Biology **312**(3):481-99.

Murphy F. V.I.V., R.M. Sweet and M.E.A. Churchill (1999). "The structure of a chromosomal high mobility group protein-DNA complex reveals sequence- neutral mechanisms important for non-sequence- specific DNA recognition." EMBO Journal **18**(23):6610-8.

Murugesapillai D., M.J. McCauley, R. Huo, M.H. Nelson Holte, A. Stepanyants, L.J. Maher, N.E. Israeloff and M.C. Williams (2014). "DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin." Nucleic Acids Research **42**(14):8996-9004.

Neidhardt F. C., P.L. Bloch and D.F. Smith (1974). "Culture medium for enterobacteria." Journal of Bacteriology **119**(3):736-47.

Ohndorf U.M., M.A. Rould, Q. He, C.O. Pabo and S.J. Lippard (1999). "Basis for recognition of cisplatin- modified DNA by high-mobility-group proteins." Nature **399**(6737):708-12.

Olmsted M. C., J.P. Bond, C.F. Anderson and M.T Record Jr. (1995). "Grand canonical Monte Carlo molecular and thermodynamic predictions of ion effects on

binding of an oligocation (L8+) to the center of DNA oligomers." Biophysical Journal **68**(2):634-47.

Park S. and S.J. Lippard (2011). "Redox state-dependent interaction of HMGB1 and cisplatin-modified DNA." Biochemistry **50**(13):2567-74.

Perez-Iratxeta C. and M.A. Andrade-Navarro (2011). "K2D2: estimation of protein secondary structure from circular dichroism spectra." BMC Structural Biology 13;8:25.

Pil P. M. and S.J. Lippard (1992). "Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin." Science **256**(5054):234-7.

Privalov PL, A.I. Dragan and C. Crane-Robinson (2011). "Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components." Nucleic Acids Research **39**(7):2483-91.

Qi D. and K.B. Scholthof (2008). "A one-step PCR-based method for rapid and efficient site-directed fragment deletion, insertion, and substitution mutagenesis." Journal of Virology Methods **149**(1):85-90.

Reeves R. (2010). "HMG Nuclear Proteins: Linking Chromatin Structure to Cellular Phenotype." Biochimica et biophysica acta **1799**(1-2):3.

Remenyi A., K. Lins, L.J. Nissen, R. Reinbold, H.R. Scholer and M. Wilmann (2003). "Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers." Genes & Development **17**(16):2048-59.

Rodríguez Lombardero S., A. Vizoso Vázquez, E. Rodríguez Belmonte, M.I. González Siso and M.E. Cerdán (2012). "SKY1 and IXR1 interactions, their effects on cisplatin and spermine resistance in Saccharomyces cerevisiae." Canadian Journal Microbiology **58**(2):184-8.

Schaffner W. (1988). "Gene regulation. A hit-and-run mechanism for transcriptional activation." Nature **336**(6198):427-428.

Schuck P. (2000). "Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." Biophysics Journal **78**(3):1606-19.

Stott K., G.S. Tang, K.B. Lee and J.O. Thomas (2006). "Structure of a complex of tandem HMG-boxes and DNA." Journal of Molecular Biology **360**(1):90-104.

Stros M. (1998). "DNA bending by the chromosomal protein HMG1 and its high mobility group box domains. Effect of flanking sequences." Journal of Biological Chemistry **273**(17):10355-61.

Stros M. (2010). "HMGB proteins: interactions with DNA and chromatin." Biochimica et Biophysical Acta. **1799**(1-2):101-13.

Stuhmeier F., A. Hillisch, R.M. Clegg and S. Diekman (2000). "Fluorescence energy transfer analysis of DNA structures containing several bulges and their interaction with CAP." Journal of Molecular Biology **302**(5):1081-100.

Toney J. H., B.A. Donahue, P.J. Kellett, S.L. Bruhn, J.M. Essigmann and S.J. Lippard (1989). "Isolation of cDNAs encoding a human protein that binds selectively to DNA

modified by the anticancer drug cis-diamminedichloroplatinum(II)" Proceedings of the National Academy of Sciences U.S.A. **86**(21):8328-32.

Travers A. (2000). "Recognition of distorted DNA structures by HMG domains." Current Opinion in Structural Biology **10**(1):102-9.

Treiber D. K., X. Zhai, H.M. Jantzen and J.M. Essigmann (1994). "Cisplatin-DNA adducts are molecular decoys for the ribosomal RNA transcription factor hUBF (human upstream binding factor)." Proceedings of the National Academy of Sciences U.S.A. **91**(12):5672-6.

Tsaponina O., E. Barsoum, S.U. Aström and A. Chabes (2011). "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools." PLoS Genetics **7**(5):e1002061.

Webb M., D. Payet, K.B. Lee, A.A. Travers AA and J.O. Thomas (2001). "Structural requirements for cooperative binding of HMG1 to DNA minicircles." Journal of Molecular Biology **309**(1):79-88.

Weir H. M., P.J. Kraulis, C.S. Hill, A.R. Raine, E.D. Laue and J.O. Thomas (1993). "Structure of the HMG-box motif in the B-domain of HMG1." EMBO Journal (4):1311-9.

Werner M.H., J.R. Huth, A.M. Gronenborn and G.M. Clore (1995). "Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex." Cell **81**(5):705-14.

Williams D.C. Jr., M. Cai and G.M Clore (2004). "Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex." Journal of Biological Chemistry **279**(2):1449-57.

Williamson M. (2012). "How proteins work." Garland Science/Taylor & Francis. ISBN: 9780815344469

Wong B., J.E. Masse and Y.M. Yen (2002). "Giannikopoulos P, Feigon J, Johnson RC. (2002) Binding to cisplatin-modified DNA by the Saccharomyces cerevisiae HMGB protein Nhp6A." Biochemistry **41**(17):5404-14.

Wright P.E. and H.J. Dyson (1999). "Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm." Journal of Molecular Biology **293**(2): 321-331.

Chapter 5

Ixr1 is an intrinsically disordered protein of prion-prone nature

SUMMARY

Ixr1 is a transcriptional regulatory factor previously identified as a player in the response to several stress conditions, such as oxidative stress or hypoxia, as well as in the resistance to cisplatin treatment. Nevertheless, little is known about the structure and mechanism of action of this protein. Here, we cloned and purified several domains of Ixr1. We show that Ixr1 is an intrinsically disorder protein with high aggregation propensity and large disorganized regions flanking the HMG boxes, which conform the DNA binding domains. Indeed Ixr1 aggregation is highly ordered and the protein is able to form amyloids. Amyloid fibrils are one of the most frequent self-templating replicative states among the prions characterized until now. The relationship between Ixr1 function and its prion capabilities is also discussed.

1.- INTRODUCTION

The term prion refers to protein isoforms with self-perpetuating capabilities and dominant phenotypes transmitted through the cytoplasm in a non-Mendelian fashion. A prion is formed when a native protein adopts an alternative conformation (known as the prion fold), which is able to recruit soluble protein isomers and induce them to form the same prion fold. The self-templating replicative state of most biochemically characterized prions is amyloid (Glover *et al.*, 1997; Alberti *et al.*, 2009), although other types of self-propagating protein conformations may also give rise to prion phenomena (Wickner *et al.*, 2007; Brown and Lindquist, 2009).

Aggregation of proteins is produced by several stress conditions affecting protein expression or folding, genetic mutations producing proteins variants or proteins that are natively unfolded. Also, fragments of proteins generated by proteolysis, and unable to correctly fold in the absence of the remaining part of the polypeptide chain, are vulnerable to aggregate. Aggregated forms of proteins are

generally amorphous at the ultrastructural level. They are formed by assemblies of interacting chains of the same of different protein variants, more or less disordered, but that retain the ability to dissociate again. In some cases these aggregates are able to reorganize and form amyloid fibrils, protein aggregates structured in closely packed and highly ordered conformations. One of the most frequent structures characterized in the amyloid fibrils contains a common 'cross-β' hydrogen-bond formation pattern, composed by β-strands oriented perpendicularly to the fibril axis, and therefore providing great stability to the fibrils (Aguzzi *et al.*, 2014). Amyloid fibrils can adopt several proto-filament-twisted morphologies with different radii and stiffness (Glover *et al.*, 1997; Diaz-Avalos *et al.*, 2005). Prion formation and propagation is facilitated by its stability, the mechanism by which the assembly of the fiber happens through nucleated-polymerization and the high specificity of new protein chains for binding to the growing template.

As noted above, some proteins under normal physiological conditions do not fold into globular conformations, the so-called Intrinsically Disordered Proteins (IDPs), or they are disordered only in certain regions of their sequences (i.e., Intrinsically Disordered Regions, IDRs) (Aguzzi *et al.*, 2014). Furthermore, studies about protein sequence conservation estimate that about 30% of eukaryotic proteins contain regions of more than 30 amino acids that do not adopt a defined structure (Aguzzi *et al.*, 2014). IDPs cannot develop the typical functions of folded proteins and their existence defies the classical "structure-function" paradigm. In general, IDRs are characterized by low sequence complexity, with few aromatic and bulky hydrophobic residues and many polar or charged residues (arginine, glutamate, lysine, glutamine, and serine) or structure-breaking amino acids (glycine and proline) (Aguzzi *et al.*, 2014). The structural disorder property of IDPs/IDRs offers several functional advantages, such as fast binding kinetics, versatile regulation by post-translation-modifications (PTMs), due to easy accessibility of modifying enzymes, broad functional spectrum, and specificity without strong

binding (reversibility). All these characteristics are potentially beneficial in signalling and regulatory tasks, when fast responses to changes in the external or internal environment are critical for cellular fitness and survival. As a result, structural disorder is usually high in transcription factors, chromatin-organizing proteins, signalling adaptors and scaffolds, cytoskeletal proteins, and cell-cycle regulatory proteins.

Since IDPs/IDRs play crucial roles in biological processes, misfolding and de-regulation of these proteins, tightly controlled and sequestered in amyloids and amyloid-like inclusions, are related to the pathogenesis of several human and other mammal neurodegenerative diseases. In Alzheimer's disease there are depositions of amyloid-β, tau-protein, and the α-synuclein fragment (non-Abeta component, NAC) (Glenner & Wong, 1984; Ueda *et al.*, 1993). In Down's syndrome non-filamentous amyloid-β deposits are observed (Wisniewski *et al.*, 1985). Parkinson's disease is characterized by the deposition of α-synuclein in a form of Lewy bodies (LBs), or Lewy neurites (LNs); in others, such as Niemann-Pick disease type C, subacute sclerosing panencephalitis, argyrophilic grain disease, myotonic dystrophy and motor neuron disease, an accumulation of tau-protein in form of neurofibrillary tangles (NFTs) is produced (Lee *et al.*, 1991). Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTD) are both characterized by the presence of cytoplasmic inclusions, which are rich in the trans-active response element DNA-binding protein of 43 kDa (Tdp43) (Nass *et al.*, 2012; Barmada *et al.*, 2014)]. In canonical prion diseases, deposition of PrPSC was identified (Prusiner, 2001); Huntington's diseases, neurodegenerative disorders caused by the expansion of GAC trinucleotide repeats that code for polyQ tracts in the gene products (Haass & Selkoe, 2007); transmissible spongiform encephalopathies ('mad cow' disease, Creutzfeldt-Jacob disease, etc.) (Aguzzi & O'Connor, 2010).

Yeast prions have been frequently utilized in prion studies, since they are

non-hazardous and easy to *in vitro* handling using recombinant proteins and faster incubation times. By using yeast prions, various issues have been resolved, such as amyloid formation and growth, the infectivity of amyloid fibrils, and the relationship between amyloid structure and prion strain. Similarly to mammalian prions, several yeast proteins are able to exist either in a normal soluble or in an abnormal amyloid conformation that propagates stably in a single yeast culture and causes characteristic phenotypes (Liebman & Chernoff, 2012). The best-studied prion is the yeast translation-termination factor Sup35 [*PSI*[+]] (Glover *et al*., 1997; Chernoff *et al*., 2000; Tanaka *et al*., 2004). As a consequence of Sup35 sequestration in amyloid fibers, the limited amount of soluble protein is not enough to perform its function and the "stop codon read-through" phenotype is increased, thus producing a variety of new traits that depend upon previously cryptic genetic variation. In addition to Sup35, Rnq [*PIN*[+]]/[*RNQ*[+]] (Derkatch *et al*., 1997; Sondheimer *et al*., 2000; Patel & Liebman, 2007), Ure2 [*URE3*] (Wickner, 1994; Brachmann *et al*., 2005; Nakayashiki *et al*., 2005), Swi1 [*SWI*[+]] (Du *et al*., 2008; Crow *et al*., 2011), Cyc8 [*OCT*[+]] (Patel *et al*., 2009), Mot3 [*MOT3*] (Alberti *et al*., 2009), [*ISP*[+]] (Volkov *et al*., 2002; Rogoza *et al*., 2010) and at least twenty more proteins not yet well characterized (Alberti *et al*., 2009) can form prions in yeast. They are transmitted through the partitioning activity of the chaperone Hsp104 (Crow *et al*., 2011), a homohexameric AAA ATPase that breaks amyloid polymers, thus allowing that replicating prion templates can be faithfully inherited by daughter yeast cells.

Whether prions represent a disease in yeast, or are actually advantageous to cell-population, is a topic of considerable debate (Liebman & Chernoff, 2012). It is known that overproduction of Sup35 protein its toxic to [*PSI*[+]] strains (Chernoff *et al*., 1992; Vishveshwara & Liebman, 2009) and Rnq1 overproduction is also toxic to [*PIN*[+]] strains (Douglas *et al*., 2008). In the same way, [URE3] prions (Schwimmer & Masison, 2002) or a combination of the [*PSI*[+]] prions with the tRNA suppressor *SUQ5* can induce the stress response to heat or ethanol (Eaglestone *et al*., 1999;

Jung *et al*., 2000). On the other hand, prions can be regarded as "bet-hedging devices" (Halfmann & Lindquist, 2010), increasing the reproductive fitness of yeast exposed to fluctuating environments by creating different subpopulations with distinct phenotypic states.

The group of approximately two dozens of prionogenic proteins discovered to date in yeast is enriched in proteins with functions associated to the processing of information, including transcription factors and RNA-binding proteins (Alberti *et al*., 2009; Halfmann & Lindquist, 2010). When yeast cells experience stress that affects protein homeostasis and the protein folding machinery is compromised, the rates at which prions appear and disappear abruptly increase (Tyedmers *et al*., 2008). This facilitates the appearance and exploration of alternative phenotypes that, with reasonable frequency, are beneficial (Halfmann & Lindquist, 2010). The first prion protein that was proposed to be advantageous for increased survival to stress was Sup35, since more than 25% of derived phenotypes are advantageous under particular growth conditions (Tyedmers *et al*., 2008). Thus, prion-containing yeast cells are able to grow in the presence of antibiotics, metals and other toxic conditions, or with different carbon or nitrogen sources, depending on the genetic background (Tyedmers *et al*., 2008). Also, the [URE3] prion state utilizes constitutively poor nitrogen sources (Shorter & Lindquist, 2005).

The *IXR1* gene of *Saccharomyces cerevisiae* encodes for a protein of 67 kDa that contains three poly-glutamine regions and two HMG-box domains that bind to DNA (Lambert *et al*., 1994). It was first characterized by its ability to bind to modified DNA containing intrastrand cross-links formed by the anticancer drug cisplatin (McA'Nulty and Lippard, 1996; McA'Nulty *et al*., 1996), masking the DNA adduct to the action of the Nucleotide Excision Repair (NER) system and conferring cellular resistance to the drug (McA'Nulty & Lippard, 1996). Nevertheless, Ixr1 protein is mainly known as a transcription factor that participates in the response

to several stress conditions, such as oxidative stress (Castro-Prego *et al*., 2010a) or hypoxia (Bordineaud *et al*., 2000; Castro-Prego *et al*., 2010b). Ribonucleotide reductase (Rnr) activity represents the rate limiting step in the biosynthesis of the dNTPs that are required for DNA synthesis, for example in repair processes that follow after a DNA damage. In *S. cerevisiae*, *RNR1* and *RNR3* encode the largest subunit of RNR, while *RNR2* and *RNR4* encode the small subunit. Recently it has been reported that cisplatin resistance observed in mutants *Δixr1* follows a pre-activation of the control systems of genomic integrity through the Mec1-Rad53-Dun1 via, regulating transcriptionally the levels of intracellular Rnr1 (Tsaponina *et al*., 2011; Tsaponina *et al*., 2013). Ixr1 also participates in the transcriptional regulation of genes that control the response to changes in the intracellular oxygen levels (Vizoso-Vázquez *et al*., 2012) and oxidative stress (Castro-Prego *et al*., 2010a).

Glutamine/asparagine-rich (Q/N-rich) domains are found in all known yeast prions as part of the "prion-forming-domain" (PFD) required for prion formation and propagation. Alberti and co-workers were the first to systematically test whether compositional similarity to known PFDs is sufficient to distinguish between Q/N-rich proteins that form prions and those that do not (Alberti *et al*., 2009). Implementing a hidden Markov model based on PFDs experimentally described, they predicted several prion-like proteins with high probability to form prions and 18 among them were experimentally verified as prion proteins by diverse approaches (Alberti *et al*., 2009). Although Ixr1 appeared as a good candidate with a high score (13th position of one hundred candidates) in the work of Alberti and col. (2009) experimental evidences of its capacity to form amyloids or prion characteristics were not reported. Based on this prediction and also considering the difficulties found in our attempts to get ordered crystals for this protein, in the present study we check the possibility that Ixr1 protein could act as a prion. We first show that purified Ixr1 protein is partially unstructured with large IDRs. Furthermore, we show that these IDRs can aggregate into fibrillar amyloids *in*

vitro, and that can act as a "prion-forming-domain" *in vivo*, showing the formation of intracellular "foci" that can be transmitted, as evidenced using the Ade+ phenotypic switching system. The prion characteristics of Ixr1 are discussed in relation to its trans-activation functions inside the cell.

2.- MATERIAL AND METHODS

2.1.- Strains and culture media

The *Saccharomyces cerevisiae* and *Escherichia coli* strains obtained from different sources are summarized in table 1.

Table 1. Yeast and bacteria strains used in this work

Strain	Genotype	Source
W303	*MATa; ura3-52 trp1D2 leu2-3,112 his3-1 ade2-1 can1-100*	EUROSCARF[a]
BJ3505	*pep4::HIS3 prb-D1.6R HIS3 lys2-2008 trp1-D101 ura3-52 gal2 can1*	Eastman Kodak Company
YJW509	*MATα, leu2- 3,112; his3-11,-15; trp1-1; ura3-1; ade1-14; can1-100; [psi-]; [pin-]*	S. Lindquist's Lab. Osherovich *et al*., 2004
YJW584	*MATa, leu2-3,112; his3-11,-15; trp1-1; ura3-1; ade1-14; can1-100; [psi-]; [PIN+]*	S. Lindquist's Lab. Osherovich *et al*., 2004
780-1D	*MATα kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 sup35*	D. Ross's Lab. Song *et al*., 2005
YRS100	*MATa, leu2-3,112; his3-11,-15; trp1-1; ura3-1; ade1-14; can1-100; [psi-]; [PIN+]; sup35::KanMX4; pAG426GPD-SUP35C*	S. Lindquist's Lab. Alberti *et al*., 2009
YER282	*MATa kar1-1 SWQ5 ade2-1 his3 leu2 trp1 ura3 arg1::HIS3 sup35::KanMx pER186*	D. Ross's Lab. Ross *et al*., 2005
PJ69-4A	*MATa trpl-901 leu2-3,112 ura3-52 his3-200 ga14A ga18OA LYSZ::GALl-HIS3 GAL2-ADE2 metZ::GAL7-lacZ*	EA Craig's Lab. James *et al*., 1996
BL21 (DE3)	F- *ompT hsdS$_B$ (r$_B$- m$_B$-) gal dcm* (DE3)	Novagen®
Rosetta 2(DE3)pLysS	F- *ompT hsdS$_B$(r$_B$- m$_B$-) gal dcm* (DE3) pLysSpRARE2 (Cam^R)	Novagen®

[a]http://www.rz.uni-frankfurt.de/FB/fb16/mik ro/euroscarf

Yeast cells were grown at 30 ^{o}C in rich media, YPD (2% glucose, 2% Bacto peptone, 1% yeast extract) or complete media (CM) lacking particular amino acids and containing either 2 % D-glucose (SD), 2 % D-galactose (SGal), or a mix of raffinose and galactose (SRafGal, 1 % each), prepared as described previously (Zitomer & Hall, 1976). Plates used for prion curing contained 5 mM guanidine hydrochloride (GdmHCl).

Bacterial cells were grown at 37 ^{o}C in Luria-Bertani (LB) or 2xYT (1,6% Bacto tryptone, 0,8% Yeast Extract, 0,5% sodium chloride) supplemented with suitable selectable antibiotics.

2.2.- Molecular cloning

DNA oligonucleotides were purchased from Isogen Lifes Sciences, Inc. The primers used were as follows (table 2):

Table 2. Oligonucleotides used in this study.

Name	Sequence[a,b,c]	Strand (W/C)[d]	Added site	Gene	Hybridization position[e]
ECV681ixr1f	cgga**gagctc**AACACCGGTATCTCGCCC	W	SacI	*IXR1*	+4
ECV682r	cggc**ctcgag**TTATTCATTTTTTATGATCGAACC	C	XhoI	*IXR1*	+1794
AVV150	cggc**ctcgag**TTATTACTGTTGTTGCTGCTGTTGC	C	XhoI	*IXR1*	+1008
ECV683hmgf	cgga**gagctc**CCAGTGGTGAAGAAATTATCTTC	W	SacI	*IXR1*	+1012
ECV785AV	cggc**ctcgag**GGTTGGGTTACCGTTTGG	C	XhoI	*IXR1*	+1527
AVV295	**cagggacccggt**ATGTCGGATTCAAACCAAGGCAAC	W	-	*SUP35*	+1
AVV296	**cgaggagaagcccgg**TTATTACTGCTTTTGTTGCTTTTGAAAGTCGTTC	C	-	*SUP35*	+405
AVV290	**cagggacccggt**AACACCGGTATCTCGCCCAAAC	W	-	*IXR1*	+4
AVV291	**cgaggagaagcccgg**TTATTATTCATTTTTTATGATCGAACCATTTG	C	-	*IXR1*	+1794
AVV292	**cgaggagaagcccgg**TTATTACTGTTGTTGCTGCTGTTGCTGTTG	C	-	*IXR1*	+1008
AVV293	**cgaggagaagcccgg**TTATTAGTTATTGCCGCTACTAACATTGTTGTTG	C	-	*IXR1*	+600
AVV294	**cagggacccggt**AGTAACAACAACAGTAACAACAACAATGTTAG	W	-	*IXR1*	+541
ECV657	gggg**ggatcc**ACACCGGTATCTCGCCCAAA	W	BamHI	*IXR1*	+5

Table 2. (continued)

Name	Sequence[a,b,c]	Strand (W/C)[d]	Added site	Gene	Hybridization position[e]
ECV658	gggg**gtcgac**CATTTTTTATGATCGAACCATTTGTAGTGTAGTG	C	SalI	*IXR1*	+1786
AVV329	**gtcaaggagaaaaaaccccggattctagaactagt**<u>gaaggagataacaaaatg</u>GTCGGATTCAAACCAAGGCAAC	W	-	SUP35	+1
AVV330	**accaccccggtgaacagctcctcgccccttgctcac**CTGCTTTTGTTGCTTTTGAAAGTCG	C	-	SUP35	+405
AVV331	**gtcaaggagaaaaaaccccggattctagaactagt**<u>gaaggagataacaaaatg</u>AACACCGGTATCTCGCCCAAAC	W	-	*IXR1*	+1
AVV332	**accaccccggtgaacagctcctcgccccttgctcac**TTCATTTTTTATGATCGAACCATTTGTAG	C	-	*IXR1*	+1794
AVV333	**accaccccggtgaacagctcctcgccccttgctcac**CTGTTGTTGCTGCTGTTGCTGTTG	C	-	*IXR1*	+1008
AVV334	**accaccccggtgaacagctcctcgccccttgctcac**GTTATTGCCGCTACTAACATTGTTGTTG	C	-	*IXR1*	+600
AVV335	**gtcaaggagaaaaaaccccggattctagaactagt**<u>gaaggagataacaaaatg</u>AGTAACAACAACAGTAACAACAACAATGTTAG	W	-	*IXR1*	+541
AVV336	**cagcacttccaccacctggtccacctaagctt**TTCATTTTTTATGATCGAACCATTTGTAG	C	-	*IXR1*	+1794
AVV337	**cagcacttccaccacctggtccacctaag**CTTCTGTTGTTGCTGCTGTTGCTGTTG	C	-	*IXR1*	+1008
AVV338	**cagcacttccaccacctggtccacctaagctt**GTTATTGCCGCTACTAACATTGTTGTTG	C	-	*IXR1*	+600
AVV339	**ggggacaagtttgtacaaaaaagcaggcttc**g<u>aaggagataacaaaatg</u>AACACCGGTATCTCGCCCAAAC	W	-	*IXR1*	+1
AVV340	**ggggaccactttgtacaagaaagctgggtc**TTCATTTTTTATGATCGAACCATTTGTAG	C	-	*IXR1*	+1794
AVV341	**ggggaccactttgtacaagaaagctgggtc**CTGTTGTTGCTGCTGTTGCTGTTG	C	-	*IXR1*	+1008
AVV342	**ggggaccactttgtacaagaaagctgggtc**GTTATTGCCGCTACTAACATTGTTGTTG	C	-	*IXR1*	+600
AVV343	**ggggacaagtttgtacaaaaaagcaggcttc**g<u>aaggagataacaaaatg</u>AGTAACAACAACAGTAACAACAACAATGTTAG	W	-	*IXR1*	+541
AVV344	**ggggacaagtttgtacaaaaaagcaggcttc**g<u>aaggagataacaaaatg</u>GTCGGATTCAAACCAAGGCAAC	W	-	*SUP35*	+1
AVV345	**ggggaccactttgtacaagaaagctgggtc**CTGCTTTTGTTGCTTTTGAAAGTCG	C	-	SUP35	+405

[a]Lower case, sequences added for restriction enzyme digestions.
[b]Restriction sites and homology regions are in bold.
[c]Shine-Dalgarno and Kozak sequences are underlined.
[d]Numbering is considering +1 for the adenine in the first start codon.
[e]W: Watson strand; C: Crick strand.

