Universidade de Aveiro Departamento de Ciências Médicas
Ano 2020

BÁRBARA FILIPA TEIXEIRA DE ARAÚJO MIRANDA

IDENTIFICATION OF REAGENTS INTENDED TO THE STUDY OF THE LONG ISOFORM OF THE NF-YA TRANSCRIPTION FACTOR

IDENTIFICAÇÃO DE REAGENTES DESTINADOS AO ESTUDO DA ISOFORMA LONGA DO FACTOR DE TRANSCRIÇÃO NF-YA

Universidade de Aveiro Departamento de Ciências Médicas
Ano 2020

BÁRBARA FILIPA TEIXEIRA DE ARAÚJO MIRANDA

IDENTIFICATION OF REAGENTS INTENDED TO THE STUDY OF THE LONG ISOFORM OF THE NF-YA TRANSCRIPTION FACTOR

IDENTIFICAÇÃO DE REAGENTES DESTINADOS AO ESTUDO DA ISOFORMA LONGA DO FACTOR DE TRANSCRIÇÃO NF-YA

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biomedicina Molecular, realizada sob a orientação científica do Doutor Bruno Miguel Bernardes de Jesus, Professor Auxiliar do Departamento de Ciências Médicas da Universidade de Aveiro e co-orientação do Doutor Roberto Mantovani da Universidade de Milão.

Dedico este trabalho à minha mãe por todo o esforço e apoio ao longo de toda a minha vida.

o júri

Presidente	Professora Doutora Catarina Rodrigues de Almeida Professora Auxiliar em Regime Laboral, Universidade de Aveiro
Vogal – Orientador	Professor Doutor Bruno Miguel Bernardes de Jesus Professor Auxiliar em Regime Laboral, Universidade de Aveiro
Vogal – Arguente Principal	Doutora Manuela Ferreira Investigadora Doutorada, Universidade de Coimbra

agradecimentos

Obrigada à minha família por me ajudar a realizar os meus sonhos, sempre acreditarem em mim e me apoiarem.

Obrigada ao professor Dr. Roberto por me aceitar no laboratório e me confiar neste projeto cativante.

Obrigada à professora Dr. Nerina por me ajudar e apoiar durante todo o processo.

Um especial obrigada à Dana e a todos os membros do laboratório por me integrarem e ajudarem durante o percurso.

Obrigada ao Dr. Bruno Jesus por aceitar fazer parte deste projeto tão prontamente.

Obrigada aos meus amigos italianos, Gianluca, Riccardo e Vittorio, por tornarem esta jornada mais divertida.

Obrigada às minhas amigas do secundário por crescerem comigo e me apoiarem em todos os momentos.

E obrigada aos meus amigos “aquáticos” por sempre acreditarem em mim e me ajudarem sempre que precisei.

acknowledgments

Thank you to my family for helping me make my dreams come true and always believe in and support me.

Thank you to Professor Roberto for accepting me in the lab and trust me with this captivating project.

Thank you to Professora Nerina for helping and supporting me in the process.

Special thanks to Dana and all the members of the Mantovani lab for welcoming me and helping me throughout the way.

Thank you to Dr. Bruno Jesus for accepting to be part of this project so promptly.

Thank you for my Italian friends Gianluca, Riccardo and Vittorio for making this journey funnier.

Thank you to my high school friends for growing up with me and supporting me at all times.

And thank you to my “aquatic” friends for believe in me and supporting me every time that I needed.

palavras chave

Fator de Transcrição, Fator Nuclear Y, NF-YA, *Splicing* Alternativo, NF-YA longo.

resumo

Fatores de transcrição (FT) são proteínas envolvidas e vitais em diversos processos celulares e na própria identidade celular. A ligação destas proteínas a elementos específicos de ADN é o que faz com que sejam tão essenciais, regulando o transcriptoma de todas as células. O fator nuclear Y é uma proteína classificada como um fator de transcrição pioneiro, ligando-se à sequência CCAAT do promotor. Este FT é composto por três diferentes e preservadas subunidades NF-YA, NF-YB e NF-YC. O NF-YA é a subunidade reguladora, composta por 347 aminoácidos na sua isoforma longa (NF-YAl). A expressão genética do NF-YA gera, por *splicing* alternativo, duas fundamentais isoformas, que diferem em 28 aminoácidos, correspondentes ao exão 3.