Table 3. Constructions created in the present work

Construct name	Plasmid
pKLSL_IXR1_complete[a]	kan[r] pBR322 ori lacI His-tag LS150-tag T7-lac-promoter IXR1(1-597)[f]
pKLSL_IXR1_HMGs[a]	kan[r] pBR322 ori lacI His-tag LS150-tag T7-lac-promoter IXR1(338-510)[f]
pKLSL_IXR1_Nterm[a]	kan[r] pBR322 ori lacI His-tag LS150-tag T7-lac-promoter IXR1(2-336)[f]
LIC1.2_SUP35N[b]	amp[r] ColE1 ori His-tag 3C-protease-site T7-lac-promoter SUP35(1-135)[g]
LIC1.2_IXR1_PrD1[b]	amp[r] ColE1 ori His-tag 3C-protease-site T7-lac-promoter IXR1(1-597)[f]
LIC1.2_IXR1_PrD2[b]	amp[r] ColE1 ori His-tag 3C-protease-site T7-lac-promoter IXR1(2-336)[f]
LIC1.2_IXR1_PrD3[b]	amp[r] ColE1 ori His-tag 3C-protease-site T7-lac-promoter IXR1(2-200)[f]
LIC1.2_IXR1_PrD4[b]	amp[r] ColE1 ori His-tag 3C-protease-site T7-lac-promoter IXR1(181-336)[f]
pGBD-C$_2$_IXR1_complete[c]	amp[r] ori 2μm *GAL4*-BD *TRP1* IXR1(1-597)[f]
pAG426GAL_SUP35N_EGFP[d]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site GAL1-promoter Cterm-EGFP *URA3* SUP35(1-135)[g]
pAG426GAL_IXR1_PrD1_EGFP[d]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site GAL1-promoter Cterm-EGFP *URA3* IXR1(1-597)[f]
pAG426GAL_ IXR1_PrD2_EGFP[d]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site GAL1-promoter Cterm-EGFP *URA3* IXR1(1-336)[f]
pAG426GAL_ IXR1_PrD3_EGFP[d]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site GAL1-promoter Cterm-EGFP *URA3* IXR1(1-200)[f]
pAG426GAL_ IXR1_PrD4_EGFP[d]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site GAL1-promoter Cterm-EGFP *URA3* IXR1(181-336)[f]
pAG415ADH_SUP35N_SUP35C[e]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site ADH1-promoter Cterm-EGFP *LEU2* SUP35(1-135)[g]
pAG415ADH_IXR1_PrD1_SUP35C[e]	amp[r] cam[r] ori 2μm ori pBR322 ccdB-site ADH1-promoter Cterm-EGFP *LEU2* IXR1(1-597)[f]

[a]pKLSL150 plasmid obtained from JM Mancheño's lab (Mancheño *et al*., 2005)
[b]pETNKI-his3C-LIC-Amp plasmid obtained from the Protein Facility of The Netherlands Cancer Institute (NKI) Amsterdam (http://proteinfacility.nki.nl)
[c]pGBD-C$_2$ plasmid obtained from EA Craig's lab (James *et al*., 1996)
[d]pAG426GAL-ccdB-EGFP plasmid obtained from S. Lindquist's lab (Alberti *et al*., 2009) through Addgene (www.addgene.org)
[e]pAG415ADH-ccdB-SUP35C plasmid obtained from S. Lindquist's lab (Alberti *et al*., 2009)
[f]P33417 Uniprot accession number for IXR1 protein
[g]P05453 Uniprot accession number for SUP35 protein

PCR reactions were performed using Vent polymerase (*NewEngland Biolabs*) and genomic DNA from *Saccharomyces cerevisiae* BJ3505 strain as template. The PCR products were gel-purified and inserted into the different plasmids by several procedures. The constructs obtained in this work are summarized in table 3.

Constructs obtained from pKLSL150 and pGBD-C$_2$ plasmids were cloned by ligation between *SacI* and *XhoI* or *BamHI* and *SalI* restriction sites respectively, with Speedy ligase (NZYTECH).

Constructs obtained from pETNKI-his3C-LIC-Amp were cloned following the Ligation-Independent Cloning (LIC) procedure. After the plasmid was linearized by *Kpn*I, compatible ends between plasmid and inserts were generated by T4 polymerase treatment in presence of dTTP or dATP, respectively, and transformed in *E. coli* competent cells to repair the nicks.

Constructs obtained from pAG426GAL-ccdB-EGFP were cloned by homologous recombination (Alberti *et al.*, 2009) in *S. cerevisiae* BJ3505 competent cells, using primers for PCR amplification carrying 30-35 nucleotide tails with the specific sequence required for homologous recombination with the corresponding plasmid.

2.3.- Protein expression and purification

The proteins obtained from pKLS150 constructions were expressed in BL-21(DE3) cells (Novagen) or Rosetta^{TM} 2(DE3)pLysS cells (Novagen). Cultures growing in 2xYT medium were treated with 1 mM IPTG to induce the expression during 3 hours at 37 $^{\circ}$C and 200 rpm of shaking. After expression, cell pellets were collected and lysed by sonication in high salt lysis buffer (50 mM sodium phosphate buffer pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 2X complete protease inhibitor cocktail from Roche™). After clarification by centrifugation 30 minutes at 23000 x

g, lysates were passed through Sepharose CL-6b resin (Sigma-Aldrich) packed into a Tricorn column (GE Healthcare) equilibrated in wash buffer A (50 mM sodium phosphate buffer pH 6.9, 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) and the aid of a peristaltic pump (GE Healthcare). Proteins were eluted in an AKTA-prime plus (GE Healthcare) by linear gradient from 0% to 100% of buffer elution A (50 mM potassium phosphate buffer, pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 300 mM lactose) and loaded on a HisTrap HP 5 mL column (GE Healthcare) equilibrated in wash buffer B (50 mM sodium phosphate buffer, pH 6.9, 200 mM NaCl, 2 mM dithiothreitol). After elution by linear gradient from 0% to 100% of buffer elution B (50 mM potassium phosphate buffer pH 6.9, 1 M NaCl, 2 mM dithiothreitol and 300 mM imidazole), the different polypeptides were dialyzed (50 mM sodium phosphate buffer, pH 6.9), 200 mM NaCl, 2 mM dithiothreitol and 1 mM EDTA) and simultaneously digested with TEV protease (Sigma-Aldrich) for 16 hours and 8 ºC. Next day, proteins digested were further purified using gel filtration chromatography using a Hi-load Superdex 200 16/60 column (GE Healthcare) pre-equilibrated with running buffer (50 mM potassium phosphate buffer, pH 6.9, 100 mM KCl, 2 mM dithiothreitol and 1 mM EDTA). The proteins were concentrated by ultrafiltration using Amicon® Ultra-15 Centrifugal Filters, 10 kDa (Merk-Millipore). The homogeneity of the purified protein sample was checked by SDS-PAGE (Laemmli, 1970) and protein concentrations were determined by absorbance at 280 nm using extinction coefficients calculated with Protparam (http://web.expasy.org/protparam/).

The proteins obtained from pETNKI-his3C-LIC-Amp constructions were expressed in Rosetta™ 2(DE3)pLysS cells (Novagen) and purified as previously described (Alberti *et al.*, 2009). Cultures growing in 2xYT medium were treated with 1 mM IPTG to induce the expression during 3 hours at 37 ºC and 200 rpm of shaking. After expression, cell pellets were collected and re-suspended in lysis buffer (7 M guanidine hydrochloride, 100 mM K_2HPO_4 pH 8.0, 5 mM imidazole, 300 mM NaCl, 5 mM 2-mercaptoethanol) for 1 hour at room temperature and constant

circular rotation. Lysates were then cleared for 30 min at 23000 x g and loaded onto Econo-Pac® Chromatography columns (Bio-Rad) for purification using HisPur[TM] Ni-NTA resin (Thermo Scientific), according to the manufacturers' instructions. Proteins were eluted (8 M urea, 100 mM NaOAc/HOAc pH 4, 5 mM β-mecaptoethanol), precipitated by methanol-chloroform and resuspended finally in resuspension buffer (7 M GdmHCl, 100 mM K_2HPO_4, pH 5.0, 300 mM NaCl, 5 mM EDTA, 5 mM TCEP). Protein concentrations were determined by measuring absorption at 280 nm using calculated extinction coefficients.

2.4.- Size-exclusion chromatography

The molecular weight and Stokes radius of Ixr1 protein were determined by size-exclusion chromatography. The chromatographic separation was done on a Superdex 200 increase 3.2/300 column (GE Healthcare) connected to a ÄKTA-prime purifier system (GE Healthcare), pre-equilibrated with running buffer (50 mM potassium phosphate buffer, pH 6.9, 100 mM KCl, 2 mM dithiothreitol and 1 mM EDTA) and calibrated using the protein standards from Bio-Rad (bovine thyroglobulin bovine (670 kDa; Stokes radius: 7.8 Å); bovine γ-globulin (158 kDa; 51 Å); chicken ovalbumin (44 kDa; Stokes radius: 31.2 Å); equine myoglobin (17 kDa; Stokes radius: 19 Å); Vitamin B_{12} (1.35 kDa)). Blue dextran (≈2 MDa) was used to determine the void volume (V_0) of the column. Flow rate was 100 μL/min and elution of proteins was monitored by absorbance at 280 nm. The relative molecular weight of the elution peak of Ixr1 protein was then estimated using a standard curve obtained by plotting the logarithm of the molecular weight of the known proteins against K_{av} ($K_{av} = (V_e-V_0)/(V_t-V_0)$) being V_e the volume value of elution for each protein and Vt the total volume. Estimation of the Stokes radius (R_S) of Ixr1 protein was obtained plotting the linear dependency of log R_S versus K_{av} (Uversky *et al*., 1993; Le Maire *et al*., 2008).

2.5.- Analytical ultracentrifugation

Sedimentation velocity experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance and interference optics. 400 μL of a protein solution at 0.8 mg/mL extensively dialyzed was placed in the sample compartment of an epon double-sector centerpiece and 400 μL of the same solution without protein was placed in the reference compartment. The samples were centrifuged at 293 K and 13100 rpm using an An60-Ti rotor. Scans were acquired using both interference and absorbance optical systems without any significant systematic variation, indicating the absence of signal contributions from residual unmatched buffer components. Sedimentation coefficients were corrected using the software package SEDNTERP (http://www.rasmb.bbri.org/) (Laue *et al*., 1992). Sedimentation coefficient distributions c(s) and molar mass distribution c(M) were calculated using the program SEDFIT (Schuck *et al*., 2000).

2.6.- Blue native PAGE

Blue Native PAGE (BN-PAGE) was performed as previously described (Wittig *et al*., 2006), running linear 4%-13% (w/v) polyacrylamide gradient (gel dimensions of 83 mm x 64 mm x 0.75 mm). Protein samples were supplemented with a tenfold concentrated loading dye (75 mM Imidazole, 0,1% Ponceau S, 50% glycerol, (w/v)). The electrophoresis was started at 120 V for thirty minutes and continued for four hours at 6 mA at room temperature. Gels were stained with Coomassie brilliant blue.

Additional supporting SDS-PAGE was performed. Dilutions 1:1 of Ixr1 protein were done and then, each dilution was treated with 150 μM glutaraldehyde to stabilize protein oligomerization state. Polyacrylamide gel (5% w/v, 0.5% SDS) (gel dimensions of 160 mm x 200 mm x 1.5 mm) was running at 150 V during 8 hours and stained with Coomassie brilliant blue.

2.7.- Far-UV circular dichroism

Circular dichroism (CD) experiments were carried out in a Jasco J-810 spectropolarimeter equipped with a Peltier PTC-423S system. Samples were dialyzed extensively against 10 mM K_2HPO_4, pH 7, 50 mM NaF, 1 mM β-mercaptoethanol and transferred to a cuvette with a 1-mm path length. Spectra were collected at 15 ^{o}C and a bandwidth of 1 nm in continuous scanning mode, with 50 nm/s scanning rate, and a response time of 2 s. From raw data, collected spectra were buffer-subtracted and converted from millidegrees (Θ_{obs}, mdeg) to molar ellipticity ([Θ], degrees cm^2 $dmol^{-1}$ $residue^{-1}$) using the equation [Θ] = (Θ_{obs} x M) / (10 x l x C), where M is the protein mean residue molecular weight, l is the optical path length of the cuvette in cm, and C is the concentration of the protein in mg mL^{-1}. Estimations of secondary structure content were calculated by deconvolution of the far-UV CD spectra using the CDNN program (Gerald Böhm, Institut für Biotechnologie, Martin- Luther-Universität Halle-Wittenberg, Germany) as well as the CONTIN (Provencher and Glöckner, 1981), SECON3 (Sreerama and Woody, 2000) and CDSSTR (Johnson, 1999) procedures contained in the Dichroweb utilities (http://dichroweb.cryst.bbk.ac.uk/html/ home.shtml) (Whitmore and Wallace, 2004; Whitmore and Wallace, 2008). The SP43 basis set was used to deconvolute the CD spectrum. Estimates of percent secondary structure obtained from the four methods were averaged and standard deviations calculated.

2,2,2-Trifluoroethanol (Sigma-Aldrich) titrations were done adding increasing concentrations to the same sample. DNA[ROX1] was obtained by mixing the complementary oligonucleotides (AVV190, 5'-AGGGCCTATTGTTGCTGCCT; and AVV191, 5'-AGGCAGCAACAATAGGCCCT) in equivalent molar amounts, heating to 95 ^{o}C for 5 minutes and cooling slowly to room temperature. Solutions of DNA for the experiments were prepared by extensive dialysis against the solvent, as required. Ixr1+DNA[ROX1] complex was prepared by mixing 1 protein equivalent with 1,5 DNA[ROX1] equivalents and incubated 1 hour at 15 ^{o}C.

The temperature dependence of the circular dichroism spectra of the different proteins was determined upon continuous heating with a rate of 1 K min^{-1}, in the temperature range 5-95 ^{o}C, at 222 nm wavelength using a 0.1 cm path length sealed cell. Ellipticity signals were plotted as a function of temperature and a non-linear Boltzmann fit was performed; the melting temperature was calculated as the maximum of the first derivative of this curve using GraphPad Prism 6.0 (GraphPad software).

2.8.- Limited proteolysis

Ixr1 protein was prepared at 1 mg mL^{-1} and dialyzed into 50 mM Tris-HCl, pH 8, 150 mM NaCl, 0,5 mM β-mercaptoethanol and 0,2 mM EDTA (Proteinase K digestion buffer). Proteinase K (Sigma-Aldrich) was dissolved at 1 mg mL^{-1} in the same digestion buffer and diluted 1:500 and 1:1000 (w/w). 10 µL of each proteinase K dilution was incubated with 10 µL of Ixr1 protein for 5, 10, 30 and 60 minutes on ice. The reactions were quenched with 4 x Laemmli SDS loading buffer, and the samples were boiled for 10 minutes and loaded onto a 12% SDS-PAGE gel and electrophoresed. Protein bands were visualized by staining the gel with Coomassie Brilliant Blue (Bio-Rad) and images were recorder in a Gel Doc^{TM} XR+ system (Bio-Rad).

Gel bands were manually excised and transferred to microcentrifuge tubes. Samples selected for analysis were in-gel reduced, alkylated and digested with trypsin according to the method of Sechi and Chait (Sechi & Chait, 1998). The samples were analyzed using the MALDI (Matrix-assisted laser desorption/ionization)-TOF (Time of Flight)/TOF mass spectrometer 4800 Proteomics Analyzer (AB Sciex, Framingham, MA, USA) and 4000 Series ExplorerTM Software (AB Sciex). Data Explorer version 4.2 (AB Sciex) was used for spectra analyses and for generating peak-picking lists. All mass spectra were internally calibrated using autoproteolytic trypsin fragments and externally calibrated using a

standard peptide mixture (Sigma-Aldrich). TOF/TOF fragmentation spectra were acquired by selecting the 10 most abundant ions of each MALDI-TOF peptide mass map (excluding trypsin autolytic peptides and other known background ions).

2.9.- One-hybrid system

The construction pGBD-C$_2$_IXR1_complete obtained was transformed in the *S. cerevisiae* PJ69-4A (Johnson *et al.*, 1996), which contains the *HIS3*, *ADE2* and *LacZ* reporter genes fused to *GAL1*, *GAL2* and *GAL7* promoter regions, respectively. *HIS3* and *ADE2* reporter systems were analyzed semi-quantitatively by incubation on CM-Trp-His or CM-Trp-Ade, respectively. β-galactosidase activity measurements were performed according to the method of Guarente (Guarente, 1983) modified for implementation in eppendorf tubes and following the release of the colored product o-nitrophenol (ONP) from the synthetic substrate o-nitrophenyl-β-D-galactopyranoside (ONPG) (Sigma-Aldrich). The activity was measured in cell-free extracts from 10 mL of culture after resuspending the cells in buffer Z (100 mM Na$_2$HPO$_4$, 40 mM NaH$_2$PO$_4$, 10 mM KCl, 1.6 mM MgSO$_4$ and 2.7 ml of β-mercaptoethanol per liter of solution, pH 7). Mechanic lysis was performed with glass beads (Sigma-Aldrich) and 100 μL of sample were diluted into 400 μL of buffer Z pre-incubated at 30 °C for 5 minutes. Then 250 μL of substrate solution (ONPG 4 mg mL^{-1} dissolved in distilled water) were added. The reaction was stopped by adding 250 μL of a Na$_2$CO$_3$ 1 M solution and the mixture was centrifuged at 17000 x g for 5 minutes and the O-nitrophenol released into the supernatant was determined spectrophotometrically by absorbance at 420 nm. β-galactosidase activity is expressed in nanomoles ONPG min^{-1} mg^{-1} of total protein in the extract. A molar extinction coefficient of 4500 M^{-1} cm^{-1} was used in the calculation (Inchaurrondo *et al.*, 1994).

2.10.- Differential Scanning Fluorometry

Differential Scanning Fluorometry (DSF) was used in order to analyze the thermal stability of the full-length protein and the HMG-box domains in tandem, and compare the results of melting temperature with the data obtained by circular dichroism (Ericcson *et al*., 2006; Niesen *et al*., 2007). A pre-screening of dye (Sypro Orange, Sigma-Aldrich) and protein was done in order to find the best dye and protein concentrations. This was found to be 20x Sypro Orange for 30 µM full-length protein and 20x Sypro Orange for 50 µM HMG-box domains protein. 96-well thin-wall PCR plates (Thermo Scienctific) sealed with Optical-Quality Sealing Tapes (Bio-Rad) were used. Samples were incubated for 5 minutes at 25 ^{o}C and then heated to 95 ^{o}C in increments of 0.5 ^{o}C x min^{-1} using an iCyclerIQ real-time PCR machine (BioRad). Fluorescence of the dye was monitored simultaneously using λ_{ex}: 490 nm and λ_{em}: 530 nm filtered wavelengths. Fluorescence intensities were plotted as a function of temperature and a non-linear Boltzmann fit was performed using GraphPad Prism 6.0 (GraphPad software). Melting temperature was calculated as the maximum of the first derivative of this curve (Niesen *et al*., 2007).

2.11.- Thioflavin T fluorescence

For reactions monitoring the rate of amyloid formation, purified proteins dissolved in resuspension buffer were heated for 5 min at 95 ^{o}C before being diluted to 20 µM in assembly buffer (5 mM K_2HPO_4, pH 6.6, 150 mM NaCl, 5 mM EDTA, 2 mM TCEP) plus 20 µM ThT (freshly prepared and added), in 96-well black with clear flat bottom polystyrene microplates (Corning) with 150 µl per well. Fluorescence measurements (λ_{ex}: 450 nm; λ_{em}: 482 nm) were made with a Multi-modal Synergy^{TM} H1 plate reader (Biotek$^{®}$). *De novo* and seeded amyloid assembly reactions monitored by ThT (Thioflavin T, Sigma-Aldrich) fluorescence were incubated at 25 ^{o}C and double orbital continuous shaken and readings were recorded every 2 min for a period 16 h. To generate amyloid seeds for comparisons

of fiber elongation rates, proteins were assembled using continuous end-over-end agitation for 5 days. Aggregates were then collected by centrifugation at 13000 rpm for 1 hour and they were re-suspended in fresh assembly buffer. To fragment the amyloid fibrils, samples were sonicated in a Vibra-cell VCX130 sonicator (Sonics®) at 70% amplitude during 3 minutes (10" on/10" off cycles) on ice.

After the final fluorescence time point, filter retardation assays (Scherzinger *et al*., 1999) were done to detect protein aggregates, which, due to their size, are specifically retained on the membrane surface. Samples were filtered through Nanosep 300K Omega ultracentrifuge columns (Pall Corporation) and a volume of 10 µL of the flow-through and the fraction retained was spotted on a PROTRAN® nitrocellulose membrane (Whatman®) and revealed with Ponceau S (Sigma-Aldrich) to detect immobilized proteins. In the same way, aggregated proteins were treated with non-denaturing tween-20 0,1% (v/v) or denaturing SDS 2% (w/v) during 30 minutes at room temperature and passed through a Nanosep 300K Omega ultracentrifuge columns (Pall Corporation) and a volume of 10 µL of the flow-through and the fraction retained (10 µL) was spotted on the same nitrocellulose membrane.

2.12.- Transmission Electron Microscopy

Different PrDs of Ixr1 protein were diluted to 20 µM in assembly buffer (5 mM K_2HPO_4, pH 6.6, 150 mM NaCl, 5 mM EDTA, 2 mM TCEP) and incubated one week at 25 °C and continuous rotation. Aggregates were then collected by centrifugation at 13000 rpm for 1 hour and re-suspended in fresh assembly buffer. 5 µL sample was incubated on top of a carbon-coated grid for 5 minutes. Excess solution was removed and the grid was washed in two sequential drops of ultrapure water (Sigma-Aldrich) and subsequently stained in a 2% uranyl acetate (w/v) or 2% phosphotungstic acid (w/v) (Sigma-Aldrich) drop (five minutes and thirty seconds of incubation time, respectively), followed by blotting off excess

solution. Samples were examined using a Jeol JEM-1010 transmission electron microscope an accelerating voltage of 100 kV, recorded at different magnifications and images were digitalized using a CCD camera.

2.13.- Fluorescent microscopy

The different PrDs of Ixr1 protein fused to the Enhanced Green Fluorescent Protein (EGFP) were used to investigate the amyloid propensities of cPrDs *in vivo*. The different pAG424GAL-PrD-EGFP constructions were transformed into the *S. cerevisiae* YJW509 and YJW584 strains, and selected in CM-leu plates. The pAG426GAL_SUP35N_EGFP construct was used as a positive control of protein aggregation. Cells were grown in galactose or galactose/raffinose-containing medium and subjected to fluorescence microscopy after 24, 48 and 72 hours of growth, using a Nikon Eclipse 50i microscope (Nikon®) and GFPHQ filter (λ_{ex}: 455-485 nm; dichroic mirror: 495; λ_{em}: 500-545 nm). Fluorescence images were acquired with a Nikon Digital Sight D5-5M camera (Nikon®) and processed using NIS-Elements D v2.20 (Nikon®) image analysis software.

2.14.- Semi-denaturing detergent-agarose gel electrophoresis (SDD-PAGE)

Aliquots of YJW584 cell cultures overexpressing the different PrDs fused to EGFP were taken at 24, 48 and 72 hours and harvested by centrifugation and resuspended in buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 2.5 mM EDTA, 1 % (v/v) Triton X-100, 1 mM PMSF and 1 x Complete Protease Inhibitor (Roche)). Cells were lysed using glass beads (Sigma-Aldrich) and were briefly spun at 3000 x g to sediment cell debris. The protein concentrations (determined by the Bradford reagent from Bio-Rad) of the cell lysate were adjusted and mixed with 4 x sample buffer (2 x TAE, 20 % (v/v) glycerol, 4 % (w/v) SDS and bromophenol blue). The samples were incubated at room temperature for 15 minutes and loaded onto a 1.8 % (1.5 mm thickness) agarose gel containing 1 x TAE and 0.1 % SDS. The gel was run in running buffer (1 x TAE, 0.1 % SDS) at 50 V, followed by blotting onto a

PROTRAN® nitrocellulose membrane (Whatman®). Anti-eGFP mouse IgG Antibody (MA1-952; Thermo Scientific) and goat anti-mouse IgG-HRP (sc-2005; Santa Cruz Biotechnology) primary and secondary antibodies respectively, and BM Chemiluminiscence Western Blotting kit -Mouse/Rabbit (Roche®) were used to detect the different PrD-fusion proteins in Amersham HyperfilmTM ECL High performance chemiluminiscence films (GE Healthcare®). Membrane washes, incubation times and developing procedures were done following manufacturer's instructions (BM Chemiluminiscence Western Blotting kit -Mouse/Rabbit (Roche®).

2.15.- SUP35 prion assay

A Sup35C tagging plasmid (pAG415ADH-ccdB-SUP35C) was used to generate different PrD-Sup35C expressing strains. After the strain YSR100 was transformed with the *cPrD-SUP35C* expression plasmid, plasmid shuffling was performed by plating the transformants on 5-FOA (5-Fluorotic acid). This produced a strain that expressed a particular PrD-Sup35C fusion protein as the only source of functional Sup35p. All strains were examined by Western blotting and SDD-AGE to evaluate expression and aggregation levels of the fusion proteins (data not shown). To induce the prion state, the PrD-SUP35C strains were transformed with a corresponding pAG424GAL-PrD-EGFP expression plasmid. Transformants were grown in SRafGal-Trp medium for 24 hours and then plated on YPD and SD-Ade plates with a cell number of 500 and 50000 per plate, respectively. The same strains grown in SRaf-Trp served as a control. The appearance of colonies in SD-Ade plates under inducing conditions suggested that expression of cPrD-EGFP induced a prion switch. In these cases Ade[+] colonies were re-streaked on YPD+GdnHCl plates to check prion curing.

2.17.- Bioinformatic resources

The Uniprot (http://www.uniprot.org) accession numbers of proteins used are P33417 (Ixr1 protein from *Saccharomyces cerevisiae*) and P05453 (Sup35

protein from *Saccharomyces cerevisiae*). Secondary structure predictions were done with JPred4 (http://www.compbio.dundee.ac.uk/jpred/) (Drozdetskiy *et al.*, 2015) and PSIPRED v3.3 (http://bioinf.cs.ucl.ac.uk/psipred/) (Jones, 1999; Buchan *et al.*, 2013). CH-plots and PONDR plots were obtained from http://www.pondr.com/cgi-bin/PONDR/pondr.cgi (Romero *et al.*, 1997; Uversky *et al.*, 2000; Romero *et al.*, 2001; Obradovic *et al.*, 2003; Obradovic *et al.*, 2005; Xue *et al.*, 2010). Hydrophobic cluster analysis (HCA plot) (Callebaut *et al.*, 1997) was made in the MeDor suite (Lieutaud *et al.*, 2008). IUPred (http://iupred.enzim.hu) (Dosztanyi *et al.*, 2005a; Dosztanyi *et al.*, 2005b), GlobPlot2 v2.3 (http://globplot.embl.de) (Linding *et al.*, 2002) and FoldIndex (http://bip.weizmann.ac.il/fldbin/findex) (Prilusky *et al.*, 2005) are included in the DisProt database. MoRFs were predicted with α-MoREs (http://biomine-ws.ece.ualberta.ca/MoRFpred/index.html) (Disfani *et al.*, 2012) and ANCHOR (http://anchor.enzim.hu/Theory. php) (Mészáros *et al.*, 2009; Dosztányi *et al.*, 2009) algorithms. Prion domain predictions were obtained from PAPA (http://combi.cs.colostate.edu/supplements/papa/) (Toombs *et al.*, 2012) algorithm.

3.- RESULTS

3.1.- Bioinformatic analyses show several prone-disordered regions in Ixr1 protein

The sequence of Ixr1 is unrelated to other known proteins, except for their two HMG-box DNA binding domains displayed in tandem, which are conserved between Ixr1 protein and other HMG-box proteins. Secondary structure prediction reveals unfolded regions flanking HMG-box domains, with three long poly-glutamine stretches of low complexity (figure 1). In general, repetitiveness and limited diversity in amino acid composition, low predicted secondary structure

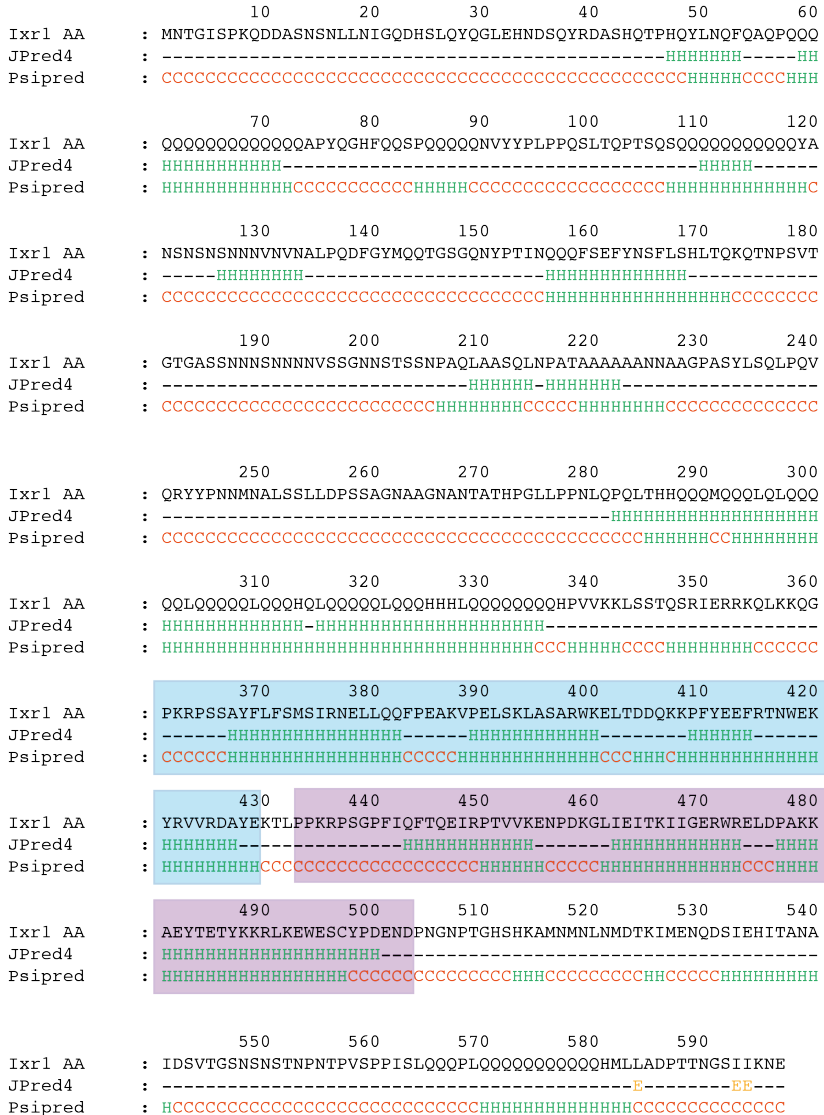content, and high sequence variability among close-related species is associated with flexibility of proteins.

```
                    10        20        30        40        50        60
Ixr1 AA   : MNTGISPKQDDASNSNLLNIGQDHSLQYQGLEHNDSQYRDASHQTPHQYLNQFQAQPQQQ
JPred4    : --------------------------------------------HHHHHH-----HH
Psipred   : CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHCCCCHHH

                    70        80        90       100       110       120
Ixr1 AA   : QQQQQQQQQQQQQQAPYQGHFQQSPQQQQQNVYYPLPPQSLTQPTSQSQQQQQQQQQQQYA
JPred4    : HHHHHHHHHH---------------------------------------HHHHH------
Psipred   : HHHHHHHHHHHCCCCCCCCCCHHHHHCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHC

                   130       140       150       160       170       180
Ixr1 AA   : NSNSNSNNNVNVNALPQDFGYMQQTGSGQNYPTINQQQFSEFYNSFLSHLTQKQTNPSVT
JPred4    : -----HHHHHHHH---------------------HHHHHHHHHHHHH------------
Psipred   : CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHCCCCCCCC

                   190       200       210       220       230       240
Ixr1 AA   : GTGASSNNNSNNNNVSSGNNSTSSNPAQLAASQLNPATAAAAAANNAAGPASYLSQLPQV
JPred4    : --------------------------HHHHHH-HHHHHH------------------
Psipred   : CCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHCCCCCHHHHHHHHHCCCCCCCCCCCCCCC


                   250       260       270       280       290       300
Ixr1 AA   : QRYYPNNMNALSSLLDPSSAGNAAGNANTATHPGLLPPNLQPQLTHHQQQMQQQLQLQQQ
JPred4    : ----------------------------------------HHHHHHHHHHHHHHHHHHHH
Psipred   : CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHCCHHHHHHHH

                   310       320       330       340       350       360
Ixr1 AA   : QQLQQQQQLQQQHQLQQQQQLQQQHHHLQQQQQQQQHPVVKKLSSTQSRIERRKQLKKQG
JPred4    : HHHHHHHHHHHHHH-HHHHHHHHHHHHHHHHHHHHH---------------------
Psipred   : HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCHHHHHCCCCHHHHHHHHCCCCCC
```

```
                   370       380       390       400       410       420
Ixr1 AA   : PKRPSSAYFLFSMSIRNELLQQFPEAKVPELSKLASARWKELTDDQKKPFYEEFRTNWEK
JPred4    : ------HHHHHHHHHHHHHHHH------HHHHHHHHHHH--------HHHHH------
Psipred   : CCCCCCHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHHHCCCHHHCHHHHHHHHHHHH
```

```
                   430       440       450       460       470       480
Ixr1 AA   : YRVVRDAYEKTLPPKRPSGPFIQFTQEIRPTVVKENPDKGLIEITKIIGERWRELDPAKK
JPred4    : HHHHHHH---------------HHHHHHHHHH-------HHHHHHHHHHH---HHHH
Psipred   : HHHHHHHHCCCCCCCCCCCCCCCCCCHHHHHCCCCHHHHHHHHHHHHHCCCHHHH
```

```
                   490       500       510       520       530       540
Ixr1 AA   : AEYTETYKKRLKEWESCYPDENDPNGNPTGHSHKAMNMNLNMDTKIMENQDSIEHITANA
JPred4    : HHHHHHHHHHHHHHHHHHH--------------------------------------
Psipred   : HHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCHHHCCCCCCCCCCHHCCCCCHHHHHHHH
```

```
                   550       560       570       580       590
Ixr1 AA   : IDSVTGSNSNSTNPNTPVSPPISLQQQPLQQQQQQQQQQQHMLLADPTTNGSIIKNE
JPred4    : ----------------------------------------E--------EE---
Psipred   : HCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHCCCCCCCCCCCCCC
```

Figure 1. Secondary structure prediction of Ixr1 protein (Uniprot accession number P33417) using JPred4 (Drozdetskiy *et al.*, 2015) and PSIPRED v3.3 (Jones, 1999; Buchan *et al.*, 2013). H (green): helix; E (orange): sheet; C (red): disordered. HMG-box domain A boxed in blue and HMG-box domain B boxed in purple.

Several tools for disorder prediction were used to analyse the possibility that Ixr1 could contain IDRs. The reliability of disorder prediction comes from the use of several methods based on different theoretical approximations, different physicochemical parameters, or different applications. Therefore, to avoid pitfalls, different predictors were combined. To check disorder propensity based on physical/chemical features of amino acids, FoldIndex (Prilusky *et al.*, 2005), GlobPlot 2 (Linding *et al.*, 2002) or HCA plot (Callebaut *et al.*, 1997) were used, integrated in the local server 'MeDor' (MEtaserver of DisORder) (Lieutaud *et al.*, 2008). PONDR-FIT (Xue *et al.*, 2010), DISOPRED2 (Ward *et al.*, 2004) and DisEMBL (Linding *et al.*, 2007) are machine-learning algorithms trained on data sets of disordered regions deposited in different databases, such as the Database of Protein Disorder (DisProt) (Sickmeier *et al.*, 2007), IDEAL database (Fukuchi *et al.*, 2012) and MobiDB (Di Domenico *et al.*, 2012). Finally, IUPred (Dosztanyi *et al.*, 2005a; Dosztanyi *et al.*, 2005b) was included to perform a better prediction of long disordered segments, avoiding the shortcomings and biases associated with disordered data sets by statistical evaluation of the energy, which results from inter-residue interactions to overcome the large decrease in configurational entropy during folding.

Figure 2 shows the results obtained for the full-length Ixr1 protein looking for characteristics of IDPs. Compositional profiling comparison of the Ixr1 protein with a set of known globular folded proteins from the Protein Data Bank (PDB) (Radivojac *et al.*, 2007) shows that Ixr1 is depleted in major order-promoting amino acids (V, I, L, F, W, Y) and is enriched in some disorder-promoting residues, particularly Q, S and P. Typically, intrinsically disordered proteins have relatively large segments of polar amino acids (Uversky *et al.*, 2000). A plot of mean residue charge against hydropathy indicates that Ixr1 clearly locates in the disordered protein cluster (figure 2a), with low values of both parameters. To a better understanding of charge and hydropathic distributions, a hydrophobic cluster analysis (HCA) was made. The distribution of hydrophobic amino acids along the

sequences of globular domains is not at random. Oppositely, specific hydrophobic clusters have been statistically associated to the internal faces of regular secondary structures and reveal signatures tightly associated with the major folds selected in nature. HCA is a two-dimensional helical representation of protein sequences focused on the residues present in the hydrophobic core of proteins that serves as a powerful tool to investigate the basis of protein stability and folding. Figure 2b shows that Ixr1 protein has extended regions, which are depleted of hydrophobic clusters and charged amino acids, mainly in the amino terminal side. On the contrary, horizontal clusters contained in the second half of the protein correspond well with the alpha helices of the HMG-box DNA binding domains. Other disorder predictors based on amino acid characteristics, such as GlobPlot 2 or FoldIndex, also show the high content of IDRs present in the Ixr1 protein (figure 2c). In the same way, several predictors based on neural networks such as IUPred (figure 2c), DISOPRED2 or the different variants of PONDR (figure 2d) also indicate that Ixr1 protein tends to be disordered in approximately 50% of its amino acid sequence; the highest degree of structural order fits in the HMG-box region, while stretches in the first half of the protein, and even in the last one hundred amino acids of the C-terminal side of Ixr1 protein, show strong disorder propensity. Predicted disorder values obtained with PONDR VL3 were lower than with the other methods, which could be attributed to the fact that it was trained to IDPs with the highest percentage of disorder.

3.2.- Ixr1 is non-globular and aggregation-prone in solution

IDPs may exist in at least three separate functional conformations: molten globule (MG), premolten globule (PMG), and random-coil-like (RC-like) (Uversky *et al*., 2002). In the MG state, disorder regions can collapse to native secondary structures, although the protein molecule lacks a well-packed core, attaining some topology and a compact state close to that of ordered globular proteins. On the contrary, the RC-like state has hydrodynamic dimensions typical of proteins

unfolded without any ordered secondary structure, meanwhile the PMG state is an intermediate state with "squeezed" and partially ordered form of the coil with some residual secondary structure.
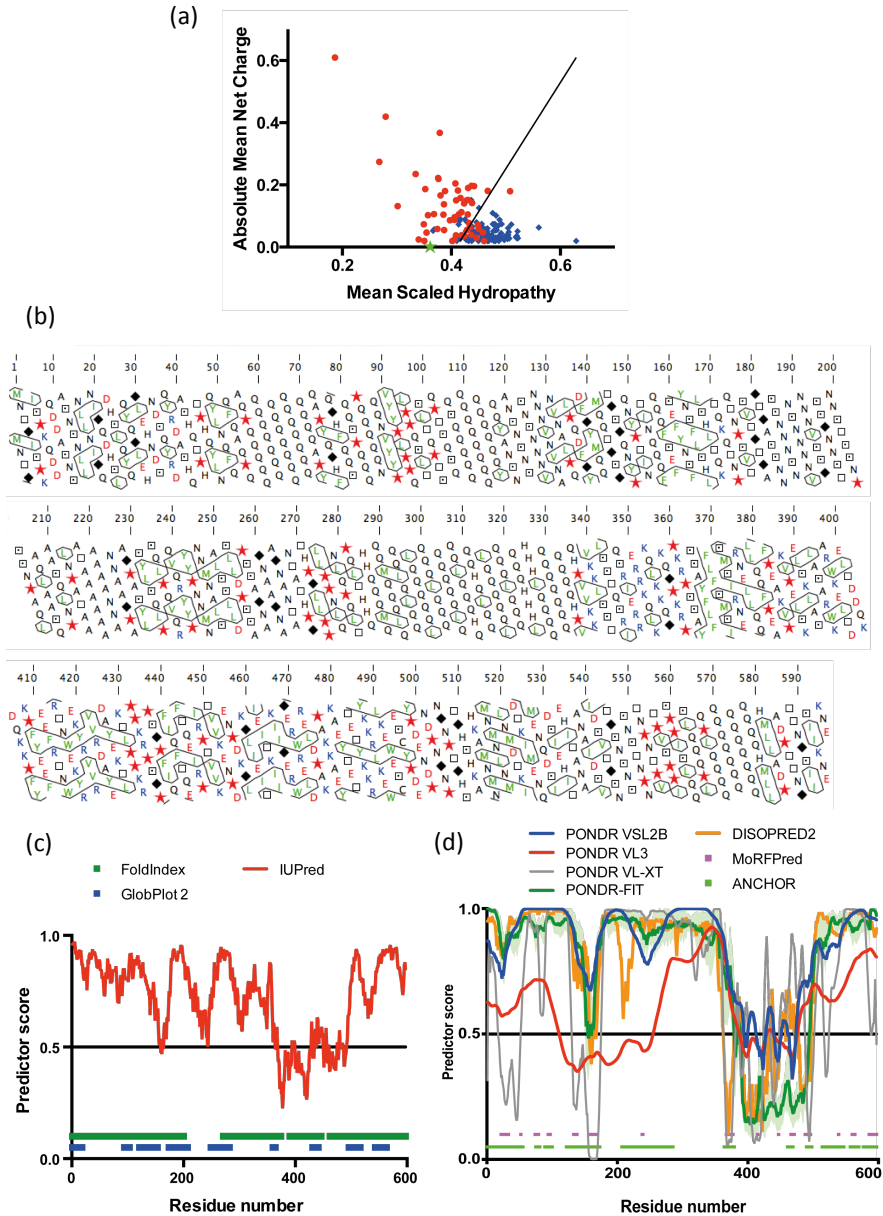
←**Figure 2.** (a) Charge-hydropathy plot (CH-plot) of Ixr1 protein (green star). Net charge versus mean hydrophobicity is plotted for disordered (red circles) and ordered (blue rhombus) proteins. A black line separates the two sets, which corresponds with the boundary equation $|<R>| = 2.785<H> - 1.151$ (being $|<R>|$ mean net charge and $<H>$ mean hydropathy) (Uversky *et al.*, 2000). (b) MeDor graphical output of the full-length Ixr1 protein (Lieutaud *et al.*, 2008). Hydrophobic cluster analysis (HCA plot) (Callebaut *et al.*, 1997) indicates enrichment in hydrophobic clusters (regions highlighted). Prolines are indicated as a red star (which confers the greatest constraint to the polypeptide chain); glycines are indicated as a black diamond (which confers the largest freedom to the chain); serines are indicated as squares with a point and threonines as squares without a point (both small polar amino acids can mask their polarity through H-bonding with the carbonyls of the main chain, particularly within helices). (c) Evaluation of intrinsic disorder in the Ixr1 protein by predictors suitable for long disordered protein segments. Three disorder prediction tools were used: IUPred (red line) (Dosztanyi *et al.*, 2005a; Dosztanyi *et al.*, 2005b), which evaluates the energy resulting from interresidue contacts in globular proteins; GlobPlot2 v2.3 (blue bars), based on a running sum of the propensity for amino acids to be in an ordered or disordered state (Linding *et al.*, 2002); and FoldIndex (green bars) (Prilusky *et al.*, 2005), based on the boundary equation of Uversky and co-workers. (d) Evaluation of intrinsic disorder in the Ixr1 protein by predictors suitable for short disordered protein segments. Four disorder prediction tools of the PONDR family were used: PONDR VSL2B (Obradovic *et al.*, 2005), which is more appropriate for proteins with disordered and structured regions); PONDR VLXT (grey) (Romero *et al.*, 1997; Romero *et al.*, 2001), useful for predicting MoREs (Molecular Recognition Elements); PONDR VL3 (red) (Obradovic *et al.*, 2003), suitable for large or fully disordered proteins; and PONDR-FIT (green) (Xue *et al.*, 2010), a meta-predictor statistically better than VSL2 when both structure and disorder are present. Light green shadows represent standard errors of disorder (*Figure 2 continued*) prediction by PONDR-FIT. Locations of α-MoREs (Disfani *et al.*, 2012) and ANCHOR-indicated binding sites (AiBSs) (Mészáros *et al.*, 2009; Dosztányi *et al.*, 2009) are shown as purple and green bars at the bottom of plot.

To investigate the conformational properties of Ixr1 protein and its oligomeric state in solution, we expressed and purified recombinant full-length Ixr1 protein (1-597) and truncated variants containing either its N-terminal region preceding the HMG-box domains (1-336), or the HMG-box domains in tandem (338-510) (figure 3a). Both full-length protein and the N-terminal truncated version showed low solubility and high tendency to aggregate at final steps during protein concentration (maximum up to ≈75 μM in case of full-length protein). Gel filtration analysis using size exclusion chromatography also showed that the hydrodynamic volume of the full-length Ixr1 protein was 1.51 fold higher than the expected for

the monomeric form, taking into account their Stoke's radii using a set pattern of globular proteins (Uversky *et al.*, 1993; Samrajnee *et al.*, 2001; Uversky *et al.*, 2002; Uversky *et al.*, 2010; Uversky *et al.*, 2012) (figure 3b). This discrepancy between conformational behaviours into the SEC matrix could be related to a more extended shape, which hydrodynamic volume increment corresponds to the expected for a MG state (Feldman and Hogue, 2000; Bernado and Svergun, 2012).

To further investigate the oligomeric state of Ixr1 in solution, analytical ultracentrifugation analysis was made (AUC). Sedimentation equilibrium (SE-AUC) experiments clearly indicated that the monomer is a predominant form of Ixr1 protein in solution at a concentration of 0,8 mg/mL (corresponding to ≈15 µM) (figure 3c). Accordingly to this data, sedimentation velocity (SV-AUC) experiments showed that full-length protein presents a frictional coefficient ratio (f/f_0) of 1.97, calculated from the molecular mass and the 2.9 c(S) sedimentation coefficient (figure 3d). This large frictional coefficient ratio could result from either a rod-like structure with a high axial ratio or a coil-like structure such as that of a denatured protein (McBryant *et al.*, 2006).

Protein oligomerization is a concentration-dependent event that increases logarithmically with increasing concentrations (Shriver *et al.*, 2009). Accordingly, many proteins were reported to maintain a dynamic equilibrium between monomers and oligomers (Dengra-Pozo *et al.*, 2009; Gruber *et al.*, 2009; Lunn *et al.*, 2008). To check this possibility, the full-length Ixr1 protein was titrated from its highest protein concentration (≈75 µM to avoid aggregation) to 290 nM and resolved by Blue-native PAGE (Wittig *et al.*, 2006). This technique allows the separation of proteins in acrylamide gels according to their size, instead of to their charge/mass ratio. The addition of Coomassie blue G-250, an anionic dye, imposes a homogeneous charge on the proteins. Figure 3e shows that the Ixr1 protein varies its stoichiometry, depending on protein concentration, from a monomeric to a homodimeric state. Furthermore, protein treatment with 150 µM glutaraldehyde

to stabilize and fix weak protein-protein interactions showed that Ixr1 has high tendency to aggregate and form larger oligomeric states (figure 3f).
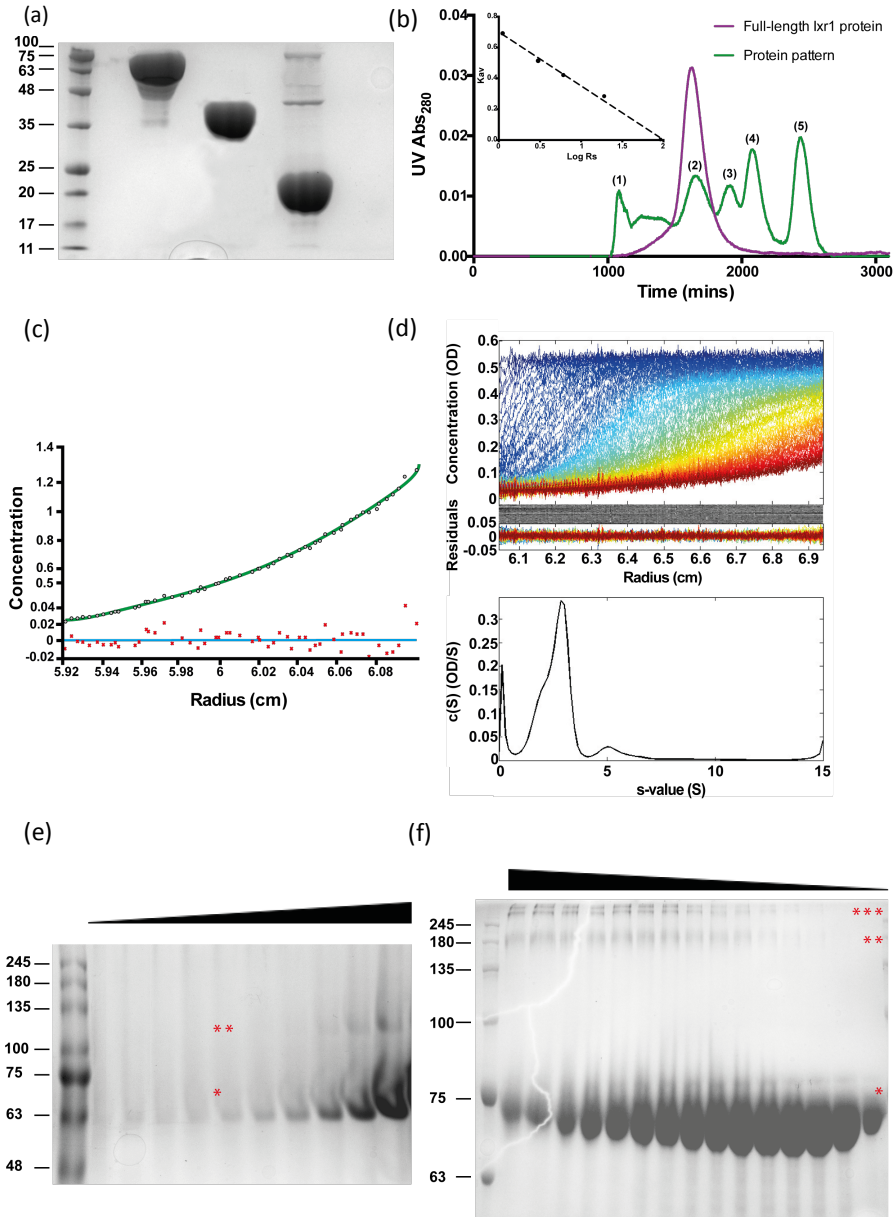

Figure 3. (a) SDS-PAGE (12% polyacrylamide gel) of full-length Ixr1 (lane 1), N-terminal region (lane 2) and tandem HMG-box domains (lane 3) purified proteins. (stained with Coomassie Blue G-250) (b)

(*Figure 3 continued*) Sepharose S200 16/70 size-exclusion chromatogram showing the elution times (mins[-1]) of full-length Ixr1 protein (purple) and a protein standard from Bio-Rad (green); (1): bovine thyroglobulin bovine (670 kDa; 78,8 Å); (2): bovine γ-globulin (158 kDa; 51 Å); (3): chicken ovalbumin (44 kDa; 31,2 Å); (4): equine myoglobin (17 kDa; 19 Å); (5): Vitamin B_{12} (1,35 kDa). Inset plot of Kav against $Log_{10}R_s$ and linear regression (K_{av}= -0,527$Log_{10}R_s$ + 1.1935). (c) Sedimentation equilibrium data of purified full-length Ixr1 protein obtained (d) Sedimentation velocity data of purified full-length Ixr1 protein obtained by optical absorbance at 280 nm and fitted to the Lamm equation with SEDFIT program. The continuous sedimentation coefficient distribution of the protein is displayed below. (e) Blue-native (4-12%) PAGE of full-length Ixr1 protein, showing monomer (*) -homodimeric (**) equilibrium. Lanes 2-15 contains growing protein concentrations, ranging from 195 nM to 50 µM (f) SDS-PAGE of full-length Ixr1 protein after 150 µM glutaraldehyde treatment showing monomer (*) -homodimeric (**) -higher oligomeric states (***) equilibrium. Lanes 2-15 contains decreasing protein concentrations, ranging from 50 µM to 6 nM.

3.3.- Far-UV circular dichroism spectra of the N-terminal region of Ixr1 is characteristic of intrinsically disordered regions (IDRs)

Spectroscopic analysis using Circular Dichroism (CD) was made to evaluate the residual structure of full-length Ixr1 protein or their N-terminal and HMG-box domains truncated regions separately. In general, far-UV circular dichroism spectra of unfolded polypeptide chains are characterized by a prominent minimum ellipticity in the vicinity of 190 to 200 nm and values close to zero in the vicinity of 222 nm (Daughdrill *et al*., 2005; Receveur-Bréchot *et al*., 2006). The far-UV CD spectrum of full-length Ixr1 protein (figure 4a, black line), as well as its N-terminal truncated region (figure 4a, blue dotted line), revealed a typical profile for disordered proteins, meanwhile its HMG-box domains in tandem (figure 4a, green dashed line) showed predominately α-helical structure as indicated by the characteristic minima at 208 and 222 nm. Indeed, deconvolution analysis using different programs showed that HMG-box domains in tandem contains 67-71% of α-helix and 8-9% of disordered regions, meanwhile full-length Ixr1 protein only contains 19-26% of α-helix but 41-69% of disordered regions. Displaying the Mean Residue Ellipticity (MRE) at 200 and 222 nm in a double wavelength plot allowed discriminate between full-length Ixr1 protein as an extended MG or PMG state and

N-terminal truncated region as a RC-like state, as described by Uversky *et al.*, 2002, 2010b; Habchi *et al.*, 2014.

Further analyses to check the tendency of Ixr1 protein to form secondary structures were done. Trifluoroethanol (TFE) is widely used in order to stabilize regions with pre-existent alpha helical structure by lowering the dielectric constant of the solution (Buck *et al.*, 1998). When full-length Ixr1 protein was titrated with TFE, it folded into a helical conformation with increasing TFE concentrations (Figure 4b), diminishing negative ellipticity at 200 nm and growing negative values at 222 nm (figure 4b inset). This could indicate that, under some conditions, particular disordered tracts may fold, for example upon binding to an interaction partner or binding to a ligand. However, no changes were observed with the addition of a 20 pb DNA duplex coding for -400 to -380 promoter region of *ROX1* gene (figure 4a, red dotted-dashed line) (Castro-Prego *et al.*, 2010a). Other additives, including guanidine hydrochloride to check Polyproline type II (PII) structure (Shi *et al.*, 2002; Liu *et al.*, 2004) or submicellar concentrations of detergents such as sodium dodecyl sulfate (SDS) to check beta sheet transitions mixtures (Zhong and Johnson, 1992), did not induce significant conformational changes (data not shown).