Devido à função essencial que este fator nuclear desempenha em vários processos celulares é crucial o alargamento do conhecimento sobre as suas funções fisiológicas e patológicas, em particular no que diz respeito à expressão diferencial das diferentes isoformas do NF-YA. Neste projeto, que inicialmente começou após dois coelhos serem imunizados com NF-YAl, desenvolveram-se ensaios para obter reagentes de anticorpos específicos para isoforma longa de NF-YA, essenciais para distinção de ambas as isoformas em experiências futuras. O objetivo deste projeto inclui a produção e purificação de proteínas recombinantes para a purificação de afinidade de anticorpos específicos.

De modo a alcançar o objetivo principal, o projeto começou com a preparação de células bacterianas competentes para transformação com os plasmídeos de expressão relevante, seguido pela configuração dos protocolos de indução e purificação para a produção proteica. Após a aquisição da quantidade suficiente de proteína, procedemos às etapas finais do projeto, com ensaios de *western blot* e cromatografia de afinidade do sera de coelho. As soluções de eluições obtidas pelas purificações de cromatografia de afinidade foram, posteriormente, testadas e analisadas através de experimentos de *western blot*, de modo a verificar se os anticorpos específicos para NF-YAl eram, de facto, específicos.

O principal objetivo do projeto foi atingido com sucesso, alcançando como resultado final positivo duas eluições correspondentes a um pH específico da cromatografia de afinidade, contendo anticorpos específicos para a isoforma longa de NF-YA.

keywords

Transcription Factor, Nuclear Factor Y, NF-YA, Alternative Splicing, NF-YA long.

abstract

Transcription factors are proteins involved and vital to several cellular processes and cell identity. The binding of these proteins to the specific DNA elements is what makes it so important, regulating the transcriptome of all cells. The nuclear factor Y is a protein classified as a pioneer TF, binding to the CCAAT box promoter element. This NF is composed by three different and preserved subunits NF-YA, NF-YB and NF-YC. NF-YA is the regulatory subunit and is composed by 347 aminoacids in its long isoform (NF-YAl). The NF-YA gene expression generates in fact, by alternative splicing, two main isoforms, differing in 28 aminoacids, corresponding to the exon 3.

Due to the essential role this nuclear factor plays in many cellular processes it is crucial the deepening of the knowledge about its physiological and pathological functions, in particular regarding the differential expression of the different isoforms of NF-YA. In this project, which initially started after two rabbits were immunized with NF-YAl, we performed experiments to obtain NF-YA long specific antibody reagents, essential to distinguish both isoforms in further experiments. This project aims included the production and purification of recombinant proteins for the subsequent affinity purification of the specific antibodies.

To accomplish the main objective, the project started with the preparation of bacterial competent cells for transformation with the relevant expression plasmids, followed by setting the induction and purification protocols for protein production. After the required amount of protein was obtained, we proceeded to the final steps of the project with experiments of western blot and affinity chromatography of rabbit sera. The elution fractions obtained by affinity chromatography purifications were subsequently tested and

analysed by western blot, to verify the specificity of NF-YAl antibodies.

The principal aim was achieved with success, having as a final positive result in two elution fractions, corresponding to a specific pH of the affinity chromatography, which included antibodies specific to the long isoform of NF-YA.

1 INDEX

2 LIST OF FIGURES

3 LIST OF TABLES

4 ABREVIATIONS

αHis	Anti-Histidine Tagged Antibody;
aa	Aminoacids;
Ab	Antibody;
AB	Acetate Buffer;
Amp	Ampicillin;
APS	Ammonium Persulfate;
Arg	Arginine;
AS	Alternative Splicing;
Asp	Aspartic acid;
BA	Buffer A;
BA2	Buffer A2;
BA100	Buffer A100;
BA250	Buffer A250;
BMP2	Bone Morphogenetic Protein 2;
BSA	Bovine Serum Albumin;
CB	Coupling Buffer
CBF	CCAAT Binding Factor;
CDK2	Cyclin-dependent Kinase 2;
Cm	Chloramphenicol;
CREs	Cis-Regulatory Elements;
CV	Column Volume;
DBD	DNA Binding Domain;
dH_2O	Distilled water;
DNA	DeoxyriboNucleic Acid;
DSPP	Dentin sialophosphoprotein;