Melting curves for full-length Ixr1 protein and their HMG-box domains were recorded at 222 nm (figure 4c). No large differences were observed between both proteins, with estimated melting temperatures of 48.35 ^{o}C for Ixr1p protein and 46.36 ^{o}C for HMG-box domains, indicating that thermal stability is provided only by the HMG-box domain folding and the rest of protein may be unstructured. Similar results were obtained from differential scanning fluorimetry, with estimated melting temperatures of 46.75 ^{o}C for Ixr1 protein and 43.75 ^{o}C for HMG-box tandem domain (figure 4d).
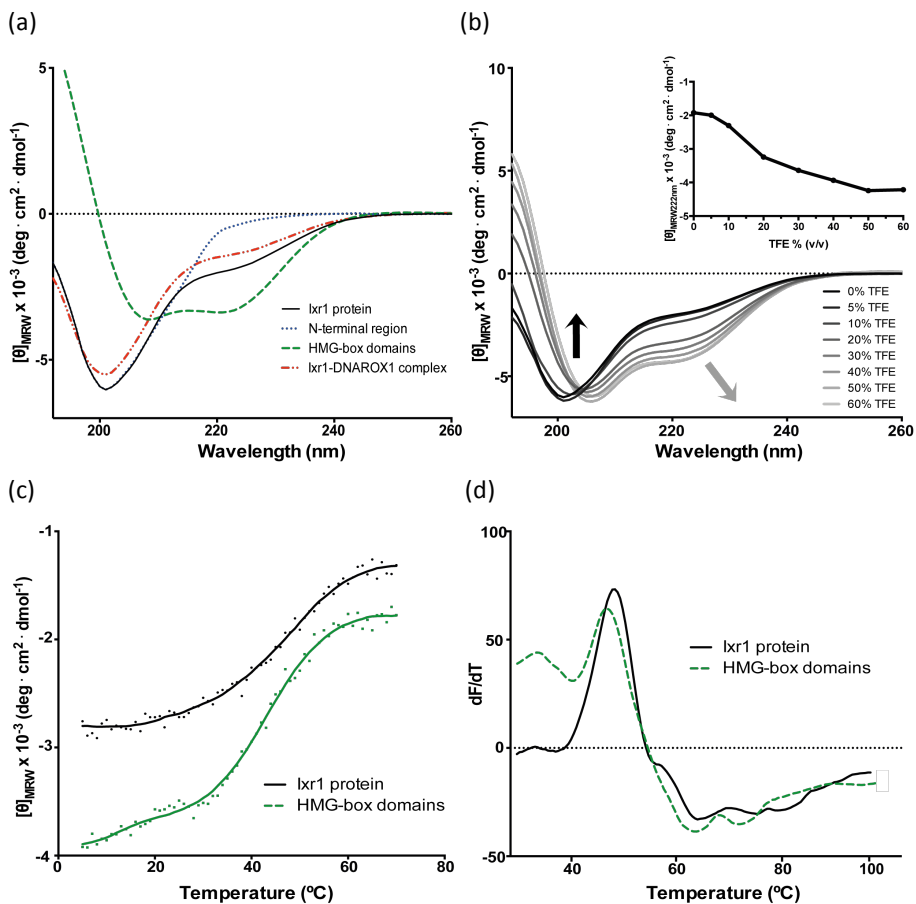
Figure 4. (a) CD spectroscopic analysis of Ixr1 protein. CD spectrum comparison of (Ixr1 protein (black solid line), Ixr1 N-terminal truncated region (blue dotted line), Ixr1 HMG-box domain (green dashed lines) and Ixr1-DNA^{ROX1} complex (red dotted-dashed line). (b) TFE titration into Ixr1 protein. The concentrations range from 0% to 60% (from black to grey), with intermediate concentration of 5, 10, 20, 30, 40 and 50%. (c) CD melting curve comparison of full-length Ixr1 protein with their HMG-box domains, starting from 5 ^{o}C and finishing at 75 ^{o}C. (d) Differential scanning fluorimetry of full-length Ixr1 protein and their HMG-box domains. Plot converted into first derivative of SyproOrange fluorescence as a function of time.

3.4.- The HMG-box domains are folded in the purified Ixr1 protein

Limited proteolysis experiments can be successfully used to probe conformational features of proteins. In general, each target sites of a determined

protease used in a limited proteolysis experiment require a large conformational change (local unfolding) of a stretch of up to 12 residues to produce a hydrolysis event (Hubbard *et al.*, 1994), allowing to discriminate between loops and disordered regions from properly folded domains, where protease action is not possible. This approach implies that the proteolysis of a protein is dictated by the stereochemistry and flexibility of the protein substrate (Fontana *et al.*, 2004). According to this, the most suitable proteases are those that display broad substrate specificity, such as proteinase K, thermolysin, pepsin or subtilisin (Fontana *et al.*, 1997).

There are 208 consensus proteinase K cleavage sites predicted throughout the full-length Ixr1 protein primary sequence, producing peptide-digested products of less than 3 kDa (Figure 5a). Thus, proteinase K accessibility is a sensitive probe for the global tertiary structure of this protein. Purified full-length Ixr1 protein was digested with increasing amounts of proteinase K, resolving the proteolytic products by SDS-PAGE (figure 5b). At 30 and 60 minutes of 1:500 Ixr1:proteinase K ratios, most of the high molecular weight bands were depleted, showing high proteolysis rates that indicates large extended regions of the protein (figure 5a, lanes 9 and 10). At the highest proteinase K digestions times employed, three prominent protease resistant bands were observed with apparent sizes of 25.6, 23.9 and 15.6 kDa (figure 5a, lanes 6 and 10). These proteinase K resistant bands were excised from the gel and subjected to peptide mass fingerprint to identify to which regions correspond along the polypeptide chain of Ixr1 protein. Resistant peptides were identified within the positions that correspond with HMG-box DNA binding domains of the Ixr1 protein (figure 5b).

By sequence homology, these domains are predicted to adopt characteristic L-shaped fold, containing three α-helices with an angle of ≈80° between the arms. The long arm or minor wing is composed by the extended N-terminal strand and helix III, while helix I and II form the short arm or major wing

(Ross *et al.*, 2005). This relative native rigid structure avoids proteinase K digestion, and effectively, the 75 proteinase K digestion sites identified within the HMG-box regions were not accessible to the enzyme in our experiments.
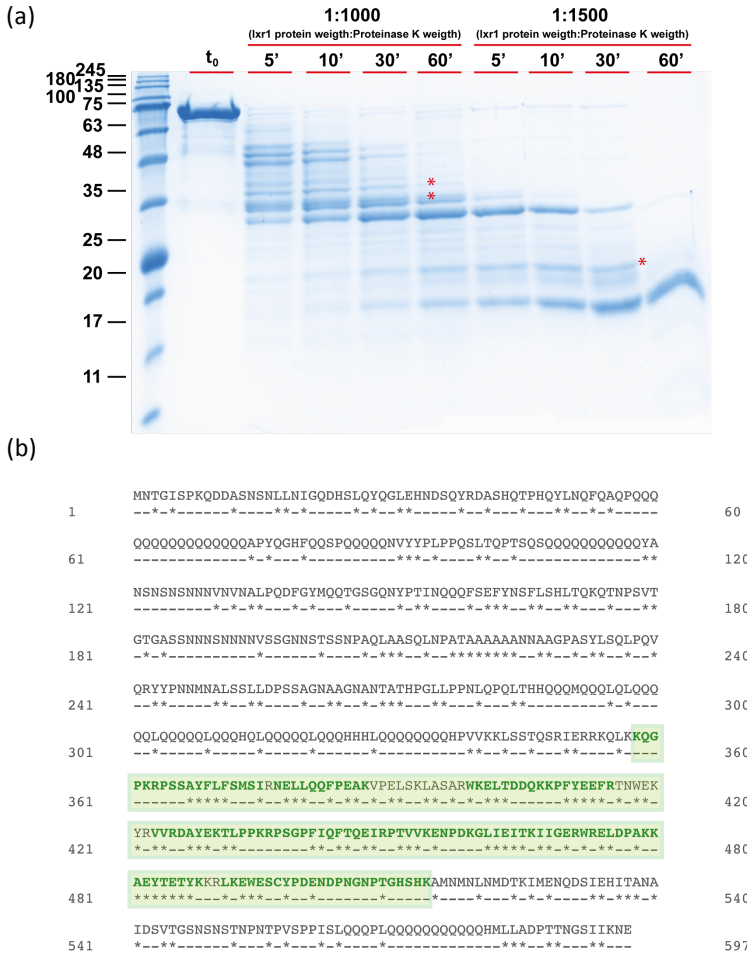

Figure 5. (a) Time-course analysis by SDS-PAGE (12% acrylamide gel) of the peptic digestion of full-length Ixr1 protein by two different proteinase K concentrations (1:1000 and 1:500 (w/w); see materials and methods). Size standards were loaded in lane 1. A red asterisk indicates digestion bands rescued and analysed by MALDITOF fingerprint. (b) Sequence scheme of Ixr1 protein (Uniprot code P33417) defining globular and extended regions based on peptides identified by MALDITOF fingerprint. Asterisks indicate possible cutting sites for proteinase K (obtained from Protparam server). Amino acids identified by proteomic analysis are indicated in bold green. Region expected to be folded are included into the green box.

3.5.- Ixr1 N-terminal polyglutamine domains may act as a trans-activation domain

Transcription factors (TFs) regulate the activation or repression of transcription, coupled with the recruitment and assembly of the transcription machinery, via the recognition of specific DNA sequences. This implies that both protein-DNA and protein-protein recognition play key roles in transcription factor function (Ward *et al.*, 2004). Their intrinsic plasticity enables IDRs to recognize and bind many biological targets with high specificity and low affinity (Dyson *et al.*, 2002; Dyson *et al.*, 2005), which altogether with their rapid turnover and reduced lifetime in the cell, make IDPs crucial in regulatory mechanism (Fink, 2005). Examples such as p53, GCN4, CBP, BRCA1 or HMGA proteins strongly support this concept (Dunker *et al.*, 2001; Dunker *et al.*, 2005; Uversky *et al.*, 2007).

In this sense, accumulated experimental evidence indicates that many trans-activating domains, which are flanking the DNA binding domains, have a significantly higher degree of disorder than the rest of the protein (Liu *et al.*, 2006). To check the possible activator function of the IDRs found in the Ixr1 protein, we measured the transcriptional activation strength of the IDRs fused to the DNA binding domain of Gal4 (Hirst *et al.*, 1999) using three different reporter genes (*HIS3, ADE2* and lacZ) whose promoters had been modified to be regulated by Gal4. Screening of protein fragments fused to the DNA binding domain of Gal4 has been used before to identify both prokaryotic and eukaryotic transcriptional activators based on a yeast system (Ruden *et al.* 1991; Sadowski *et al.*, 1992; Gstaiger & Schaffner, 1994: Escher & Schaffner, 1996).

PJ69-4a cells (James *et al.*, 1996) transformed with pGBD-C_2_IXR1_complete construction (containing the Ixr1 protein fused to the DNA binding domain of Gal4 (Gal4BD)) were selected in CM-His-Leu and in CM-His-Ade plates. Figure 6a shows that cells expressing Ixr1-Gal4BD chimera are able to grow in leucine selectable plates, indicating transcriptional activation capacities.

Nevertheless, growth in adenine plates was not observed. β-galactosidase enzyme assays were then performed to measure indirectly the level of transcriptional activation activity of Ixr1 (see Materials and methods), showing LacZ gene expression only when Gal4BD are fused to Ixr1 (figure 6b).
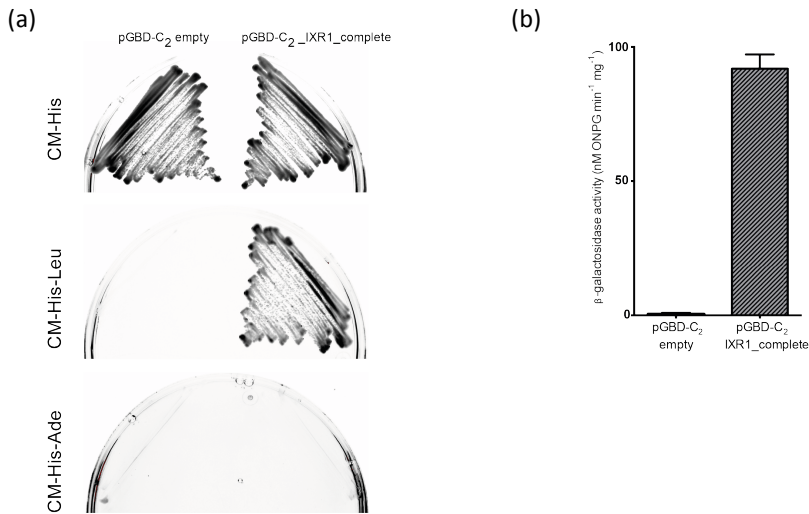

Figure 6. (a) PJ69-4a cells transformed with either pGBD-C$_2$ empty plasmid or with pGBD-C$_2$_IXR1_complete construction and incubated at 30 ^{o}C in CM-His (positive growth control), CM-His-Leu or CM-His-Ade plates. (b) β-galactosidase enzyme measurements of PJ69-4a cell extracts after growth in CM-His medium at 30 ^{o}C and continuous shaking, carrying either pGBD-C$_2$ empty plasmid or with pGBD-C$_2$_IXR1_complete construction.

Activation domains vary greatly in their sequence and consequently are difficult to predict studying homologies. Therefore, it remains an important challenge to map more precisely these domains in the large IDR here found as important for transcriptional activation mediated by Ixr1. An interaction-prone short segment of IDR that becomes ordered upon specific binding was described as 'molecular recognition element' or "MoRE" (Oldfield, 2002; Bourhis *et al*. 2004; Dunker *et al*., 2005). There are three basic types of MoREs: those that form α-helical structures, those that form β-strands (the protein partner has to provide additional b-sheet forming elements), and those that form irregular structures

upon binding. The names of α-MoRE, β-MoRE and I-MoRE respectively, were proposed (Uversky *et al*., 2005). As noted in figure 2b, PONDR VL-XT often identifies regions as short downward spikes flanked by regions predicted to be disordered (Garner *et al*., 1999). Two different algorithms, MoRFPred (Disfani *et al*., 2012) and ANCHOR (Mészáros *et al*., 2009; Dosztányi *et al*., 2009), were used to find disordered regions of the Ixr1 protein that can bind to specific protein or DNA partners, and potentially undergo the disorder-to-order transitions as a result of this binding. Figure 2e shows that Ixr1 protein regions that are flanking the HMG box domains could be possible sites to provide an interaction platform with various binding partners.

3.7.- Ixr1 are able to form amyloids *in vitro*

The presence of conformational flexible regions of IDPs, highly enriched for uncharged polar residues, can lead to the self-assembly into amyloids, one of the most highly ordered structures in biology. To assess the amyloidogenic properties of the Ixr1 protein, four different constructs were designed, corresponding to different regions of full-length Ixr1 protein (PrD1-4, figure 7a). Sup35 prion domain (Alberti *et al*., 2009) was purified as a positive control assay. Different PrDs cloned into pETNKI-his3C-LIC-Amp plasmid were expressed in bacteria and purified under denaturing conditions. After that, dilutions into physiological conditions to follow amyloid formation as a function of time were done in presence of Thioflavin T (ThT), a benzathiole dye that emits fluorescence upon amyloid binding without interfere in amyloid assembly (LeVine *et al*., 1997).

In general, despite differences in nucleated polymerization times and fibrillation rates among Ixr1 PrDs, all of them show positive amyloidogenic properties (figure 7, green points). PrD2 (figure 7d) and PrD4 (figure 7f) of Ixr1 protein show the highest aggregation rates, completing amyloid formation within the first 10 hours of the experiment time-course, similar to Sup35 PrD (figure 7b).

(a)

(b)

(c)

(d)

(e)

(f)

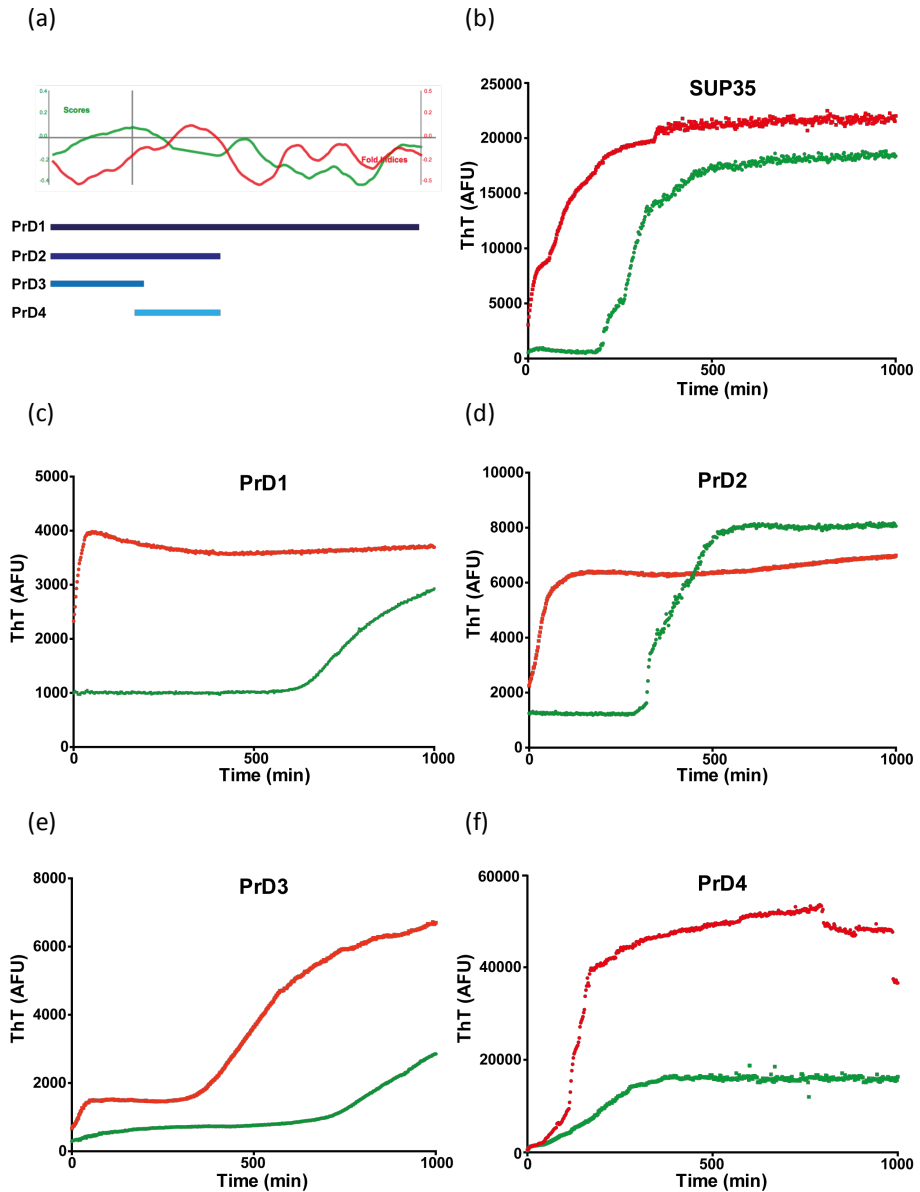**Figure 7.** (a) Scheme of the different Ixr1 PrDs purified based on PAPA algorithm (Toombs *et al.*, 2012). (b), (c), (d), (e) and (f) plots amyloid kinetics formation of Sup35 PrD, Ixr1 PrD1, Ixr1 PrD2, Ixr1 PrD3 and Ixr1 PrD4, respectively, monitored by ThT fluorescence during 16 hours at 30 ^{O}C and continuous shaking. Green squares correspond to no-seeded and red circles to seeded assays.

PrD1 (figure 7c) and PrD3 (figure 7e), despite being able to nucleate within the time-frame examined (16 hours), did not reach their maximum at final maturation plateau phase and show large initial lag phases.

To accelerate the kinetic rates of aggregation, seeds of each PrD were prepared and added to the mixture. PrD1 (figure 7c) and PrD2 (figure 7d) of Ixr1 protein changed dramatically their kinetic profile, as well as Sup35 PrD (figure 7b), without a lag phase and reaching the plateau phase very early.

Because of some non-amyloid β-sheet structures can alter ThT fluorescence (LeVine *et al.*, 1993), additional experiments were done. Samples from amyloid kinetic assays were treated with 2% SDS (denaturing conditions) or 0.1% Tween 20 (non-denaturing conditions) and passed through non-binding 300 kDa membrane spin column. Non-amyloids aggregates can then differentiated from amyloids because non-amyloid aggregates become solubilized in denaturing conditions and flow through. Figure 8 shows that all PrD aggregate formed are SDS resistant, although it was observed less resistance to denaturing conditions in the case of Ixr1 Prd1.

Transmission electron microscopy images of the different PrDs were also taken to confirm the fibrillar morphology of the protein aggregates obtained (Figure 9). The overall morphology of amyloid aggregates depends on the conditions in which fibrillogenesis takes place, and different fibril morphologies are often observed in the same preparation, with a variable number and arrangement of protofilaments.

3.6.- Different tracts of Ixr1 coalesce into foci aggregates in the cytosol

Most misfolded polypeptides have an inherent tendency to form self-templated amyloid structures (Chiti & Dobson, 2006). Outmosting, prion-forming proteins have an unusual conformational flexibility that allows access to the

amyloid fold even under physiological conditions (Alberti *et al.*, 2009; Uversky, 2009). This property derives in part from a greatly reduced amino acid complexity compared to globular proteins (Romero *et al.*, 2001; Alberti *et al.*, 2009).

To investigate the propensity of the Ixr1 protein to form foci in living yeast cells, the different Ixr1 PrDs were cloned into the pAG426GAL-ccdB-EGFP plasmid to overexpress them as chimeras with green fluorescent protein (PrD-EGFP), as well as Sup35 prion domain as positive control experiment (Sup35N-EGFP). YJW509 [pin-] and YJW584 [PIN+] transformed cells were incubated in SRafGal medium and samples were taken at 24, 48 and 72 hours, since amyloid formation is time and concentration dependent. In general, all the PrDs under study were capable to form foci with a punctate form and close to the vacuole (figure 10). However, PrD2 and PrD3 showed higher frequencies to form aggregates into the cell. A striking feature observed was the capability of PrD1 and PrD2 to induce cell flocculation when the protein construction was overexpressed, meanwhile PrD3 and PrD4 did not show this characteristic.
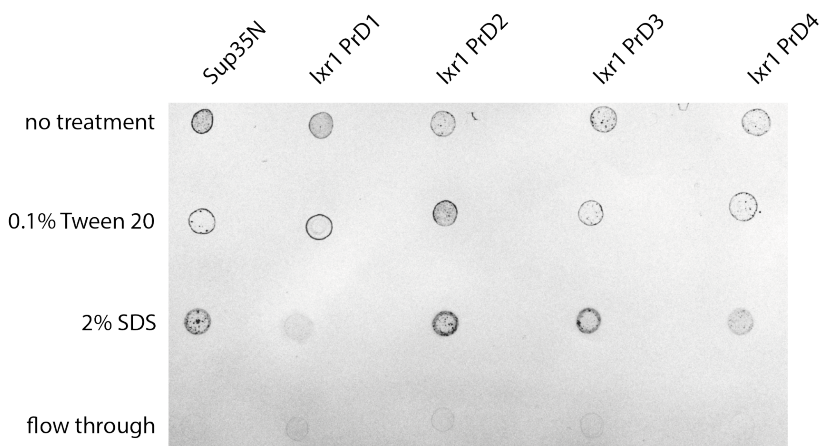
Figure 8. Dotblot of amyloid kinetic samples after the final measurement. Reactions were analyzed for detergent-resistant aggregation. Aliquots from ThT fluorescence experiments were spotted directly onto nitrocellulose ('no treatment'), or treated with either 0.1% Tween 20 or 2% SDS and filtered through a 300 kDa nonbinding membrane. Retained protein was visualized with Ponceau S.
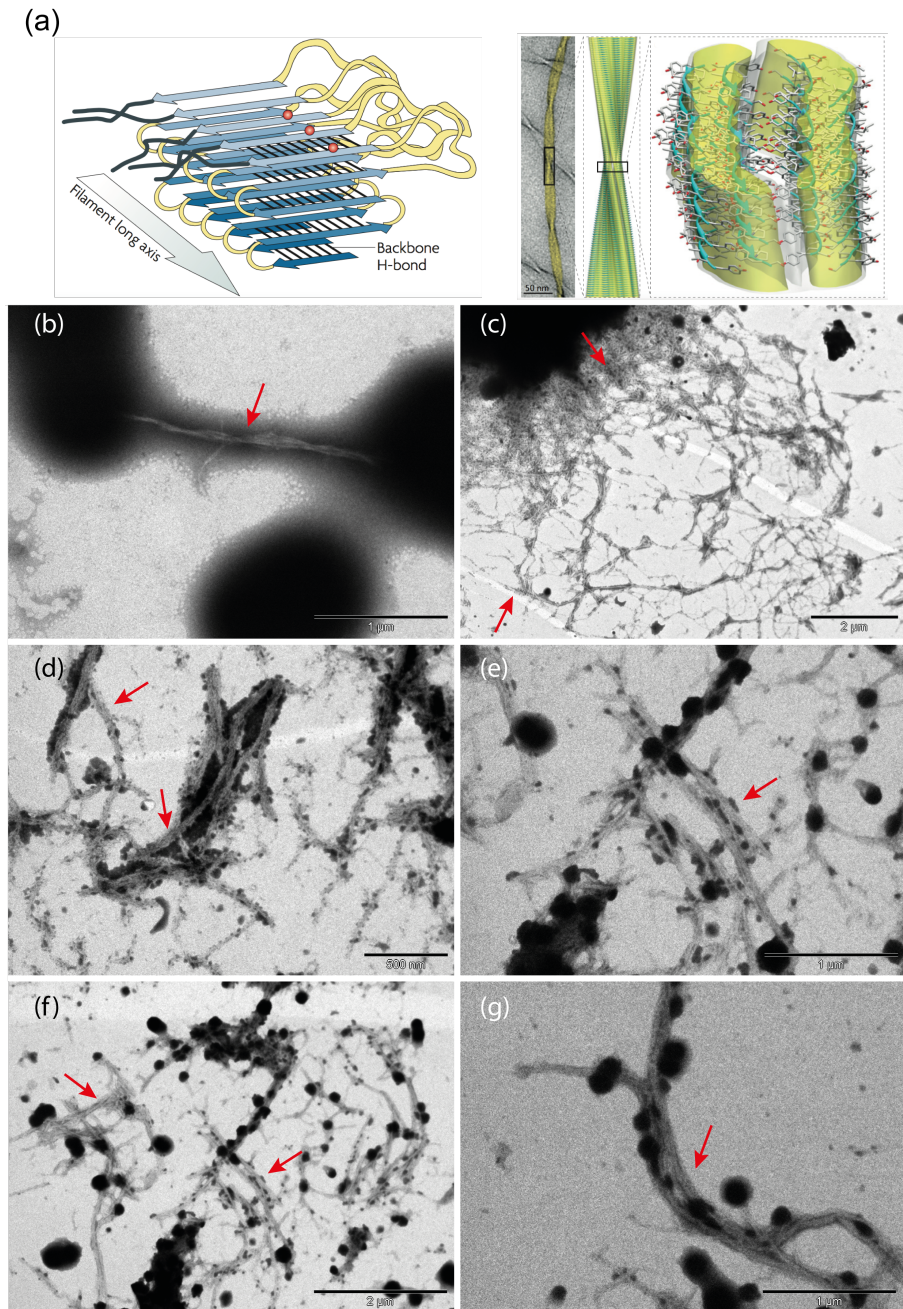
Figure 9. (a) Amyloid structure model. The parallel β-sheet structure of the prion domain of the yeast protein Sup35p. β-strands (blue arrows) run perpendicular to the long axis of the filaments and are connected by loops (yellow). This structure can explain the transmission of prion-variant information, as

(*Figure 9 continued*) the entirety of each prion domain contacts those of the next and previous molecules in the filament. (Jiménez *et al*., 2002; Wickner *et al*., 2005). Transmission electron microscopy pictures of (b) Sup35N and (c), (d), (e), (f), (g) Ixr1 PrD4 amyloids with different aggregation morphologies.

Figure 10. (a) PrD1 24h, (b) PrD2 24h, (c) PrD3 24h, (d) PrD4 24h, (e) PrD3 Dapi 72h stained with DAPI, (f) PrD3 72h.

Since foci formation can also be produced by overexpression of many non-prion proteins, cells from the different cultures and incubation times were lysed and analyzed by semi-denaturing detergent-agarose gel electrophoresis (SDD-AGE) to determine if *in vivo* Ixr1 PrDs formed aggregates correspond to highly ordered amyloid fibrils or to amorphous superstructures.
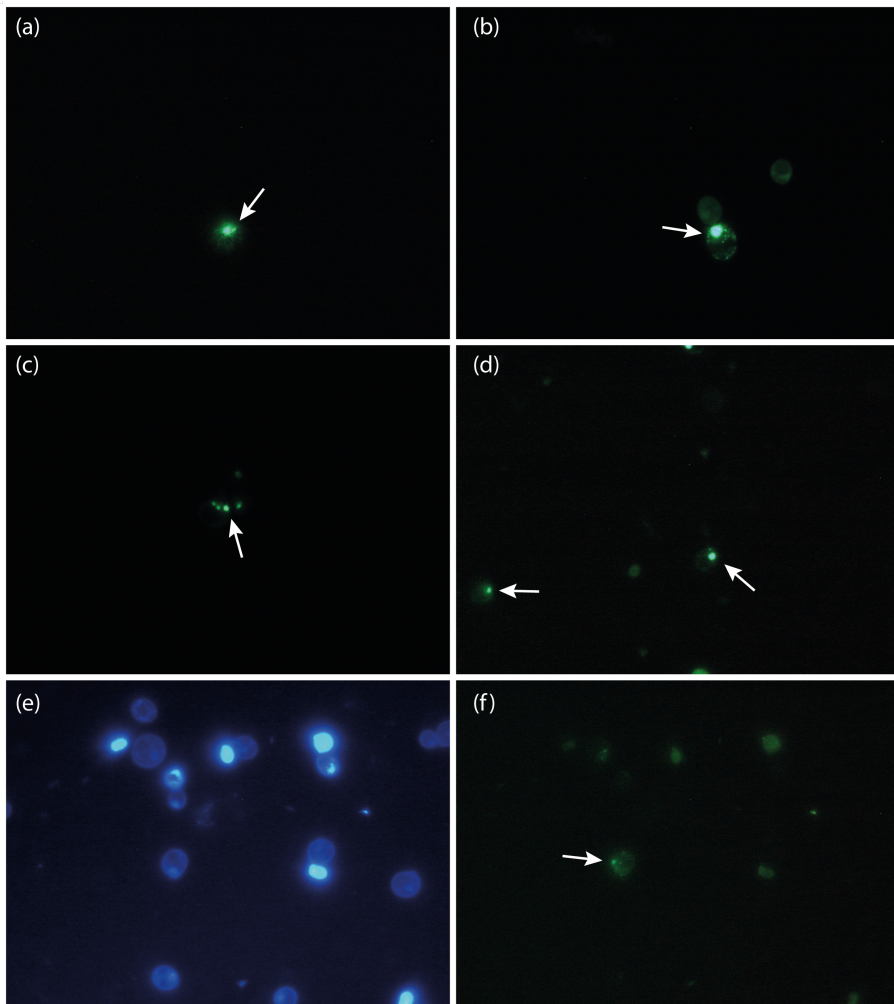
SDD-AGE allows the identification of amyloid SDS-resistant aggregates and the resolution in a wide size-frame, ranging from oligomeric species to polymers assembled from hundreds of individual polypeptides (Bagriantsev *et al.*, 2006). Figure 11 shows that despite the differences in amyloid formation rates and flocculation phenotypes, all PrDs aggregates show resistance to the presence of 0.1% SDS detergent.
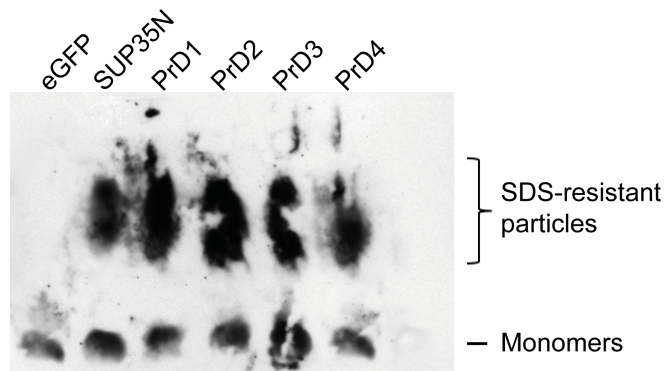
Figure 11. Detection of SDS-resistant aggregates by SDD-AGE in cell lysates of yeast strains expressing PrD1-4-eGFP fusions. Expression of the proteins was induced for 48 hours. eGFP alone and SUP35N were used as control proteins. Proteins were detected with an eGFP-specific antibody (see Materials and methods section).

3.8.- A Sup35p-Based Prion Assay Identifies Phenotypic Switching Behavior

In order to test the ability of each domain to form prions, we used the method previously described by taking advantage of the fact that [*PSI*$^+$] increases read-through of stop codons (Cox, 1965), and therefore [*PSI*$^+$] detection is possible

by nonsense suppression of the *ade2-1* allele. Non-prion *ade2-1* mutants do not grow in medium deprived of adenine or turn red when adenine is limiting; but when these cells express the prion [*PSI*[+]] they can grow in the absence of adenine and remain white on limiting adenine (figure 12a). Sup35p, the protein determining [*PSI*[+]], is modular and consists of an N-terminal PrD (N), a highly charged middle domain (M) and a C-terminal domain (C), which provides the translation termination function. A PrD domain from a prion might be fused to other proteins conferring them the elements of inheritance (Li and Lind- quist, 2000). We take advantage of this property to fusion each PrD of Ixr1 with the C terminal domain of Sup35 to create *PrD-SUP35C* chimeras (under the control of the constitutive *ADH1* promoter) and test them for their ability to suppress the *ade2-1* allele as previously assayed with other prions (Alberti *et al.*, 2009). In the *ade2-1* cells transformed with the chimeras Ade[+] colonies are rare but could result from either DNA mutation or [*PSI*[+]] formation. To distinguish between these two possibilities we tested the effect of PrDs overexpression, since the frequency of events necessary for [*PSI*[+]] formation is dependent on Sup35 concentration, whereas the frequency of chromosomal mutations is not. Overexpression of the different PrDs increased Ade[+] colony formation (figure 12f), suggesting that the Ade[+] phenotype was a result of a prion. Almost no Ade[+] colonies were seen for the PrDs expressed at lower levels (in SRaf medium without PrD overexpression).

We tested the ability of the Ade[+] colonies to maintain that state on nonselective medium after loss of the overexpression plasmid. Indeed, on complete medium, four *PrD-SUP35C* strains displayed a colony color change from red to white or pink that was maintained over several rounds of re-streaking.

All known fungal prions are dependent on chaperones to induce and maintain a prion state, although they differ in the chaperon nature. Generally, amyloid-based fungal prions are critically dependent on the protein remodeling factor Hsp104p, which is inhibited by guanidine hydrochloride (Shkundina and Ter-

Avanesyan, 2007). Therefore we tested yeast cells for prion cure by streaking them in plates containing low concentrations of guanidine hydrochloride (5 mM). Figure 12f shows a positive result for GdmHCl curation (number 4).

(a)

(b)

(c)

(d)

(e)

(f)
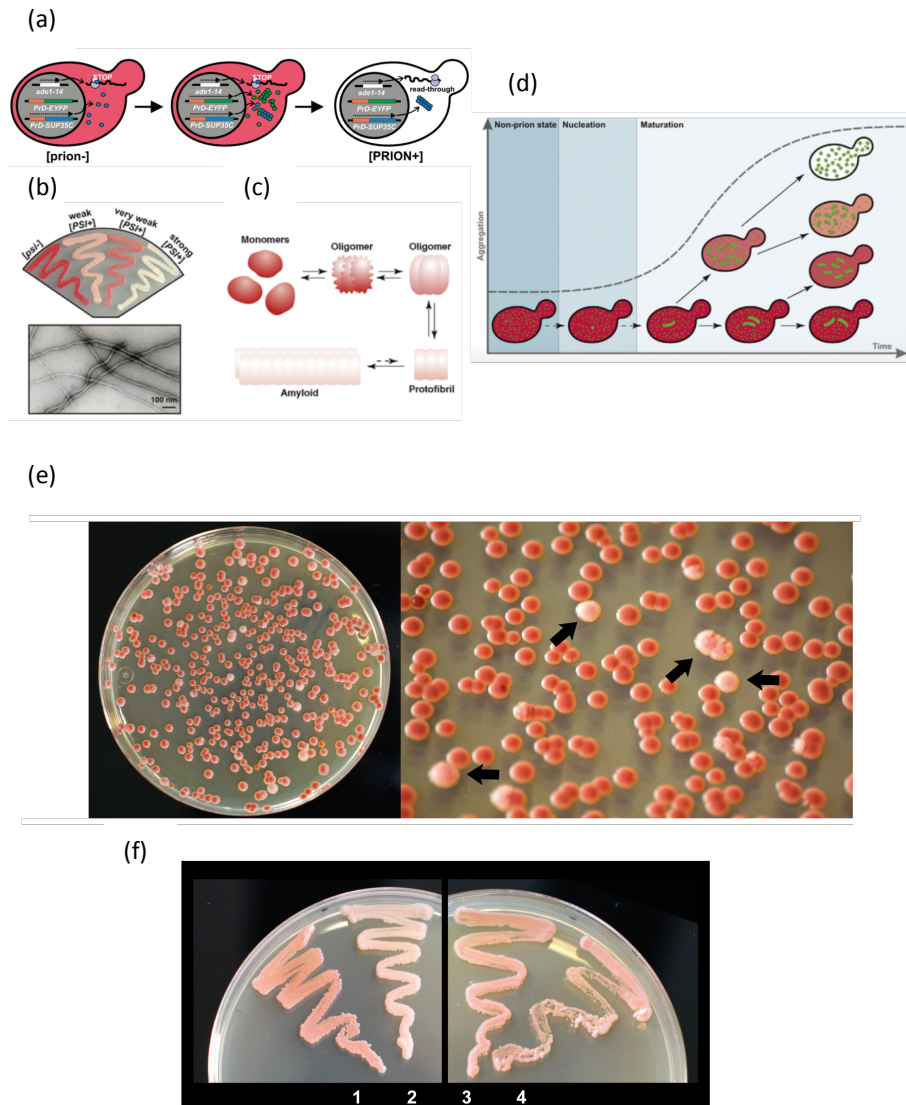
Figure 12. (a) A schematic overview of the genetic manipulations employed to identify PrDs with prion properties. When Sup35 is soluble, translation termination occurs, conferring red colour to cell colonies. When the intracellular protein levels of Sup35 increase as a consequence of transient overproduction, prion formation occurs and avoid Sup35 termination activity, allowing translation and overcoming *ade2-1* mutation, conferring white colour to the cell colonies. (b) Conformational and temporal diversity of

(*Figure 12 continued*) prion states. Prions create multiple stable phenotypic states or ''strains''. [PSI+] strains differ in their levels of nonsense suppression, with stronger strains having lower quantities of functional Sup35 available to fulfill its role in translation termination, giving rise to a whiter coloration in a particular genetic background (top). At the molecular level, strains are determined by amyloid conformational variants (bottom) that arise during nucleation but then stably propagate themselves. (c) Along with the conformational diversity apparent in the end products of amyloid formation, multiple conformational variants are also transiently populated during the early stages of amyloid assembly and might constitute integral on-pathway species. These oligomeric intermediates probably have limited self-templating capacity, but nevertheless may contribute to the weak phenotypes associated with incipient prion states. (d) Incipient prion states acquire progressively stronger phenotypes and stabilities, possibly via mass-action population dynamics of prion particles. Several elegant studies have correlated the phenotypic strength of the prion state with the intracellular number of prion particles. Upon de novo nucleation within a prion-free cell, prion polymerization onto limiting fiber ends proceeds during the ''maturation'' phase under pre-steady state conditions. During each cell division, prion particles are distributed passively and asymmetrically to daughter cells. Progeny that inherit more particles will have faster total prion polymerization rates and correspondingly stronger phenotypes, and will tend to accumulate more prion particles, which will in turn strengthen the prion phenotype in subsequent generations (light pink and white cells). Conversely, cells that inherit fewer particles will have slower polymerization rates and weaker phenotypes (red and pink prion-containing cells) and they will tend to accumulate fewer particles to pass on to their progeny. Such 'noise' in prion distribution might allow prions to stratify protein functionality along a continuum of semistable phenotypes (e.g. red cells, pink cells and white cells) within a small number of cell generations (Halfmann *et al*., 2009). (e) YPD plates after of YRS100 strain after 24 hours overexpression of the PrD1-SUP35C chimera, showing colonies that switched from red to pink (black arrows). (f) Pink colonies were streaked onto YPD plates containing 5 mM GdmHCl to check Hsp104-depence of prion heritance. Numbers indicate (1) YRS100 growth in SDRaf medium and plated in a YPD plate, (2) YRS100 growth in SDRafGal medium and plated in a YPD plate; (3) and (4) cells from (1) and (2) re-streaked in YPD plates containing 5 mM GdmHCl, respectively.

4.- Discussion

In this study we have shown that Ixr1 protein possesses several intrinsically disordered regions flanking two HMG-box domains displayed in tandem. Until the early 1990s, the unique accepted concept of protein function was the well-known protein sequence-structure-function paradigm. According to this concept, a protein could only achieve its biological function upon folding into an exclusive energetically favourable (global energy minimum) conformation determined by its amino acid sequence (referred as "native state"). Loss of specific three-dimensional fold was known as denatured state, in which the protein lost its function. However, the discovery of IDPs (Weinreb *et al*., 1996; Wright *et al*., 1999; Dunker *et al*., 2001, Daughdrill *et al*., 2007), defying structure-function paradigm, have significantly broadened the view of the scientific community, increasing the interest to study these intriguing class of proteins.

Genes encoding IDR amino acid sequences are under reduced selective pressure, which is manifest in a higher sequence diversity compared to genes of structured proteins or domains (Liu *et al.*, 2008). Whereas the functionality of a protein segment is often approached by investigating the evolutionary history of its primary sequence (Capra *et al.*, 2007), this is difficult with IDRs because of their generally high sequence diversity (Chen *et al.*, 2006). There are several bioinformatics tools that help in detect and characterize IDRs. Prediction algorithms of secondary structure in the Ixr1 protein sequence revealed that it only possesses α-helical folded content in their HMG-box domain regions, with extended disordered region in the first half of protein. This region was identified as harbouring IDRs by different IDP analysis algorithms based on amino acid features or machine learning algorithms that used databases of IDP and IDRs.

Dunker and Obradovic (Dunker & Obradovic, 2001) proposed that functional IDRs may exist in two different structural forms: molten globule-like (collapsed) and random coil-like (extended) forms, whereas Uversky suggested the existence of another extended form, the pre-molten globule (Uversky, 2002), distinguishable from the other two by the presence of unstable secondary structure. Several experimental evidences in the present work support the idea that Ixr1 protein possesses IDRs switching between random coil-like to pre-molten and molten globule forms. Analysis of Ixr1 protein size and Stoke's radius by size-exclusion chromatography showed an increment in its hydrodynamic volume of 1.5 times with respect of the theoretically expected, despite behave as a monomer in solution at low protein concentrations (determined by equilibrium sedimentation (SE-AUC) experiments). In this sense, the frictional ratio (f/f_0 = 1.97) of full-length Ixr1 protein obtained from sedimentation velocity (SV-AUC) experiments indicates that it dispose a rod-like structure, with a high axial ratio or a coil-like structure such as that of a denatured protein (McBryant *et al.*, 2006). Moreover, far-UV CD data obtained from full-length Ixr1 protein and their N-terminal truncated region

showed a typical non-folded profile, with a characteristic minimum ellipticity in the vicinity of 190 to 200 nm and values close to zero in the vicinity of 222 nm (Daughdrill *et al.*, 2005; Receveur-Bréchot *et al.*, 2006), meanwhile its HMG-box domains in tandem exhibited predominately α-helical structure as indicated by the characteristic minima at 208 and 222 nm. In this sense, deconvolution using several analysis programs showed up to 69% of disordered contain, meanwhile α-helix regions represent lower percentage (ranging 19-26%). In an ordered protein region, the Ramachandran angles and backbone atoms of each amino acid residue undergo nonisotropic small-amplitude motions relative to their local neighbourhood and are characterized by the equilibrium positions defined by the time-averaged values. In contrast, IDRs do not exhibit dynamic ensembles in which atom positions and backbone Ramachandran angles vary significantly over time with no specific equilibrium angles. In this sense, techniques based on conformational phase transitions upon temperature-induced unfolding may help to identify lack of globularity (and thereby, disorder), detecting no-cooperative or random conformational changes measuring melting temperatures and/or enthalpy of melting of a particular protein. Melting temperatures of full-length Ixr1 protein obtained from far-UV circular dichroism (222 nm fixed wavelength) and differential scanning fluorimetry showed melting temperatures very close to those of their HMG-box domains in tandem. Similar results were obtained from limited proteolysis with proteinase K, since resistant peptides identified by MALDITOF fingerprint matched uniquely the HMG-box domain regions. Taken together, all these results indicate that the Ixr1 protein is partially unfolded, showing large IDRs flanking the HMG-box tandem region.

In general, eukaryotes exhibit higher proteome content in IDRs/IDPs than either prokaryotes or archaea. In *S. cerevisiae* 52-67% of its proteins are predicted to have such long regions of disorder, while bacteria and archaea were predicted to have 16-45 and 26-51% of their proteins with such long disorder regions, respectively (Dunker *et al.*, 2000). It was hypothesized that the higher abundance

of intrinsic disorder in eukaryotes could be a consequence of the increased need for cell signalling and regulation in higher organisms (Singler, 1988; Dunker *et al.*, 2000; Liu & Rost, 2001; Dunker *et al.*, 2001; Iakoucheva *et al.*, 2002; Vucetic *et al.*, 2003). In this sense, IDPs and IDRs are frequently involved in key biological processes such as cell cycle control, transcriptional and translational regulation, signal transduction and membrane transport (Wright & Dyson, 1999; Dyson & Wright, 2002). This probably results from a number of capabilities inherently related to this protein conformation (Uversky *et al.*, 2005; Dunker *et al.*, 2007), such as: i. high specificity with moderate affinity; ii. binding diversity in which one IDR folds differently to accommodate several partner binding interfaces; iii. Binding targets and partners shared, in which various different sequences fold differently but each recognize a common binding surface; iv. larger interaction surfaces wrapping-up or surrounding its partners; v. faster association and dissociation rates. We have shown that Ixr1 protein possesses various IDRs in its N-terminal side with trans-activation activity and it is capable to exhibit disorder-to-order transitions upon the presence of the α-helix inducer agent TFE.

Regulated unfolding or disorder-to-order transitions increase functional versatility of IDPs/IDRs, which allow some proteins to functionally switch between the structured and disordered states. The interactions of IDPs/IDRs are most often mediated by short interaction modules, referred as MoREs (Molecular Recognition Elements) (Oldfield, 2002; Bourhis *et al.* 2004; Dunker *et al.*, 2005). MoREs are tipically <10 amino acids in length which, based on the presence of a few highly conserver, specificity-determining residues, they are recognized and/or modified by structured domains of their interacting partners. Due to their short length, MoREs usually are promiscuous motifs, producing weak and transient interactions. There are several different functional outcomes of MoRE motif recognition, including targeting, regulated degradation, or post-translational modifications. It was previously reported that Ixr1 protein is hyper-phosphorylated in W1588-4C cells under normoxic conditions in rich medium, and that this hyper-

phosphorylated state tightly depends on S366 localized in the long arm of the first HMG-box domain (Tsaponina *et al*., 2011). Besides, phosphorylation in S6 (Soufie *et al*., 2009) and S532 (Albuquerque *et al*., 2008) were previously described, and our proteomic analyses indicate that S36, T45, S83, T218, S232 and S559 are also phosphorylated (data not shown). Strikingly, all these modifications map in regions identified as IDRs in this work, suggesting the important role of this regions in post-translational modifications. Disordered proteins are on average targets of twice as many kinases as structured proteins in yeast (Gasponer *et al*., 2008). The number of phosphorylation events can change the net charge on an IDP and thereby the binding affinity or the kinetics of interaction with a partner (Borg *et al*., 2007; Serber & Ferrel, 2007; Malleshaiah *et al*., 2010). Some of the kinases reported to be implicated in IDP phosphorylation are related to cell-cycle regulation or to the response to certain stimuli or cellular stress condition, fine-tuning the availability of IDPs in different conditions (Gasponer *et al*., 2008).

Multiple mechanisms during transcription and translation also control the availability of IDPs, maintaining most disordered proteins and their transcripts in relatively low levels and for short periods of time compared to structured proteins. In this sense, Ixr1 protein is known to be present in very low protein levels in normal growth conditions (Ghaemmaghami, *et al.* 2003). Several studies have revealed that shorter half-life of disordered proteins could be related to the increased susceptibility of IDRs to the ubiquitin dependent and independent proteasome degradation pathways (Prakash *et al*., 2004; Asher *et al*., 2006; Tsvetkov *et al*., 2009; Schrader *et al*., 2009; Babu *et al*., 2011; Inobe *et al*., 2011).

In this study, we also show that the Ixr1 protein has a high aggregation propensity. It is known that soluble proteins become thermodynamically unstable when the protein concentration is above a few micromoles. This propensity of a protein to misfold and aggregate is determined by its amino acid sequence, where the number of repulsive and attractive interactions increases linearly with protein

concentration. In particular, intrinsically disordered regions with Q/N-rich sequences have been shown to have poor solubility in water and a high aggregation tendency (Michelitsch & Weissman, 2000), in which the collapsed state is maintained through a dynamic network of internal hydrogen bonds involving backbone-to backbone interactions.

However, solubility is not only dependent on the protein concentration but also strongly influenced by its folding state. Unlike globular proteins, IDP cannot adopt a defined structure in the absence of their ligand, so IDP integrity is maintained by interaction with several chaperones. Chaperones of the Hsp70 family constantly monitor the folding state of proteins and recognize and bind misfolded proteins with assistance from co-chaperones of the Hsp40 family. Stress conditions, aging, and pathological conditions lead to an increase in protein misfolding and overwhelming the chaperone machinery capacity. As a consequence, misfolded proteins coalesce into aggregates. Frequently, disorder-to-order transitions of polypeptides within these misfolded proteins are increased by additional intermolecular interactions, eventually leading to the formation of an oligomer with characteristic cross-β structure also known as amyloids. Once this structure is established, amyloid formation becomes self-sustaining and grows by depleting soluble conformers of the same protein. Many yeast prion proteins carry domains of low sequence complexity that are enriched for glutamines and asparagines, but also for other polar amino acids, such as serines and glycines (Alberti *et al*., 2009). Protein domains with such distinctive compositions are referred to as prion-like. Interestingly, prion-like sequence stretches are also present in aggregation-prone proteins are associated with a variety of human neurodegenerative diseases, such as Alzheimer's, Huntington's, and Parkinson's disease. A hallmark of these diseases is the presence of amyloids either in nuclear inclusions, cytoplasmic inclusions, or extracellular aggregates (Forman *et al*., 2004). This indicates that prion-like Q/N-rich proteins have a high aggregation potential and need to be kept under tight control.

Amyloids have specific physicochemical properties: they bind dyes such as Thioflavin-T or Congo Red with high affinity, they show a characteristic red-green birefringence in polarized light, and they show a strong resistance to denaturants, such as SDS. Using these characteristics, we have demonstrated that Ixr1 protein is able to form amyloids *in vitro* and *in vivo*. Full-length Ixr1 protein or different stretches of the prion-like domain expressed and purified in *E. coli* cells, aggregate into amyloid fibrils in adequate solution conditions, showing resistance to 2% SDS and ThT fluorescence signal. In this way, we also show that the different Ixr1 PrDs are able to coalesce when they were overexpressed in SDGalRaf medium. We confirmed that these foci aggregates formed *in vivo* correspond to highly stable structures, showing SDS resistance by SDD-PAGE assay.

After demonstration that Ixr1 can form amyloids, the following step was to check whether effectively this protein could behave as prion, since amyloids are the most common self-templating replicative state found in previously characterized prions (Glover *et al*., 1997; Alberti *et al*., 2009). Also other types of self-propagating protein conformations have been also related to prion phenomena (Wickner *et al*., 2007; Brown & Lindquist, 2009). In this sense, the PrD1-SUP35C chimera showed partial switching from red to light pink in YPD plates. Besides, switching is mediated by Hsp104, as indicates cell curation after guanidinium chloride treatment. This result indicates partial or weak [*PSI*[+]], as previously obtained for proteins that are described to have a prion state, like Cyc8 (Patel *et al*., 2009), or other proteins with high potential to form prions (Alberti *et al*., 2009).

It is frequent that prionogenic proteins are gene products that control gene expression, cell signaling and the response to stimuli such as stress (Alberti *et al*., 2009), in this sense Ixr1 has been characterized as a transcriptional modulator that control gene expression in response to hypoxia and oxidative stress (Castro-Prego et a., 2010a). Besides, the prionogenic proteins usually represent strategic

nodes in the yeast genetic network that control a wide response. For instance, other prionogenic proteins like the Swi1 chromatin remodeler, regulates the expression of 6% of the yeast genome (Du *et al*., 2008); Cyc8 represses 7% of the yeast genes (Green & Johnson, 2004) and other prion candidates, such as Pub1, Ptr69 and Puf2, are RNA-binding proteins that regulate the stability of hundreds of mRNAs encoding functionally related proteins (Hogan *et al*., 2008). As shown in previous chapters, also Ixr1 would fit this criteria since its deletion affect significantly the yeast transcriptome for a representative number of genes.

The prion nature of a protein can allow simple organisms to switch spontaneously between distinct phenotypic states (True & Lindquist, 2000). Starting from a stable intracellular population of non-prion conformers, a single nucleating event may allow initial elongation into a fibrillar species by acting as a template for the conformational conversion (Serio, 2000; Tessier & Lindquist, 2009). Lately, fragmentation of the protein fiber into smaller propagating entities, will allow transmission of the prion and the associated phenotype to daughter cells (Alberti *et al*., 2009). Considering that the change in protein conformation also causes a change in function, these self-perpetuating prions create heritable phenotypes unique to the determinant protein and the specific host cellular genetic background. Therefore these prions would create different phenotypes in the different genetic backgrounds caused by pre-existing genetic polymorphisms creating variant subpopulations with distinct phenotypic states (Seger & Brockmann, 1987). Thus, many prions are likely to create strong and complex phenotypes in populations upon which natural selection can act (Shorter & Lindquist, 2005; Masel & Siegal, 2009).

The variable ways in which prion conformers nucleate and propagate, give them the capacity to stratify protein functionality into several semi-stable levels, which greatly increases the phenotypic diversity created by prion-driven switches as reported both for mammalian and yeast prions. (Halfmann *et al*., 2009;

Pedersen & Oltzen, 2008; Tessier & Lindquist, 2009). For a given prion, multiple distinct yet related protein conformations can each self-perpetuate (for a detailed example of the possible mechanisms implied in generating this diversity see Figure 12b reproduced from Halfmann *et al.*, 2009). Usually phenotypes of prion formation are related to suppression of the function carried by the soluble protein and stronger suppression correlates with higher stability of the prion, which in turn depends on the kinetics of prion formation and fragmentation (King, 2001; Krishnan & Lindquist, 2005; Kryndushkin *et al.*, 2003; Kochneva-Pervukhova *et al.*, 2001).

Whatever they were generated, incipient prion states represent dynamic molecular populations upon which natural selection could operate favoring a certain number of prion particles. This could result in the clonal expansion of some prion-containing cells relative to other cells and acting to shift the distribution of phenotypes within the continuum.

In particular, stress-induced formation of prions and their iterative maturation offers a rapid route to tunable, advantageous phenotypes for adaptation to the new conditions. In this regard it is a future challenge to understand the advantages for cell survival that may derive from Ixr1 prionization. As seen in previous chapters of this PhD Thesis and as already reported (Castro-Prego *et al.*, 2010a; Castro-Prego *et al.*.2010b; Vizoso Vázquez *et al.*, 2012)) Ixr1 is an important regulator of yeast responses to several stresses such as hypoxia, oxidative stress or the addition of cisplatin. In non-stressed conditions the levels of Ixr1 are low (Ghaemmaghami *et al.*, 2003; Castro-Prego *et al.*.2010b) and this could be related to the toxic effect of high expression of this IDP. It has been shown that elevated intracellular levels of Ixr1 induce a necrosis cell-death response (Chen *et al.*, 2013) and we have also experienced, along the processes of Ixr1 purification from yeast cells, that high induced Ixr1 overexpression decrease dramatically cell viability. It is possible that stratification in Ixr1 protein among different monomer-

oligomer-amyloid states during adaptation to stress signals, which increase Ixr1 expression, could allow the population to survive. Future experiments to elucidate the functional significance of a prion state for the Ixr1 protein should contemplate these issues.

5.- REFERENCES

Aguzzi, A., and T. O'Conno (2010). "Protein aggregation diseases: pathogenicity and therapeutic perspectivas." Nature Reviews Drug Discovery **9** (3): 237-248.

Alberti, S., R. Halfmann, O. King, A. Kapila, and S. Lindquist (2009). "A systematic survey identifies prions and illuminates sequence features of prionogenic proteins." Cell **137**(1): 146-158.

Albuquerque C.P., M.B Smolka, S.H Payne, V. Bafna, J. Eng and H. Zhou (2008). "A multidimensional chromatography technology for in-depth phosphoproteome analysis." Molecular & Cellular Proteomics **7**(7):1389-96.

Alexandrov I.M, A.B Vishnevskaya, M.D. Ter-Avanesyan and VV. Kushnirov (2008). "Appearance and propagation of _polyglutamine-based amyloids in yeast: tyrosine residues enable _polymer fragmentation." The Journal of Biological Chemistry **283**(22), 15185-15192.

Asher G, N. Reuven and Y. Shaul (2006). "20S proteasomes and protein degradation ''by default''." BioEssays **28**(8):844-849.

Babu M.M., R. van der Lee , N.S de Groot and J. Gsponer (2011). "Intrinsically disordered proteins: regulation and disease." Current Opinion in Structural Biology **21**(3):432-40.

Bagriantsev, S.N., V.V. Kushnirov, and S.W. **Liebman** (2006). "Analysis of amyloid aggregates using agarose gel electrophoresis." Methods in Enzymology 412, 33-48.

Bernadó, P. and D.I. Svergun (2012). "Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering." Molecular BioSystems **8**(1):151-67

Blocquel, D.; J. Habchi, ; A. Gruet, ; S. Blangy, and S. Longhi (2012). "Compaction and binding properties of the intrinsically disordered C-terminal domain of Henipavirus nucleoprotein as unveiled by deletion studies." Molecular BioSystems **8**(1):392-410.

Borg M, T. Mittag , T. Pawson, M.Tyers, J.D. Forman-Kay and H.S Chan (2007). "Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity." Proceedings of the National Academy of Sciences U S A. **104**(23):9650-5

Bourhis J.M., K. Johansson, V. Receveur-Brechot , C.J. Oldfield , K.A. Dunker , B. Canard B and S. Longhi (2004). "The C-terminal domain of measles virus nucleoprotein belongs to the class of intrinsically disordered proteins that fold upon binding to their physiological partner." Virus Research **99**(2):157-167.

Brachmann, A., U. Baxa, and R. B. Wickner (2005). "Prion generation in vitro: amyloid of Ure2p is infectious." EMBO Journal **24**(17): 3082-3092.

Brown, J.C., and S. Lindquist (2009). "A heritable switch in carbon source utilization driven by an unusual yeast prion." Genes & Development **23**(19): 2320-2332.

Buchan D.W., F. Minneci, T.C., K. Bryson and D.T Jones (2013). "Scalable web services for the PSIPRED Protein Analysis Workbench." Nucleic Acids Research. **41** (Web Server issue): W340-W348.

Buck M. (1998). "Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins." Quarterly Reviews of Biophysics **31**(3):297-355.

Callebaut, I.; G. Labesse, P. Durand, A. Poupon, L. Canard, J. Chomilier, B. Henrissat and J.P Mornon (1997). "Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives." Cellular and Molecular Life Sciences. **53**(8), 621-645. "

Capra J.A. and M. Singh (2007). "Predicting functionally important residues from sequence conservation." Bioinformatics **23**(15):1875-82. "

Chandonia J.M. (2007). "StrBioLib: a Java library for development of custom computational structural biology applications." Bioinformatics **23**(15):2018-2020

Chen J.W., P. Romero, V.N. Uversky and A.K. Dunker (2006). "Conservation of intrinsic disorder in protein domains and families: I A database of conserved predicted disordered regions." Journal of Proteome Research **5**(4):879-87.

Chernoff, Y. O., A. P. Galkin, E. Lewitin, T. A. Chernova, G. P. Newnam S.M. Belenkly (2000) "Evolutionary conservation of prion- forming abilities of the yeast Sup35 protein." Molecullar Microbiology **35**(4): 865-876.

Cox B.S (1965). "PSI, a cytoplasmic suppressor of super-suppressor in yeast." Heredity **26**: 211-232.

Crow, E. T., Z. Du, and L. Li (2011). "A small, glutamine-free domain propagates the [SWI+] prion in budding yeast." Molecullar of Cell Biology **31**(16): 3436-3444.

Crow,E.T. and L. Li (2011). "Newly identified prions in pudding yeast, and their possible functions." Seminars in Cell & Developmental Biology **22**(5), 452-459.

Daughdrill, G. W., G. J. Pielak, V. N. Uversky, M. S. Cortese, and A. K. Dunker (2005). "Natively disordered protein. In Protein Folding Handbook". J. Buchner and T. Kiefhaber, editors. Wiley-VCH:Verlag, Weinheim, Germany. 271-353.

Dengra-Pozo J, S. Martinez-Rodriguez S, L.M. Contreras, J. Prieto, M. Andujar-Sanchez, J.M. Clemente-Jimenez, F.J. Las Heras-Vazquez, F. Rodriguez-Vico and J.L. Neira (2009). "Structure and conformational stability of a tetrameric thermostable N-succinylamino acid racemase." Biopolymers **91**(9):757-772

Derkatch, I. L., M. E. Bradley, P. Zhou, Y. O. Chernoff, and S. W. Liebman (1997). "Genetic and environmental factors affecting the de novo appearance of the [PSI+] prion in *Saccharomyces cerevisiae*." Genetics **147**(2): 507-519.

Di Domenico, T.; I. Walsh, A.J. Martin and S.C Tosatto (2012). "MobiDB: a comprehensive database of intrinsic protein disorder annotations." Bioinformatics **28**(15), 2080-1.

Diaz-Avalos, R., C. Y. King, J. Wall, M. Simon, and D. L. Caspar (2005). "Strain-specific morphologies of yeast prion amyloid fibrils." Proceedings of the National Academy of Sciences **102**(29): 10165-10170.

Dosztányi, Z.; V. Csizmok, P. Tompa and I. Simon (2005). "IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content." Bioinformatics **21**(16), 3433-4.

Dosztányi, Z.; V. Csizmok, P. Tompa, and I. Simon (2005). "The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins." Journal of Molecular Biology **347**(4), 827-39.

Dosztányi Z., B. Mészáros and I. Simon (2009). "ANCHOR: web server for predicting protein binding regions in disordered proteins." Bioinformatics **25**(20):2745-6.

Drozdetskiy A, C. Cole, J. Procter and G.J. Barton (2015). "JPred4: a protein secondary structure prediction server." Nucleic Acids Research **43**(W1):W389-94.

Du, Z., K. W. Park, H. Yu, Q. Fan, and L. Li (2008). "Newly identified prion linked to the chromatin-remodeling factor Swi1 in *Saccharomyces cerevisiae*." Nature Genetics **40**(4): 460-465.

Dunker AK, J.D. Lawson, C.J. Brown, R.M. Williams, P. Romero, J.S Oh, C.J. Oldfield, A.M. Campen, C.M. Ratliff, K.W Hipps, J. Ausio, M.S. Nissen, R. Reeves, C. Kang, C.R. Kissinger, R.W. Bailey, M.D. Griswol, W. Chiu, E.C. Garner and Z. Obradovic (2001). "Intrinsically disordered protein." Journal of Molecular Graphics and Modelling **19**(1), 26-59.

Dunker AK, Z.Obradovic, P. Romero, E.C. Garner and C.J. Brown (2000). "Intrinsic protein disorder in complete genomes." Genome Inform Series Workshop Genome Inform **11**, 161-171.

Dunker, A. K., and Z. Obradovic (2001). "The protein trinitys—linking function and disorder". Nature Biotechnology **19**(9), 805-806.

Dunker, A. K., M.S. Cortese, P. Romero, L.M. Iakoucheva, and V.N. Uversky (2005). "Flexible nets. The roles of intrinsic disorder in protein interaction Networks" FEBS Journal **272**(20), 5129-5148.

Dyson, H. J., and P.E. Wright (2002). "Coupling of folding and binding for unstructured proteins." Current Opinion in Structural Biology **12**(1), 54-60.

Dyson, H. J., and P.E. Wright (2005). "Intrinsically unstructured proteins and their functions." Nature Review of Mollecular Cell Biology **6**(3), 197-208.

Ericsson U.B., B,M. Hallberg, G.T. Detitta, N. Dekker and P.Nordlund (2006). "Thermofluor-based high-throughput stability optimization of proteins for structural studies." Analytical Biochemistry **357**(2):289-98.

Escher, D. and W. Schaffner (1996). "Improved `activator trap' method for the isolation of transcriptional activation domains from random DNA fragments." Biotechniques **21**(5), 848-854.

Chiti F. And C. M. Dobson (2006). "Protein misfolding, functional amyloid, and human disease." Annual Review of Biochemistry 75, 333-66.

Feldman, H. J. And C.W. Hogue (2000). "A fast method to sample real protein conformational space." Proteins **39**(2), 112-131."

Fontana A, P.P. de Laureto, B. Spolaore, E. Frare, P. Picotti and M. Zambonin (2004). "Probing protein structure by limited proteolysis." Acta biochimica Polonica **51**(2):299-321.

Fontana A, P. Polverino de Laureto, V. De Filippis, E. Scaramella, and M. Zambonin (1997). "Probing the partly folded states of proteins by limited proteolysis." Folding and Design **2**(2), R17 -R26

Forman MS, J.Q. Trojanowski and V.M Lee (2004). "Neurodegenerative diseases: A decade of discoveries paves the way for therapeutic breakthroughs." Nature Methods **10**(10):1055-1063.

Fukuchi, S.; S. Sakamoto, Y. Nobe, S.D. Murakami, T. Amemiya, K. Hosoda, R. Koike, H. Hiroaki and M. Ota (2012). "IDEAL: Intrinsically Disordered proteins with Extensive Annotations and Literature." Nucleic Acids Research **40**(Database issue): D507-11.

Garner E, P. Romero, A.K. Dunker, C. Brown and Z. Obradovic (1999). "Predicting binding regions within disordered proteins." Genome Information Services Workshop Genome Inform. **10**: 41-50.

Ghaemmaghami S, W.K. Huh, K. Bower, R.W. Howson, A. Belle, N. Dephoure, E.K. O'Shea and J.S. Weissman (2003). "Global analysis of protein expression in yeast." Nature **425**(6959):737-41.

Glover, J. R., A. S. Kowal, E. C. Schirmer, M. M. Patino, J. J. Liu and S. Lindquist (1997). "Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*." Cell **89**(5): 811-819.

Green, S.R. and A.D. Johnson (2004). "Promoter-dependent roles for the Srb10 cyclin-dependent kinase and the Hda1 deacetylase in Tup1- mediated repression in *Saccharomyces cerevisiae*." Molecular Biology of the Cell **15**(9), 4191-4202.

Gruber CW, M. Cemazar, A. Mechler, L.L. Martin and D.J. Craik (2009). "Biochemical and biophysical characterization of a novel plant protein disulfide isomerase." Biopolymers **92**(1):35-43

Gsponer J, M.E. Futschik, S.A. Teichmann and M.M. Babu (2008). "Tight regulation of unstructured proteins: from transcript synthesis to protein degradation." Science **322**(5906):1365-1368.

Gstaiger, M. and W. Schaffner (1994). "Strong transcriptional activators isolated from viral DNA by the `activator trap', a novel selection system in mammalian cells." Nucleic Acids Research **22**(20): 4031-4038.

Guarente, L. (1983). "Yeast promoters and lacZ fusions designed to study expression of cloned genes in yeast." Methods in Enzymology 101: 181-191.

Haass C. and D.J. Selkoe (2007). "Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide." Nature Review of Mollecular of Cell Biology **8**(2):101-12.

Habchi J., P. Tompa, S. Longhi and V.N. Uversky (2014). "Introducing protein intrinsic disorder." Chemical Society Review **114**(13):6561-88

Halfmann R, S. Alberti, R. Krishnan, N. Lyle, C.W. O'Donnell, O.D. King, B. Berger, R.V. Pappu and S. Lindquist (2011). "Opposing effects of glutamine and asparagine

govern prion formation by intrinsically disordered proteins." Molecullar and Cell **43**(1):72-84.

Halfmann, R. and Lindquist, S. (2010). "Epigenetics in the extreme: prions and the inheritance of environmentally acquired traits." Science **330**(6004): 629-32.

Hirst M, M.S. Kobor, N. Kuriakose, J. Greenblatt and I.Sadowski (1999). "GAL4 is regulated by the RNA polymerase II holoenzyme-associated cyclin-dependent protein kinase SRB10/CDK8." Molecular Cell **3**(5):673-8.

Hogan, D.J., D.P. Riordan, A.P. Gerber, D. Herschlag and P.O Brown (2008). "Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system." PLoS Biology **6**(10), e255.

Hubbard, S.J., F. Eisenmenger and J.M. Thornton (1994). "Modeling studies of the change in conformation required for cleavage of limited proteolytic sites." Protein Science **3**(5): 757-768.

Iakoucheva LM, C.J. Brown, J.D. Lawson, Z. Obradovic Z and A.K. Dunker (2002). "Intrinsic disorder in cell-signal- ing and cancer-associated proteins." Journal of Molecullar Biology **323**(3): 573-584.

Inchaurrondo V.A., O.M. Yantorno, and C.E. Voget (1994). "Yeast growth and galactosidase production during aerobic batch cultures in lactose limited synthetic medium". Process Biochemistry 29: 47-54.

Inobe T., S. Fishbain, S. Prakash and A. Matouschek (2011). "Defining the geometry of the two-component proteasome degron." Nature of Chemical Biology **7**(3):161-167.

James P., J. Halladay and E.A Craig (1996). "Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast." Genetics. **144**(4):1425-36.

Jiménez J.L, E.J. Nettleton, M. Bouchard, C.V. Robinson, C.M. Dobson and H.R. Saibil (2002). "The protofilament structure of insulin amyloid fibrils." Proceedings of the National Academy of Sciences U S A. **99**(14):9196-201.

Johnson W.C. (1999). "Analyzing protein circular dichroism spectra for accurate secondary structures." Proteins **35**(3): 307-312.

Jones DT. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." Journal of Molecullar Biology **292**(2): 195-202.

King C.Y. (2001). "Supporting the structural basis of prion strains: Induction and identification of [PSI+] variants." Journal of Molecullar Biology **307**(5):1247-60.

Kochneva-Pervukhova N. V., M.B. Chechenova, I.A. Valouev, V.V. Kushnirov, V.N. Smirnov and M.D. Ter-Avanesyan (2001). "[PSI+] prion generation in yeast: Characterization of the "strain" difference." Yeast **18**(6):489-97.

Krishnan R and S.L. Lindquist (2005). "Structural insights into a yeast prion illuminate nucleation and strain diversity." Nature **435** (7043):765-72.

Kryndushkin DS, I.M. Alexandrov, M.D. Ter-Avanesyan and V.V. Kushnirov (2003). "Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104." Journal of Biological Chemistry **278** (49):49636-43.

Kushnirov V.V, N.V. Kochneva-Pervukhova, M.B. Chechenova, N.S. Frolova and M.D Ter-Avanesyan (2000). "Prion properties of the Sup35 protein of yeast Pichia methanolica." EMBO Journal **19**(3): 324-331.

Laemmli U.K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". Nature **227**(5259):680-5.

Laue, T.M., B.D. Shah, T.M. Ridgeway and S.L.Pelletier (1992). "Computer-aided interpretation of analytical sedimentation data for proteins" in Analytical ultracentrifugation in biochemistry and polymer science (Harding, S.E., Rowe, A.J. and Horton, J.C., Eds.), pp. 90-125, The Royal Society of Chemistry, Cambridge, UK.

M. le Maire M, B. Arnou, C. Olesen, D. Georgin, C. Ebel, J.V. Møller (2008). "Gel chromatography and analytical ultracentrifugation to determine the extent of detergent binding and aggregation, and Stokes radius of membrane proteins using sarcoplasmic reticulum Ca2+-ATPase as an example." Nature Protocols **3**(11):1782-95.

Mészáros B., I. Simon and Z. Dosztányi (2009). "Prediction of Protein Binding Regions in Disordered Proteins." PLoS computational biology **5**(5): e1000376.

LeVine, H. (1997). "Stopped-flow kinetics reveal multiple phases of thiofla- vin T binding to Alzheimer beta (1-40) amyloid fibrils." Archives Biochemistry Biophysics **342**(2): 306-316.

Li, L., and S. Lindquist (2000). "Creating a protein-based element of inheri- tance." Science **287**(5453):661-664.

Lieutaud, P.; B. Canard and S. Longhi (2008). "MeDor: a metaserver for predicting protein disorder." BMC Genomics 9, S25. "

Linding, R.; L.J. Jensen, F. Diella, P. Bork, T.J. Gibson and R.B Russell (2003). "Protein disorder prediction: implications for structural proteomics." Structure **11**(11): 1453-9.

Linding, R.; R.B. Russell, V. Neduva and T.J. Gibson (2003). "GlobPlot: Exploring protein sequences for globularity and disorder." Nucleic Acids Research **31**(13): 3701-8.

Liu J., N.B. Perumal, C.J. Oldfield, E.W, V.N. Uversky VN and A.K Dunker (2006) "Intrinsic disorder in transcription factors." Biochemistry **45**(22):6873-88.

Liu J and B. Rost (2001). "Comparing function and structure between entire proteomes." Protein Science **10**(10):1970-1979.

Liu J., Y. Zhang, X. Lei and Z. Zhang (2008). "Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective." Genome Biology **9**(4):R69.

Liu Z., K. Chen, A. Ng, Z. Shi, R.W. Woody and N.R Kallenbach (2004). "Solvent dependence of PII conformation in model alanine peptides." Journal of the American Chemical Society **126**(46):15141-15150

Lunn F.A., T.J. MacLeod and S.L Bearne (2008). "Mutational analysis of conserved glycine residues 142, 143 and 146 reveals Gly(142) is critical for tetramerization of CTP synthase from Escherichia coli." Biochemical Journal **412**(1):113-121 .

Malleshaiah M.K., V. Shahrezaei, P.S. Swain and S.W. Michnick (2010) "The scaffold protein Ste5 directly controls a switch-like mating decision in yeast." Nature **465**(7294):101-105.

Mancheño J.M., H. Tateno, I.J. Goldstein, M. Martínez-Ripoll, J.A. and J.A Hermoso (2005). "Structural analysis of the Laetiporus sulphureus hemolytic pore-forming lectin in complex with sugars." Journal of Biological Chemistry. 280(17):17251-9.

Masel, J. and M.L. Siegal (2009) "Robustness: mechanisms and consequences." Trends Genetics. 25(9): 395-403.

McA'Nulty M.M. and S.J. Lippard (1996). "The HMG-domain protein Ixr1 blocks excision repair of cisplatin-DNA adducts in yeast." Mutation Research **362**(1):75-86.

McA'Nulty M.M., J.P. Whitehead and S.J. Lippard (1996). "Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts in vitro and in vivo." Biochemistry **35**(19):6089-99.

McBryant, S. J., C. Krause and J.C. Hansen (2006) Biochemistry **45**(51): 15941-15948 .

Michelitsch M.D.M and J.S. Weissma (2000). "A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions." Proceedings of the National Academy of Sciences USA **97**(22):11910-11915.

Miri Disfani F., W.L. Hsu, M.J. Mizianty, C. Oldfield, B. Xue, A.K. Dunker, V. Uversky and L.. Kurgan (2012). "MoRFpred, a computational tool for sequence-based prediction and characterization of disorder-to-order transitioning binding sites in proteins." Bioinformatics **28**(12):i75-i83.

Nakayashiki T., K. Ebihara, H. Bannai and Y. Nakamura (2001). "Yeast [PSI+] ''prions'' that are crosstransmissible and susceptible beyond a species barrier through a quasi-prion state." Mollecular Cell **7**(6), 1121-1130.

Nakayashiki, T., C. P. Kurtzman, H. K. Edskes, and R. B. Wickner (2005). "Yeast prions [URE3] and [PSI+] are diseases." Proceedings of the National Academy of Sciences USA **102**(9): 10575-10580.

Niesen F.H., H. Berglund and M. Vedadi (2007). "The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability." Nature Protocols **2**(9):2212-21."

Obradovic Z., K. Peng, S. Vucetic, P. Radivojac, C.J. Brown and A.K. Dunker (2003) "Predicting intrinsic disorder from amino acid sequence." Proteins **53** Suppl 6:566-72.

Obradovic Z., K. Peng, S. Vucetic, P. Radivojac and A.K. Dunker (2005). "Exploiting heterogeneous sequence properties improves prediction of protein disorder." Proteins **61** Suppl 7:176-82"

Oldfield C.J. (2002). "Predictions of disordered proteins and protein binding regions: applications to structural genomics." Honors Thesis, Washington State University Pullman, Washington.

Osherovich, L.Z., B.S. Cox, M.F. Tuite and J.S. Weissman (2004). "Dissection and design of yeast prions." PLoS Biology 2(4): E86.

Patel, B. K., and S. W. Liebman (2007). "Prion-proof" for [PIN+]: infection with in vitro-made amyloid aggregates of Rnq1p-(132-405) induces [PIN+]." Journal of Molecular Biology **365**(3): 773-782.

Patel, B. K., J. Gavin-Smyth, and S. W. Liebman (2009). "The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion." Nature Cell Biology **11**(3): 344-349.

Pedersen, J.S. and D.E. Otzen (2008). "Amyloid -a state in many guises: survival of the fittest fibril fold." Protein Science 17(1): 2-10.

Prakash S., L. Tian, K.S. Ratliff, R.E. Lehotzky and A. Matouschek (2004). "An unstructured initiation site is required for efficient proteasome-mediated degradation." Nature Structure of Molecullar Biology **11**(9):830-837.

Prilusky, J.; C.E. Felder, T. Zeev-Ben-Mordehai, E.H. Rydberg, O. Man, J.S. Beckmann, I. Silman and J.L. Sussman (2005). "FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded." Bioinformatics **21**(16): 3435-8.

Provencher S.W. and J. Glöckner (1981). "Estimation of globular protein secondary structure from circular dichroism." Biochemistry **20**(1): 33-37.

Radivojac P., L.M. Iakoucheva, C.J. Oldfield, Z. Obradovic, V.N Uversky and A.K. Dunker (2007). "Intrinsic disorder and functional proteomics." Biophysics Journal **92**(5): 1439-1456.

Receveur-Bréchot V., J.M. Bourhis, V.N. Uversky, B. Canard and S. Longhi (2006). "Assessing protein disorder and induced folding." Proteins 62(1): 24-25.

Romero P., Z. Obradovic A.K. and Dunker (1997). "Sequence data analysis for long disordered regions prediction in the calcineurin family." Genome Informatics Services Worshop Genome Informetion 8: 110-124.

Romero P., Z. Obradovic Z, X. Li , E. Garner, C. Brown and A.K. Dunker (2001). "Sequence complexity of disordered protein." Proteins **42**(1): 38-48.

Ross E.D., H.K. Edskes, M.J. Terry and R.B. Wickner (2005). "Primary sequence independence for prion formation." Proceedings of the National Academy of Sciences USA **102**(36):12825-12830.

Ruden,D.M., J. Ma, Y. Li, K. Wood and M. Ptashne (1991). "Generating yeast transcriptional activators containing no yeast protein sequences." Nature **350**(6315): 250-252

Sadowski, I., B. Bell, P. Broad and M. Hollis (1992). "GAL4 fusion vectors for expression in yeast or mammalian cells." Gene **118**(1): 137-141.

Dutta S. and D. Bhattacharyya (2001). "Size of Unfolded and Dissociated Subunits versus that of Native Multimeric Proteins." Journal of Biological Physics 27(1): 59-71.

Scherzinger, E., A. Sittler, K. Schweiger, V. Heiser, R. Lurz, R. Hasenbank, G.P. Bates, H. Lehrach, and E.E. Wanker (1999). "Self-assembly of polyglut- amine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology." Proceedings of the National Academy of Sciences USA **96**(8): 4604-4609.

Schrader E.K., K.G. Harstad KG and A.Matouschek (2009). "Targeting proteins for degradation." Nature of Chemical Biology **5**(1):815-822.

Schuck P. (2000). "Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." Biophysics Journal **78**(3): 1606-1619.

Sechi S. and B.T Chait. (1998) "Chait Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification." Analytical Chemistry **70**(24): 5150-5158.

Seger, J. and H.J. Brockmann (1987). "What is bet-hedging?" Oxford Surveys in Evolutionary Biology **4**:182-211.

Serber Z. and J.E. Ferrell Jr (2007). "Tuning bulk electrostatics to regulate protein function." Cell **128**(3):441-444.

Serio T.R., A.G. Cashikar, A.S. Kowal, G.J. Sawicki, J.J. Moslehi, L. Serpell M.F Arnsdorf and S.L. Lindquist (2000). "Nucleated conformational conversion and the replication of conformational information by a prion determinant." Science **289**(5483): 1317-1321.

Shi Z., C.A. Olson, G.D. Rose, R.L. Baldwin and N.R. Kallenbach (2002). "Polyproline II structure in a sequence of seven alanine residues." Proceedings of the National Academy of Sciences U S A **99**(4):9190-9195.

Shkundina, I.S., and M.D. Ter-Avanesyan (2007). "Prions." Biochemistry **72**(13): 1519-1536.

Shorter J. and S. Lindquist (2005). "Prions as adaptive conduits of memory and inheritance." Nature Review Genetics **6**(6): 435-450.

Shorter J., and S. Lindquist, S. (2008). "Hsp104, Hsp70 and Hsp40 interplay regulates formation, growth and elimination of Sup35 prions." The EMBO journal **27**(20): 2712-2724.

Shriver J.W. and S.P. Edmondson (2009). "Defining the stability of multimeric proteins." Method of Molecullar Biology **499**:57-82

Sickmeier M., J.A. Hamilton, T. LeGall, V. Vacic, M. S. Cortese, A. Tantos, B. Szabo, P. Tompa, J. Chen, V.N Uversky, Z. Obradovic and A.K. Dunker (2007). "DisProt: the Database of Disordered Proteins." Nucleic Acids Research (Database issue): D786-93.

Sigler P.B. (1988). "Transcriptional activation. Acid blobs and negative noodles." Nature **333**(6170): 210-212.

Sondheimer, N., and S. Lindquist (2000). "Rnq1: an epigenetic mod- ifier of protein function in yeast." Molecullar Cell **5**(1): 163-172.

Song Y., Y.X. Wu YX, G. Jung, Y. Tutar, E. Eisenberg, L.E. Greene, and D.C. Masison (2005). "Role for Hsp70 chaperone in *Saccharomyces cerevisiae* prion seed replication." Eukaryot Cell **4**(2):289-297.

Soufi B., C.D. Kelstrup, G. Stoehr, F. Fröhlich, T.C. Walther and J.V Olsen (2009). "Global analysis of the yeast osmotic stress response by quantitative proteomics." Molecullar Biosystems **5**(11):1337-46.

Sreerama N. and R.W. Woody (2000). "Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set." Analytical Biochemistry **287**(2):252-60.

Tanaka, M., P. Chien, N. Naber, R. Cooke, and J. S. Weissman (2004). "Conformational variations in an infectious protein determine prion strain differences." Nature **428**(6980): 323-328.

Tessier P.M. and S. Lindquist (2009). "Unraveling infectious _structures, strain variants and species barriers for the yeast prion _[PSI+]." Nature Structural Mollecular Biology **16**(6): 598-605.

Titz B., S. Thomas, S.V. Rajagopala, T. Chiba, T. Ito T and P.Uetz (2006). "Transcriptional activators in yeast." Nucleic Acids Research **34**(3):955-67.

Tompa, P. (2002). "Intrinsically unstructured proteins." Trends of Biochemical. Science 27(10): 527-33.

Toombs J.A., M. Petri, K.R. Paul, G.Y. Kan, A. Ben-Hur, E.D. Ross (2012) "De novo design of synthetic prion domains." Proceedings of the National Academy of Sciences U S A. **109**(17):6519-24.

True, H.L. and S.L. Lindquist (2000). "A yeast prion provides a mechanism for genetic variation and phenotypic diversity." Nature **407**(6803): 477-483.

Tsaponina O., E. Barsoum E, S.U. Aström, A. Chabes (2011). "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools." PLoS Genetics **7**(5):e1002061.

Tsaponina O. and A. Chabes (2013). "Pre-activation of the genome integrity checkpoint increases DNA damage tolerance." Nucleic Acids Research **41**(22):10371-8.

Tsvetkov P., N.Reuven, C. Prives and Y. Shaul (2009). "Susceptibility of p53 unstructured N terminus to 20S proteasomal degradation programs the stress response." Journal of Biological Chemistry **284**(39):26234-26242.

Tyedmers, J., M.L. Madariaga and S. Lindquist (2008). "Prion switching in response to environmental stress." PLoS Biology **6**(11): e294.

Uversky, V. N. (2002). "Natively unfolded proteins: a point where biology waits for physics." Protein Science **11**(4):739-756.

Uversky V. N. (2002). "What does it mean to be natively unfolded?" European Journal of Biochemistry **269**(1): 2-12.

Uversky V. N. (2002) "Natively unfolded proteins: a point where biology waits for physics." Protein Science **11**(4): 739-56.

Uversky V. N., C. J. Oldfield and A. K. Dunker (2005). "Showing your ID: intrinsic disorder as an ID for recognition, regulation, and cell signaling." Journal of Molecular Recognition **18**(5):343-384.

Uversky V. N., J.R. Gillespie and A.L. Fink (2000). "Why are ""natively unfolded proteins unstructured under physiologic conditions?" Proteins **41**(3): 415-427.

Uversky V. N. and A.K. Dunker (2012). "Multiparametric analysis of intrinsically disordered proteins: looking at intrinsic disorder through compound eyes." Analytical Chemistry 84(5): 2096-104.

Uversky V. N. and A.K. Dunker (2010) "Understanding protein non-folding." Biophysics Acta 1804(6): 1231-64.

Uversky V. N. and S.Longhi (2010). "Instrumental Analysis of Intrinsically Disordered Proteins" John Wiley & Sons: Hoboken, NJ.

Uversky V. (1993). "Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule." Biochemistry **32** (48): 13288-13298.

Vucetic S., C.J. Brown, A.K. Dunker and Z. Obradovic (2003). "Flavors of protein disorder." Proteins 52(4): 573-584.

Ward J. J., J.S. Sodhi, L. J. McGuffin, B.F Buxton and D.T. Jones (2004). "Prediction and functional analysis of native disorder in proteins from the three kingdoms of life." Journal of Molecullar Biology 337(3): 635-45.

Ward J. J., L.J. McGuffin, K. Bryson, B.F. Buxton and D.T. Jones (2004). "The DISOPRED server for the prediction of protein disorder." Bioinformatics 20(13): 2138-9.

Weinreb, P. H., W. Zhen, A. W. Poon, K. A. Conway, and P. T. Lansbury, Jr. (1996). "NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded." Biochemistry **35**(43):13709-13715.

Whitmore L. And B.A. Wallace (2004). "DICHROWEB: an online server for protein secondary structure analyses from circular dichroism spectroscopic data." Nucleic Acids Research 32 (Web server issue): W668-W673.

Whitmore L. and B.A. Wallace (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases." Biopolymers **89**(5): 392-400.

Wickner R. B. (1994). "[URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*." Science **264**(5158): 566-569.

Wickner R.B., H.K. Edskes, F. Shewmaker, T. Nakayashiki, A. Engel, L. McCann and D. Kryndushkin (2007). "Yeast prions: evolution of the prion concept." Prion **1**(2): 94-100.

Wiesner C., M. Hoeth, B.R. Binder and R. de Martin (2002). "A functional screening assay for the isolation of transcription factors." Nucleid Acids Research **30**(16): e80.

Wittig I., H.P. Braun and H. Schägger (2006). "Blue native PAGE." Nature Protocols **1**(1):418-28.

Wright P. E., and H. J. Dyson (1999). "Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm." Journal of Molecullar Biology **293**(2):321-331.

Xue, B., R.L. Dunbrack, R.W. Williams, A.K Dunker and V.N. Uversky (2010). "PONDR-FIT: a meta-predictor of intrinsically disordered amino acids." Biochimica et Biophysica Acta 1804(4): 996-1010.

Ye R., Q.H. Yao, Z.H. Xu and H.W. Xue (2004) "Development of an efficient method for the isolation of factors involved in gene transcription during rice embryo development." Plant Journal **38**(2):348-57.

Zhong L. and W.C. Johnson Jr (1992). "Environment affects amino acid preference for secondary structure." Proceedings of the National Academy of Sciences USA **89**(10):4462-4465.

Zitomer R.S. and B.D. Hall BD (1976). "Yeast cytochrome c messenger RNA. In vitro translation and specific immunoprecipitation of the CYC1 gene product." Journal of Biological Chemistry **251**(20): 6320-6326.

Concluding remarks

The main conclusions from the studies reported in this work can be summarized as follows:

1. Genome-wide transcriptional analyses reveal that Ixr1 participates in the yeast hypoxic response to adapt the use and production of cellular energy to oxygen disposal. In this sense, Ixr1 regulates the transcript levels of a number of genes related to oxidative stress response, cell-wall composition, cellular-energetics, the metabolism of sulphur and branched-chain (BC) amino acids and DNA synthesis.

2. Ixr1 has switching regulatory capabilities, which are dependent on oxygen availability, and activates or represses the same target genes. The genes from pathways of sulphate assimilation and branched-chain (BC) amino acids illustrate this mechanism of regulation mediated by Ixr1.

3. ChIP-on-chip analysis of Ixr1 binding sites both during normoxia and hypoxia reveal low enrichment in promoter regions, ranging from 50% to 60% of the total significant binding sites.

4. Ixr1 and Rox1 share common target promoter sequences both through the ATTGTT core defined for HMG box binding or through the redefined consensus ***AAG[G/C]GG***.

5. Regulation of transcription by Ixr1 is mainly attributable to indirect mechanisms, instead of direct binding to regulated promoters, as deduced from the low overlap obtained between transcriptome and ChIP-on-chip data.

6. Genome-wide transcriptional analyses reveal that the effect of Ixr1 in the resistance to cisplatin is attributable mainly to the lower impact of the drug treatment in the ribosome biosynthetic pathway in *Δixr1* null mutants, as well as to an enhancement of sulfur metabolism.

7. Ixr1 function in the response to cisplatin treatment is regulatory and not an adduct-masking effect, as reveals the binding enrichment of Ixr1

binding to promoter regions (from 46% to 69%) in the presence of the drug.

8. We have biochemically and thermodynamically characterized the individual HMG-box domains of Ixr1. Main conclusions are as follows:

    a. HMG-box B shows lower stability than HMG-box A.

    b. HMG-box A shows one order of magnitude higher binding affinities for DNA than HMG-box B, producing also higher bending angles.

    c. Both HMG-box domains bind to linear DNA in a non-sequence specific entropy-driven manner, sharing large entropy values.

    d. The two HMG-box domains present different thermodynamic signatures in the binding to cisplatin-modified DNA, indicating differences in the binding mode.

9. The two in tandem HMG-box domains of Ixr1 bind to DNA with a positive cooperative effect. The binding of HMG-box A domain to DNA is necessary to favour that the HMG-box B domain could adopt an optimal conformation for DNA binding.

10. Ixr1 is an intrinsically disordered protein enriched in disorder-promoting amino acids (glutamine, serine and proline) and displays two intrinsically disorder regions, flanking the HMG-box region and with α-helix propensity.

11. Ixr1 has a high aggregation propensity and is able to form amyloids *in vitro* and *in vivo*. Furthermore, it shows several features that open the possibility that Ixr1 could act as a prion.

Appendix I

Resumen

1.- INTRODUCCIÓN

Las proteínas HMG son proteínas nucleares no histonas que se asocian con cromatina y que tienen un comportamiento altamente dinámico entre todos sus sitios potenciales de unión. El dominio HMG-box está compuesto por 65-85 aminoácidos y se caracteriza por presentar un plegamiento en forma de "L" en la que tres hélices alfa que forman un ángulo de aproximadamente 80° entre los dos brazos. El brazo corto está compuesto por la tercera hélice y una región amino terminal extendida y sin estructura secundaria definida mientras que, la primera y la segunda hélice alfa forman el brazo corto. Se trata de un dominio de pequeño tamaño, con un plegamiento sencillo y único; está conservado en un gran número de proteínas de diferentes especies (Štros *et al.*, 2010).

Basándose en estudios filogenéticos y estructurales, se han definido dos grandes subfamilias de proteínas que contienen dominios HMG. Una primera clase incluye proteínas que se unen a formas de ADN distorsionado con o sin secuencia específica (_Non_ _Sequence_ _Specificity, NSS_) y tienen, en general, dos o más dominios dispuestos en tándem. Una segunda clase de proteínas que contienen dominios HMG-box son aquellas que se unen al ADN por especificidad de secuencia (_Sequence_ _Specificity_, SS) y normalmente contienen un único dominio. Frecuentemente actúan como factores transcripcionales, solo se expresan en determinados tipos celulares y contienen otros dominios reguladores asociados. Los elementos determinantes en la especificidad de secuencia al ADN recaen fundamentalmente en el brazo largo del dominio HMG-box (Štros *et al.*, 2010). A pesar de estas diferencias, ambas subfamilias de proteínas HMG son capaces de unirse con elevada afinidad al ADN lineal a través del surco menor, produciendo un doblamiento acusado en el ADN y formando complejos con una estructura bastante similar. Utilizan su superficie cóncava para intercalar las cadenas laterales de uno (en SS) o dos (en NSS) aminoácidos hidrofóbicos entre los pares de bases del surco menor. Aunque la mayoría de los contactos entre el dominio HMG-box y

el ADN se producen a través del surco menor, también se producen algunos contactos adicionales por la inserción de la cola amino-terminal del dominio en el surco mayor, estabilizando el complejo (Štros, 1998; Travers, 2000; Thomas & Travers, 2001). La interfaz o superficie de contacto DNA-proteína se compone de numerosas interacciones electrostáticas, puentes de hidrogeno y contactos *Van der Waals*, además de contactos mediados por moléculas de agua. Como resultado de estas interacciones, el ADN se dobla y ensancha, ampliando el surco menor y comprimiendo el surco mayor.

Saccharomyces cerevisiae posee 7 genes que expresan proteínas HMG: *ABF2*, *HMO1*, *NHP6A*, *NHP6B*, *NHP10*, *IXR1* y *ROX1* (Bustin, 2001). Cinco de estos genes presentan un único dominio HMG-box, mientras que Abf2 e Ixr1 poseen dos motivos HMG-box en tándem. Entre las proteínas HMG de levaduras, la proteína Ixr1 es la menos conocida, tanto en sus características funcionales como estructurales. El gen *IXR1* de *S. cerevisiae* codifica para una proteína de 67.2 kDa que contiene 3 regiones poliglutamina y dos dominios HMG-box (*High Mobility Group*, HMG) con capacidad de unión a DNA (Lambert *et al*., 1994). El modelado por homología de los dos dominios HMG-box presentes en Ixr1 basados en las estructuras determinadas experimentalmente de las proteínas SRY de mamíferos (Werner, 1995) y HMGB1 de humanos (Read, 1993; Weir, 1993; Hardman, 1995; Stott, 2006) sugiere que la unión puede ser no específica de secuencia a través del dominio HMG-box A y específica de secuencia a través del dominio HMG-box B de Ixr1 (Castro-Prego *et al*., 2010a), sin embargo esta hipótesis no había sido verificada experimentalmente.

La primera referencia sobre la participación de Ixr1 en la respuesta a hipoxia en levaduras se produjo hace más de 20 años, cuando Lambert y colaboradores publicaron que Ixr1 causa represión aerobia del gen *COX5B*, que codifica para la isoforma hipóxica de la subunidad Vb de la citocromo oxidasa *c* del complejo mitocondrial (Lambert *et al*., 1994). A lo largo de estos años este factor

de transcripción se ha relacionado también con otros genes hipóxicos como *TIR1,* una manoproteína de la pared celular perteneciente a la familia de proteínas ricas en serina-alanina (Bourdineaud *et al*., 2000) y *HEM13,* que codifica la enzima coproporfirinógeno III oxidasa en la ruta de biosíntesis del grupo hemo (Castro-Prego *et al*., 2010b). Se conoce muy poco sobre la interacción de Ixr1 con los promotores de los genes diana. Algunos datos indican la posibilidad de un mecanismo de competición o uso alternativo de Rox1 e Ixr1 por la unión al mismo sitio consenso en los promotores co-regulados por estos dos factores de transcripción, ya que ambos contienen dominios HMG-box con capacidad de unión a ADN (Castro-Prego *et al*., 2010b).

Ixr1 está también relacionada con la resistencia de las células de levadura a la droga cisplatino, utilizada en quimioterapia contra el cáncer. *S. cerevisiae* es un buen modelo eucariota para la búsqueda de genes relacionados con la sensibilidad o resistencia al cisplatino (Fox *et al*., 1994; Huang *et al*., 2005; Schenk *et al*., 2001, 2003). Aunque se han identificado 22 genes de levaduras relacionados con cisplatino en cribados de mutantes (Huang *et al*., 2005), es poco lo que se conoce sobre el papel de todos estos genes en la respuesta celular a cisplatino. Ixr1 se une a los aductos que forma el cisplatino con el ADN (Brown *et al*., 1993) y su depleción causa un aumento de la resistencia a cisplatino en levaduras (McA´Nulty & Lippard, 1996). Se ha planteado la hipótesis de que Ixr1 y otras proteínas con dominios HMG-box puedan bloquear la reparación de la mayoría de aductos DNA-cisplatino *in vivo,* desencadenando un mecanismo de muerte celular (McA´Nulty & Lippard, 1996). Efectivamente la depleción de Ixr1 no aumenta la resistencia en cepas de *S. cerevisiae* con mutaciones en los genes *RAD2*, *RAD4* y *RAD14* que están relacionados con los mecanismos de reparación de DNA(McA´Nulty & Lippard, 1996). Más recientemente, se ha observado que mutaciones en Ixr1 aumentan la probabilidad de mutagénesis espontánea producida por errores durante la replicación (Fedorov *et al*. 2010) y que Ixr1 es necesario para el mantenimiento de los niveles de dNTPs, necesarios para la síntesis y reparación del ADN (Tsaponina *et*

al. 2011). El interés en conocer los mecanismos de acción de Ixr1 en relación a la respuesta de las levaduras al cisplatino se debe a observaciones previas en las que se muestra que el gen humano HMGB1, que tiene una secuencia similar a *IXR1*, se sobre-expresa en células cancerosas resistentes a cisplatino (Nagatani *et al*., 2001).

Las regiones poli-glutamina presentes en lxr1 le confieren unas características físico-químicas y estructurales muy peculiares, que se ponen de manifiesto por su elevada tendencia a la agregación durante el proceso de purificación. La agregación de proteínas es un proceso complejo y multicausal que se puede producir por ejemplo ante diversas condiciones de estrés que afectan a la expresión proteica o plegamiento. También puede producirse agregación como consecuencia de mutaciones que producen variantes proteicas o por las propias características de las proteínas de naturaleza intrínsecamente desordenada. Los fragmentos proteicos generados por proteólisis, e incapaces de plegarse correctamente en ausencia de la parte que falta de la cadena poli-peptídica, también son vulnerables de sufrir agregación. Las formas agregadas de las proteínas son, por norma general, amorfas a nivel ultra-estructural. Sin embargo, en algunos casos estos agregados son capaces de reorganizarse y formar fibras amiloideas. Esta fibras son agregados proteicos empaquetados en conformaciones altamente ordenadas. Una de las estructuras caracterizadas más frecuente en las fibras amiloideas presenta un patrón único, '*cross-β,*' estabilizado por puentes de hidrógeno y en el que las hebras *β* están orientadas perpendicularmente al eje de las fibra, proporcionándole gran estabilidad (Aguzzi *et al*., 2014). Estas fibras amiloideas pueden adoptar varias morfologías de protofilamentos trenzados con diferente radio y rigidez (Glover *et al*., 1997; Diaz-Avalos *et al*., 2005).

Dominios ricos en glutamina/asparagina (*Q/N-rich*), similares a los hallados en Ixr1, se encuentran también en todos los priones conocidos, formando parte de los llamados "dominios formadores de priones" (*Prion Formation Domain*, PFD), necesarios para la formación de priones y su propagación. De forma similar a

los priones de mamíferos, varias proteínas de levaduras son viables tanto en forma soluble como en conformación amiloidea, propagándose de manera estable y confiriendo características fenotípicas únicas (Liebman & Chernoff, 2012). Las priones de levaduras se han utilizado frecuentemente en estos estudios, ya que no son promiscuos y son fáciles para manejar *in vitro* usando proteínas recombinantes y tiempos de incubación cortos. Además estos estudios llevados a cabo en levaduras han resuelto varias cuestiones relativas a la formación de priones y su crecimiento, la infectividad de las fibras amiloideas, y la relación entre la estructura amiloidea y la cepa priónica. Cuando las células de levadura experimentan un estímulo de estrés que compromete la homeostasis celular y a la maquinaria de plegamiento de la proteína, la frecuencia con la que los priones aparecen y desaparecen aumenta bruscamente (Tyedmers *et al*., 2008). La formación de priones y su propagación se ve favorecida por su estabilidad, su mecanismo de ensamblaje por polimerización a partir de procesos de nucleación, y la elevada especificidad de unión de las nuevas cadenas proteicas al molde en crecimiento.

2.- OBJETIVOS Y METODOLOGÍA

Los objetivos principales de esta tesis consisten en el estudio de la proteína Ixr1 de *S. cerevisiae*, tanto desde el punto de vista de su función, como regulador transcripcional en la respuesta a hipoxia y al tratamiento con cisplatino, como de las características bioquímicas y estructurales de la proteína.

Para ello, en primer lugar, se hicieron estudios de transcriptómica y ensayos de ChIP-on-chip a escala genómica para identificar promotores regulados que no habían sido todavía identificados y para caracterizar secuencias de ADN reguladoras específicas para la de unión de dos proteínas HMGB de *S. cerevisiae*, Ixr1 y Rox1, relacionadas con la respuesta a hipoxia.

En paralelo, se realizaron experimentos de transcriptómica y ChIP-on-chip

mediante *arrays* para caracterizar el papel de Ixr1 en la respuesta a cisplatino, y averiguar si existe una implicación directa de Swi6 en esta respuesta. Swi6 es un cofactor de gran importancia en la regulación del ciclo celular y de la traducción proteíca (Verma *et al*., 1992; Koch *et al*., 1995; Sidorova *et al*., 1997; Costanzo *et al*., 2003; Stillman, 2013).

La determinación de la estructura de Ixr1 es interesante, puesto que la organización en tandem de sus dos dominios HMG-box no se parece a la descrita en otras proteínas HMGB, cuyas estructuras han sido determinadas experimentalmente, ya que la secuencia linker que separa ambos dominios en Ixr1 está formada por tan sólo tres aminoácidos. Además, como ya se ha comentado, Ixr1 posee tres regiones ricas en glutamina que no presenta ninguna otra proteína HMGB de *S. cerevisiae*. En este trabajo, se ha caracterizado, desde un punto de vista bioquímico y termodinámico, la unión de los dominios HMG-box de Ixr1 a distintos tipos de ADN. Se seleccionaron a este efecto las regiones consenso previamente identificadas por binding *in vitro* e *in vivo* en los promotores de *ROX1* y *HEM13* (Castro-Prego *et al*., 2010a; Castro-Prego *et al*., 2010b) así como otros ADN lineales que no contienen el consenso definido. También se analizó la unión de Ixr1 a ADN platinado, que contiene aductos como los que se forman tras el tratamiento *in vivo* con cisplatino y a ADN cruciforme; este último importante en procesos de recombinación (Taudte *et al*., 2001). Las características de ambos dominios HMG-box de Ixr1 han sido determinadas por dicroísmo circular (CD) y resonancia magnética nuclear (NMR). Las uniones de cada HMG-box a los distintos ADN diana han sido analizadas por ensayos de retardo en gel (EMSA) y fluorescencia polarizada (FP), mientras que las constantes termodinámicas de la unión han sido calculadas por técnicas como calorimetría isotérmica (ITC) o fluorescencia polarizada (FP).

Aunque inicialmente se planteó la resolución de la estructura de la proteína mediante difracción de rayos X, diversos intentos de cristalización de Ixr1

a lo largo de este trabajo fueron fallidos. La producción en levaduras, considerada ventajosa puesto que permitía las modificaciones post-traduccionales, ofrecía un bajo rendimiento debido al carácter tóxico de la proteína al sobre-expresarla. La producción en bacterias fue mucho mayor pero, como consecuencia la peculiar secuencia y composición aminoacídica de Ixr1, los cristales generados eran desordenados, con morfología amorfa y no generaban patrones de difracción de rayos X interpretables. Nos centramos consecuentemente en otros aspectos estructurales abordables sin necesidad de obtener cristales que difractasen. Estudiamos experimentalmente el estado oligomérico, abordamos la caracterización de regiones ordenadas y desordenadas, de las regiones ricas en glutamina y la tendencia a formar amiloides. Para ello utilizamos un amplio abanico de técnicas como la cromatografía de exclusión molecular, ultracentrifugación analítica, electroforesis de geles nativos monodimensionales, dicroísmo circular, proteólisis limitada, sistema de "doble híbrido", cinéticas de formación de amiloides o microscopia electrónica de transmisión.

3.- RESULTADOS

Los resultados obtenidos en los estudios de transcriptómica muestran que Ixr1 regula genes relacionados con la respuesta al estrés oxidativo, composición de la pared celular, energética celular, metabolismo del azufre, biosíntesis de aminoácidos de cadena ramificada y síntesis de ADN de forma altamente dependiente de la disponibilidad de oxígeno. Curiosamente, al comparar la expresión en normoxia o hipoxia, algunos grupos de genes están regulados por Ixr1 de manera opuesta. En definitiva, la regulación mediada por Ixr1 contribuye a adaptar el uso y la producción de energía celular a la disponibilidad de oxígeno.

Análisis *in silico* de los genes regulados por Ixr1 en normoxia e hipoxia descritos en esta tesis revela que Ixr1 puede unirse a secuencias que contienen el *core* de unión de Rox1 ***ATTGTT*** y también a otra caracterizada en este trabajo:

AAG[G/C]GG. Los resultados obtenidos a partir de los ensayos de ChIP-on-chip, analizando los sitios de unión de Ixr1 al genoma de *S. cerevisiae,* mostraron que sólo entre el 12% y el 18% de los genes que presentan la secuencias *ATTGTT* o *AAG[G/C]GG* en sus promotores son regulados potencialmente por unión directa y estable de Ixr1. Esto indica que los efectos indirectos o transitorios basados modelos tipo "*hit and run*" son también importantes en la regulación transcripcional mediada por *IXR1*. También demuestran el bajo grado de especificidad de los dominios HMG-box a la hora de reconocer secuencias de ADN específicas. Incluso entre las interacciones estables detectadas en los promotores en normoxia o hipoxia, sólo una pequeña fracción de entre 20-30% puede considerarse específica de secuencia al integrar los consensos *ATTGTT* o *AAG[G/C]GG*.

Los datos obtenidos, tanto en los estudios de transcriptómica como los de ChIP-on-chip, permiten deducir que el incremento de la resistencia a cisplatino en mutantes *Δixr1* se debe principalmente a una reducción en la disminución que la droga produce sobre la vía de síntesis de ribosomas. Además, la deleción de *IXR1* produce un incremento de la transcripción de genes relacionados con el metabolismo de compuestos de azufre, que puede contribuir a aumentar la disponibilidad de grupos quelantes que inmovilizan el compuesto platinado y que podría incrementar la síntesis de glutatión para favorecer reacciones anti-oxidantes o para la expulsión del cisplatino fuera de la célula a través de la formación de complejos de cisplatino-glutatión. Esto, junto con el enriquecimiento observado en la unión de Ixr1 a regiones promotoras después del tratamiento con cisplatino (del 46.3% al 69.1%), permite concluir que el mecanismo por el que Ixr1 modifica la resistencia a cisplatino tiene un componente que afecta de forma específica a la regulación transcripcional de determinadas rutas metabólicas.

Para el análisis estructural de los dominios de unión a ADN de la proteína Ixr1, se clonaron, purificaron y caracterizaron los dominios HMG-box de Ixr1, tanto

en su disposición nativa en tándem como cada uno por separado. Los análisis mediante dicroísmo circular y resonancia magnética nuclear demostraron una mayor estabilidad del dominio HMG-box A con respecto al dominio HMG-box B. Los análisis de propiedades termodinámicas y de afinidad llevados a cabo mediante fluorescencia polarizada, ensayos de retardo en gel y calorimétricos utilizando distintos tipos de ADN mostraron que, en términos generales, el dominio HMG-box A tiene una mayor afinidad que el dominio HMG-box B, produciendo también mayores ángulos de doblamiento en el ADN. Un aspecto interesante observado es la diferenciación en los signos de entalpia que muestran ambos dominios en función de los tipos de ADN a que se unen. El dominio HMG-box A muestra valores de entalpia muy positivos en su unión a ADN-B-lineal que contiene secuencias específicas presentes en regiones promotoras reguladas por la proteína; mientras que la unión a ADN platinado o ADN cruciforme es termodinámicamente más favorable desde el punto de vista de los valores de entalpía. Sin embargo, el signo de los valores de entalpía no varía en el caso del dominio HMG-box B en su unión a ADN B-lineal o platinado. El estudio de los dos dominios HMG-box en tándem demostró que ambos se unen de manera secuencial con cooperatividad positiva y permitió establecer un modelo de unión de Ixr1 al ADN lineal por el cual la unión del dominio HMG-box A al ADN es necesaria para favorecer que el dominio HMG-box B pueda adoptar una óptima conformación para su posterior unión al ADN.

Finalmente, se clonó, expresó y purificó toda la secuencia aminoacídica que compone Ixr1 en bacterias con rendimiento suficiente para su estudio bioquímico y estructural. Ixr1 mostró una estructura caracterizada por extensas regiones ricas en glutamina desestructuradas dispuestas a ambos lados de los dominios HMG-box en tándem que sí poseen una estructura globular definida. Estudios hidrodinámicos mediante cromatografía de exclusión molecular y ultracentrifugación analítica mostraron el carácter extendido de la proteína, con un radio hidrodinámico 1.5 veces mayor de lo esperado y un ratio en el coeficiente de

fricción (f/f_0) de 1.97. Análisis por proteólisis limitada y posterior identificación de bandas mediante huella peptídica confirmaron estos resultados. Un aspecto muy interesante es su capacidad de formar agregados amiloideos. Se realizaron ensayos *in vitro*, a partir de distintas regiones desordenadas de Ixr1 purificadas, mediante cinéticas monitorizadas con un fluoróforo. La formación de fibras se visualizó mediante microscopía electrónica de transmisión y también en ensayos *in vivo*, mediante la sobreexpresión en *S. cerevisiae* de distintas regiones de Ixr1 fusionadas a GFP. Puesto que la fibra amiloidea es uno de los principales mecanismos de propagación de proteínas priónicas, en este trabajo se realizaron los primeros ensayos fenotípicos para determinar si Ixr1 puede actuar como prion, una característica de enorme importancia y de gran actualidad tanto para comprender aspectos de la evolución y adaptación de las poblaciones de organismos unicelulares ante cambios ambientales, como por su aplicación a diversos aspectos relacionados con enfermedades de origen priónico.

4.- CONCLUSIONES

1. El análisis transcripcional basado en *arrays* revela que Ixr1 participa en la respuesta a hipoxia en levaduras para adaptar el uso y producción de la energía celular a la disponibilidad del oxígeno. Ixr1 regula los niveles de expresión de genes relacionados con la respuesta a estrés oxidativo, la composición de la pared celular, el metabolismo del azufre, los aminoácidos de cadena ramificada y la síntesis de ADN.

2. Ixr1 tiene capacidades reguladoras conmutables y dependientes de la disponibilidad de oxígeno, activando o reprimiendo los mismos genes diana. Los genes de las vías metabólicas de asimilación del azufre y los aminoácidos de cadena ramificada ilustran este mecanismo.

3. Los análisis ChIP-on-chip de los lugares de unión a ADN de Ixr1 en normoxia e hipoxia revelan un bajo enriquecimiento en la unión estable a regiones promotoras, solamente representando el 50-60% del total de

lugares de unión significativos, en función de los niveles de oxígeno intracelulares.

4. Ixr1 y Rox1 comparten secuencias promotoras comunes para su unión a través del *core **ATTGTT*** definido para la unión del dominio HMG-box, o a través de la secuencia consenso ***AAG[G/C]GG*** redefinida en el presente trabajo.

5. La regulación transcripcional mediada por Ixr1 es principalmente atribuible a mecanismos de unión transitoria inestable o indirectos, en lugar de uniones estables directas a los promotores regulados, como se deduce del bajo solapamiento obtenido entre los datos de transcriptómica y de ChIP-on-chip.

6. El análisis transcripcional basado en *arrays* revela que el efecto de Ixr1 en la resistencia a cisplatino se debe principalmente al menor impacto del tratamiento con la droga en los genes relacionados con la vía de síntesis de ribosomas en los mutantes *Δixr1*, así como un mayor aumento de la expresión de los genes del metabolismo de azufre.

7. La función de Ixr1 en la respuesta de las células W303 al tratamiento con cisplatino es también reguladora y no sólo un efecto debido a que enmascare lesiones del DNA, como lo revela el enriquecimiento de las uniones de Ixr1 a regiones promotoras (pasando del 46% al 69%) observado en presencia de la droga

8. Hemos caracterizado bioquímica y termodinámicamente los dominios individuales HMG-box de Ixr1. Las principales conclusiones son las siguientes:

    a. El dominio HMG-box B muestra una menor estabilidad que el dominio HMG-box A.

    b. El dominio HMG-box A muestra una afinidad de unión al ADN un orden de magnitud más alta que el dominio HMG-box B, produciendo ángulos de doblamiento de ADN mayores.

c. Los dos dominios HMG-box se unen al ADN con una reducida especificidad de secuencia en un proceso conducido principalmente por fuerzas entrópicas, mostrando valores de entropía elevados y similares.

d. Los dos dominios HMG-box presentan diferentes características termodinámicas en la unión a ADN modificado por cisplatino, revelando diferencias en el modo de unión.

9. Los dos dominios HMG-box en tándem de Ixr1 se unen al ADN con un efecto cooperativo positivo. La unión del dominio HMG-box A al ADN es necesaria para favorecer que el dominio HMG-box B pueda adoptar una óptima conformación para su posterior unión al ADN.

10. Ixr1 es una proteína intrínsecamente desordenada enriquecida en aminoácidos que promueven el desorden espacial (glutamina, serina y prolina) y muestra 3 regiones a ambos lados de los dominios HMG-box que son intrínsecamente desordenadas, con tendencia a formar hélices alfa.

11. Ixr1 tiene una alta tendencia a la agregación y es capaz de formar estructuras amiloideas *in vitro* e *in vivo*. Además, muestra varias características que abren la posibilidad a que Ixr1 pudiera actuar como un prión en levaduras.

5.- REFERENCIAS BIBLIOGRÁFICAS

Aguzzi, A., and T. O'Conno (2010). "Protein aggregation diseases: pathogenicity and therapeutic perspectivas." Nature Reviews Drug Discovery **9** (3): 237–248.

Bourdineaud, J. P., G. De Sampaio & G.J. Lauquin (2000) "A Rox1- independent hypoxic pathway in yeast. Antagonistic action of the repressor Ord1 and activator Yap1 for hypoxic expression of the SRP1/TIR1 gene" Molecular Microbiology **38**(4):879-90.

Bustin M. (2001) "Revised nomenclature for high mobility group (HMG) chromosomalproteins" Trends in Biochemical Sciences **26** (3) 152–153.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, I. Carneiro, I. González-Siso and M.E. Cerdán (2010a) "Regulatory factors controlling transcription of

Saccharomyces cerevisiae IXR1 by oxygen levels: a model of transcriptional adaptation from aerobiosis to hypoxia implicating ROX1 and IXR1 cross-regulation." Biochemical Journal **425**(1):235-43.

Castro-Prego R., M. Lamas-Maceiras, P. Soengas, R. Fernández-Leiro, I. Carneiro, M. Becerra, M.I. González-Siso and M.E. Cerdán (2010b) "Ixr1p regulates oxygen-dependent HEM13 transcription." FEMS Yeast Research **10**(3):309-21.

Costanzo M., O. Schub and B. Andrews (2003). "G1 transcription factors are differentially regulated in Saccharomyces cerevisiae by the Swi6-binding protein Stb1." Molecular of Cell Biology **23**(14):5064-77.

Diaz-Avalos, R., C. Y. King, J. Wall, M. Simon, and D. L. Caspar (2005). "Strain-specific morphologies of yeast prion amyloid fibrils." Proceedings of the National Academy of Sciences **102**(29): 10165–10170.

Glover, J. R., A. S. Kowal, E. C. Schirmer, M. M. Patino, J. J. Liu and S. Lindquist (1997). "Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of *S. cerevisiae*." Cell **89**(5): 811–819.

Fedorov D.V., S.V. Koval'tsova, V.T. Peshekhonov, V.G. Korolev (2010). "[*IXR1* and *HMO1* genes jointly control the level of spontaneous mutagenesis in yeast Saccharomyces cerevisiae]." Genetika **46**(6):750-7.

Fox M.E., B.J. Feldman, G. Chu (1994). "A novel role for DNA photolyase: binding to DNA damaged by drugs is associated with enhanced cytotoxicity in *Saccharomyces cerevisiae*." Molecular Cell Biology **14**(12):8071-7.

Hardman C.H., R.W. Broadhurst, A.R. Raine, K.D. Grasser, J.O. Thomas, E.D. Laue (1995). "Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy." Biochemistry **34**(51):16596-607.

Huang R.Y., M. Eddy, M. Vujcic, D. Kowalski (2005). "Genome-wide screen identifies genes whose inactivation confer resistance to cisplatin in *Saccharomyces cerevisiae*." Cancer Research **65**(13):5890-7.

Koch C., T. Moll, M. Neuberg, H. Ahorn, and K. Nasmyth (1993). "A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase." Science **261**(5128):1551–1557.

Lambert JR, V.W. Bilanchone & M.G. Cumsky (1994) "The ORD1 gene encodes a transcription factor involved in oxygen regulation and is identical to IXR1, a gene that confers cisplatin sensitivity to Saccharomyces cerevisiae" Proceedings of the National Academy of Sciences of the United States of America **91**(15):7345-9.

Liebman S.W., Y.O. Chernoff (2012). "Prions in yeast." Genetics **191**(4):1041-72.

McA'Nulty M.M. and S.J. Lippard (1996). "The HMG-domain protein Ixr1 blocks excision repair of cisplatin-DNA adducts in yeast." Mutation Research **362**(1):75-86.

Nagatani G., M. Nomoto, H. Takano, T. Ise, K. Kato, T. Imamura, H. Izumi, K. Makishima, K. Kohno (2001). "Transcriptional activation of the human *HMG1* gene in cisplatin-resistant human cancer cells." Cancer Research **61**(4):1592-7.

Read C.M., P.D. Cary, C. Crane-Robinson, P.C. Driscoll, D.G. Norman (1993). "Solution structure of a DNA-binding domain from HMG1." Nucleic Acids Research

21(15):3427-36.

Schenk P.W., A.W. Boersma, J.A. Brandsma, H. den Dulk, H. Burger, G. Stoter, J. Brouwer, K. Nooter (2001). "*SKY1* is involved in cisplatin-induced cell kill in *Saccharomyces cerevisiae*, and inactivation of its human homologue, SRPK1, induces cisplatin resistance in a human ovarian carcinoma cell line." Cancer Research **61**(19):6982-6.

Schenk P.W., M. Brok, A.W. Boersma, J.A. Brandsma, H. den Dulk, H. Burger, G. Stoter, J. Brouwer, K. Nooter (2003). "Anticancer drug resistance induced by disruption of the *Saccharomyces cerevisiae NPR2* gene: a novel component involved in cisplatin- and doxorubicin-provoked cell kill." Molecular Pharmacology **64**(2):259-68.

Sidorova JM, L.L. Breeden (1997) "Rad53-dependent phosphorylation of Swi6 and down-regulation of CLN1 and CLN2 transcription occur in response to DNA damage in *Saccharomyces cerevisiae*". Genes Development **11**(22):3032-45.

Stillman DJ. (2013). "Dancing the cell cycle two-step: regulation of yeast G1- cell-cycle genes by chromatin structure." Trends in Biochemical Sciences **38**(9):467-75.

Stott K., G.S. Tang, K.B. Lee and J.O. Thomas (2006). "Structure of a complex of tandem HMG-boxes and DNA." Journal of Molecular Biology **360**(1):90-104.

Stros M. (1998). "DNA bending by the chromosomal protein HMG1 and its high mobility group box domains. Effect of flanking sequences." Journal of Biological Chemistry **273**(17):10355-61.

Stros M. (2010). "HMGB proteins: interactions with DNA and chromatin." Biochimica et Biophysical Acta. **1799**(1-2):101-13.

Taudte S., H. Xin, A.J. Jr. Bell, N.R. Kallenbach (2001). "Interactions between HMG boxes." Protein Engineering **14**(12):1015-23.

Thomas J.O. and A.A. Travers (2001). "HMG1 and 2, and related 'architectural' DNA-binding proteins." Trends in Biochemical Science **26**(3):167–174.

Travers A. (2000). "Recognition of distorted DNA structures by HMG domains." Current Opinion in Structural Biology **10**(1):102-9.

Tsaponina O., E. Barsoum, S.U. Aström and A. Chabes (2011). "Ixr1 is required for the expression of the ribonucleotide reductase Rnr1 and maintenance of dNTP pools." PLoS Genetics **7**(5):e1002061.

Tyedmers, J., M.L. Madariaga and S. Lindquist (2008). "Prion switching in response to environmental stress." PLoS Biology **6**(11): e294.

Verma R, J. Smiley, B. Andrews, J.L. Campbell (1992) "Regulation of the yeast DNAreplication genes through the Mlu I cell cycle box is dependent on SWI6" Proceedings of the National Academy of Sciences U S A. **89**(20):9479- 83.

Weir H. M., P.J. Kraulis, C.S. Hill, A.R. Raine, E.D. Laue and J.O. Thomas (1993). "Structure of the HMG-box motif in the B-domain of HMG1." EMBO Journal **12**(4):1311-9.

Werner M.H., J.R. Huth, A.M. Gronenborn and G.M. Clore (1995). "Molecular basis of human 46X,Y sex reversal revealed from the three- dimensional solution structure of the human SRY-DNA complex." Cell **81**(5):705-14.

Appendix II

Curriculum vitae

Publications

2015

Barreiro-Alonso A., M. Lamas-Maceiras, E. Rodríguez-Belmonte, A. Vizoso-Vázquez, M. Quindós and M.E. Cerdán. (2015) "High Mobility Group B Proteins, Their Partners, and Other Redox Sensors in Ovarian and Prostate Cancer." Oxidative Medicine and Cellular Longevity. Article ID 403282, in press.

González-Siso M.I., A. Touriño, A. Vizoso, A. Pereira-Rodríguez, E. Rodríguez-Belmonte, M. Becerra, M.E. Cerdán. (2015) "Improved bioethanol production in an engineered *Kluyveromyces lactis* strain shifted from respiratory to fermentative metabolism by deletion of *NDI1*." Microbial Biotechnology **8**(2):319-30.

2014

Rodríguez-Lombardero S., A. Vizoso-Vázquez, L.J. Lombardía, M. Becerra, M.I. González-Siso, M.E. Cerdán. (2014) "Sky1 regulates the expression of sulfur metabolism genes in response to cisplatin." Microbiology **160**(7):1357-68.

Rodríguez-Lombardero S., M.E. Rodríguez-Belmonte, M.I. González-Siso, A. Vizoso-Vázquez, V. Valdiglesias, B. Laffón, M.E. Cerdán. (2014). "Proteomic analyses reveal that Sky1 modulates apoptosis and mitophagy in *Saccharomyces cerevisiae* cells exposed to cisplatin." International Journal of Molecular Sciences **15**(7):12573-90.

Rico-Díaz A., A. Vizoso Vázquez, M.E. Cerdán, M. Becerra, J. Sanz-Aparicio. (2014). "Crystallization and preliminary X-ray diffraction data of β-galactosidase from *Aspergillus niger*." Acta Crystallographica section F **70**(11):1529-31.

2012

Vizoso Vázquez A., M. Blanco, J. Zaborowska, P. Soengas, M.I. González-Siso, M. Becerra, E. Rodríguez-Belmonte, M.E. Cerdán. (2012) "Two proteins with different functions are derived from the *KlHEM13* gene." Eukaryotic Cell **10**(10):1331-9.

Vizoso-Vázquez A., M. Lamas-Maceiras, M. Becerra, M.I. González-Siso, E. Rodríguez-Belmonte, M.E. Cerdán (2012) "Ixr1p and the control of the *Saccharomyces cerevisiae* hypoxic response" Applied Microbiology and Biotechnology **94**(1):173-84.

Rodríguez Lombardero S., A. Vizoso Vázquez, E. Rodríguez Belmonte, M.I. González Siso and M.E. Cerdán (2012). "*SKY1* and *IXR1* interactions, their effects on cisplatin and spermine resistance in *Saccharomyces cerevisiae*." Canadian Journal Microbiology **58**(2):184-8.

García Leiro A., S. Rodríguez Lombardero, A. Vizoso Vázquez, M.I. González Siso and M.E. Cerdán. (2012) "The yeasts genes *ROX1*, *IXR1*, *SKY1* and their effect upon enzymatic activities related to oxidative stress." OXIDATIVE STRESS ISBN: 979-953-307-574-6

Saheb Shaik K., F. Meyer, A. Vizoso-Vazquez, M. Flötenmeyer, M.E. Cerdán, B. Moussian. (2012) "Delta- aminolevulinate synthase is required for apical transcellular barrier formation in the skin of the Drosophila larva." European Journal of Cell Biology **91**(3):204-15.

González Siso M.I., M. Becerra, M. Lamas Maceiras, A. Vizoso Vázquez, M.E. Cerdán. (2012) "The yeast hypoxic responses, resources for new biotechnological opportunities." Biotechnology Letters **34**(12):2161-73.

Conference communications

2014

Rico A., A. Vizoso-Vazquez, M.E. Cerdán, M. Becerra, Sanz-Aparicio J "Purificación y cristalización de la β-galactosidasa de *Aspergillus niger*" Poster communication to the *XXIV* Symposium of the Crystallography and Crystal Growth Specialized Group (GE3C), "Crystallography and Sustainability". (23-26 June 2014) (Bilbao, Spain).

Vizoso-Vazquez A., M. Lamas-Maceiras., R. Fernández Leiro, M.E. Cerdán. "Molecular characterization of ixr1 in terms of dna binding through their hmg-box domains" Poster communication to The European Molecular Biology Organization (EMBO) Conference 2014 : "Gene Transcription in Yeast : From regulatory networks to mechanisms". (14-19 June 2014) (Sant Feliu de Guixols, Spain).

2013

Vizoso-Vazquez A., M. Lamas-Maceiras., R. Fernández Leiro, M.E. Cerdán. "Caracterización de la unión de los dominios HMG-box de Ixr1p a la región promotora de ROX1" Poster communication to the *XXXVI* Spanish Society of Biochemistry and Molecular Biology Conference. (03-06 September 2013) (Madrid, Spain).

2012

Vizoso Vázquez A., S. Rodríguez Lombardero, M. Becerra, L.J. Lombardía, M.E. Cerdán Villanueva "Transcriptome analysis of cisplatin treatment in yeast and the effect of *Δixr1* and *Δsky1Δixr1* mutations" Poster communication to the 22nd Congress The International Union of Biochemistry and Molecular Biology (IUBMB) and 37th Federation of European Biochemical Societies (FEBS). (04-09 September 2012) (Sevilla, Spain).

2011

Vizoso-Vazquez A., M.I. González-Siso, M. Becerra Fernández, E. Rodríguez-Belmonte, M.E. Cerdán "Klhem13p y *moonlighting* en *Kluyveromyces lactis*" Poster communication to the *XXXIV* Spanish Society of Biochemistry and Molecular Biology Conference. (05-08 September 2009) (Barcelona, Spain).

Rico A., A. Vizoso-Vazquez, M.E. Cerdán "Regulación del gen *KlHEM13* por Klixr1p en *Kluyveromyces lactis*" Poster communication to the *XXXIV* Spanish Society of Biochemistry and Molecular Biology Conference. (05-08 September 2009) (Barcelona, Spain).

2010

Vizoso-Vazquez A., M. Lamas, M.E. Cerdán "Mechanism of cross-regulation between Rox1p and Ixr1p" Oral communication to the Yeast, an evergreen model – Tribute to P. Slonimski. (23-25 September 2010) (Rome, Italy).

Vizoso-Vazquez A., M. Lamas, M.E. Cerdán "Papel regulador de Ixr1p en la respuesta a hipoxia en *Saccharomyces cerevisiae*" Poster communication to the *XXXIII* Spanish Society of Biochemistry and Molecular Biology Conference. (14-17 September 2009) (Cordoba, Spain).

2009

Vizoso-Vazquez A., M.I. González-Siso, M. Becerra Fernández, M.E. Cerdán "Interacciones de Ixr1p con otras proteínas implicadas en la regulación transcripcional" Poster communication to the *XXXII* Spanish Society of Biochemistry and Molecular Biology Conference. (23-26 September 2009) (Oviedo, Spain).

2008

Vizoso-Vázquez A., M. Blanco, J. Zaborowska, P. Soengas, M.I. González-Siso, M.E. Cerdán "El papel de KlCpo como regulador transcripcional en relación con la utilización de sitios alternativos para la transcripción y traducción del gen *KlHEM13*" Poster communication to the *XXXI* Spanish Society of Biochemistry and Molecular Biology Conference. (10-14 September 2008) (Bilbao, Spain).

Research Fellowships / Contracts

2015

Predoctoral Contract (University of A Coruña) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (15 January – 15 May 2015) (A Coruña, Spain).

2014

Research stay fellowship (University of A Coruña - INDITEX) Biochemistry Department of University of Cambridge. (03 September 2014 – 21 December 2013)

2013

Research stay fellowship (University of A Coruña) Crystallography and Structural Biology area. Physics-Chemistry Institute "Rocasolano" (IQFR). (01 September 2013 – 19 December 2013)

Travel fellowship to the 22nd Congress The International Union of Biochemistry and Molecular Biology (IUBMB) and 37th Federation of European Biochemical Societies (FEBS). (04-09 September 2012) (Sevilla, Spain).

2011

Predoctoral fellowship (Xunta de Galicia) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (01 November 2011 – 31 October 2014) (A Coruña, Spain).

Predoctoral Contract (University of A Coruña) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (01 September – 31 October 2011) (A Coruña, Spain).

Travel fellowship to the *XXXIV* Spanish Society of Biochemistry and Molecular Biology Conference (SEBBM). (05-08 September 2009) (Barcelona, Spain).

2009

Graduate fellowship "Lucas Labrada program" (Xunta de Galicia) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (01 January 2009 – 31 December 2010) (A Coruña, Spain).

Travel fellowship to the *XXXII* Spanish Society of Biochemistry and Molecular Biology Conference (SEBBM). (23-26 September 2009) (Oviedo, Spain).

2008

Graduate fellowship (University of A Coruña) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (01 November – 31 December 2008) (A Coruña, Spain).

Predoctoral Contract (University of A Coruña) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (01 September – 31 October 2008) (A Coruña, Spain).

Travel fellowship to the *XXXI* Spanish Society of Biochemistry and Molecular Biology Conference (SEBBM). (10-14 September 2008) (Bilbao, Spain).

Undergraduate fellowship (City council of A Coruña) Aquarium Finisterrae (September 2007 – June 2008) (A Coruña, Spain).

Undergraduate fellowship (Spanish Ministry of Education and Science) Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña (September 2007 – June 2008) (A Coruña, Spain).

2007

Undergraduate fellowship (City council of A Coruña) Aquarium Finisterrae (September 2006 – June 2007) (A Coruña, Spain).

Participation in Research Grants / Projects

2009

Regulation mediated by Ixr1p and two homologous human proteins in *Saccharomyces cerevisiae*.
Funded by the Spanish Ministry of Science and Technology (BFU2009-08854).
Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña
2008-2014
Principal Investigator: María Esperanza Cerdán Villanueva

Regulation mechanisms related to oxidative stress and low oxygen availability in yeast.
Funded by the Spanish Ministry of Science and Education (BFU2006-03961)

Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña

2006-2009

Principal Investigator: María Esperanza Cerdán Villanueva

2008

Cheese whey valorization by producing functional food ingredients and production of bioethanol

Funded by the Xunta de Galicia (PGIDIT 07TAL006E)

Biochemistry area. Department of Cellular and Molecular Biology. University of A Coruña

2007-2008

Principal Investigator: Manuel Becerra Fernández

Research Stays

2014

Predoctoral stay during four months in the Department of Biochemistry (University of Cambridge) under the supervision of Dra. Jean O. Thomas (03 September 2014 – 21 December 2014) (Cambridge, United Kingdom)

2013

Predoctoral stay during four months in the Physics-Chemistry Institute "Rocasolano" (IQFR) under the supervision of Dra. Juliana Sanz Aparicio (01 September 2013 – 19 December 2013) (Madrid, Spain)

2011

Short stay of one week in the National Center of Oncological Research (CNIO) under the supervision of Dr. Jose Luis Lombardia (25 June 2011 – 1 July 2011) (Madrid, Spain)

Other merits

Official Master of Bioinformatic certified by the International University of Andalucia (UNIA) (September 2008 – January 2011) (Sevilla, Spain).

Minor thesis entitled "ScIxr1p y ScRox1p, dos proteínas con dominios HMG-box en la respuesta a los niveles de oxígeno intracelular" and certified by the University of A Coruña (February 2011) (A Coruña, Spain).

"Biotechnology and dairy industry: waste recovery" seminar organized by the University of A Coruña (2011) (A Coruña, Spain).

DEA (Advanced Studies Diploma) certified by the University of A Coruña (September 2008 – June 2010) (A Coruña, Spain).

"2nd International School on Biological Crystallization" course organized by the IACT, CSIC-UGRA (May 2009) (Granada, Spain).

2nd Bioinformatic Galician seminar organized by the University of Santiago de Compostela and the Bioinformatic Galician Network (2009) (Santiago de Compostela, Spain).

"Structural Bioinformatic" course organized by the University of Santiago de Compostela and the Bioinformatic Galician Network (2009) (Santiago de Compostela, Spain).

"1st research session of RNA" seminar organized by the University of A Coruña (2009) (A Coruña, Spain).

English level title B1 certified by the Official School of Languages of A Coruña (29 July 2008) (A Coruña, Spain).

"Gene expression analysis and it s applications" course organized by the University of A Coruña (2008) (A Coruña, Spain).

"UPLC: Instrumentation and applications" seminar organized by Waters Chromatography company (2008) (Santiago de Compostela, Spain).

 "Initiation to Biochemistry and Molecular Biology Research" course organized by the Spanish Society of Biochemistry and Molecular Biology Conference (SEBBM) (12 September 2007) (Malaga, Spain).

Job training in "AMBICAL projects" during 300 hours (19 July 2007 – 21 September 2007) (A Coruña, Spain).