EB	Elution Buffers;
ESC	Embryonic Stem Cells;
EtOH	Ethanol;
Gln	Glutamine;
GTF	General Transcription Factors;
HAP	Heme-activated Protein;
HCl	Hydrochloric acid;
HFD	Histone-fold Domain;
His-Tag	PolyHistidine-Tag;
HSM	Hematopoietic Stem Cells;
Htt	Huntington;
$MgCl_2$	Magnesium chloride;
mRNA	Messenger RNA;
NaCl	Sodium chloride;
NaP	Sodium phosphide;
NF	Nuclear Factor;
NF-Y	Nuclear Factor Y;
NF-YA	Nuclear Factor Y, subunit A;
NF-YB	Nuclear Factor Y, subunit B;
NF-YC	Nuclear Factor Y, subunit C;
NF-YAl	Nuclear Factor Y, subunit A, long isoform;
NF-YAs	Nuclear Factor Y, subunit A, short isoform;
OD	Optical Density;
OD600	Optical Density at 600 nm;
ON	Overnight;
PBS	Phosphate-Buffered Saline;

Pre-mRNA	Messenger RNA precursor;
RT	Room Temperature;
SB 2x	Sample Buffer 2x;
SDS	Sodium Dodecyl Sulphate;
ST	Supernatant;
tAD	Trans-Activation Domain;
TBP	TATA Binding Protein;
TEMED	Tetramethylethylenediamine;
TF	Transcription Factor;
TFBSs	Transcription Factor Binding Sites;
TG	Tris-Glycine;
TGF-β	Transforming Growth Factor-beta;
Trp	Tryptophan;
TSS	Transcription Start Site;
WB	Washing Buffer.

CHAPTER 1

INTRODUCTION

5 INTRODUCTION

5.1 TRANSCRIPTION FACTORS

Transcription factors (TFs) are the main elements in the regulation of gene transcription, being proteins responsible to provide and/or even preserve the cellular identity, through the direct interpretation of the genome, operating the initial step in DNA sequence decoding, in control regions of genes (Lambert et al., 2018; Mitsis et al., 2020). Although not always was considered this way, TFs are used to describe exclusively proteins able to bind DNA in a sequence-specific mode and to regulate the transcription (Fulton et al., 2009; Lambert et al., 2018; Vaquerizas et al., 2009). The precise and often synergistic sequence-specific binding of TFs to, in the regulatory regions - the promoters (section of DNA that controls the initiation of RNA transcription, positioned in proximity to the genes) and enhancers (short region of DNA that can increase transcription of genes, acting at distance) - of genomes is what determines the transcriptome of every and each living cell (R Mantovani, 1998).

TFs do this regulation mostly through the binding to DNA elements, such as motifs or *cis*-regulatory elements (CREs), that are specific short DNA base pair patterns. Besides that, they are also able to interact in different genomic locations that may be distant to the primary DNA sequence, that are described as the gene regulatory regions. CREs include promoters and sequences, those are termed enhancers when they lead to transcriptional activation and silencers when in transcriptional repression. DNA binding domain (DBD) corresponds to the specific domain in TFs where DNA can bind (Kolomeisky, 2011; Mitsis et al., 2020). Besides to DBD, which allows sequence specific binding of DNA regulatory elements previously mentioned, TFs also include separate transactivating region, the tAD (trans-activation domain), that also interact with downstream factors and coactivators for transcriptional regulation (Arnold et al., 2018).

It is known that the binding is done in order to regulate their transcription, whether positive or negatively (Latchman, 1997; Ly et al., 2013), that is vital in processes promoting development, differentiation and tissue maintenance (Boyer et al., 2005; Dolfini, Minuzzo, et al., 2012) and in which the specific expression is transmitted through cell division (Oestreich & Weinmann, 2012). Upon binding activity cofactors will be recruited, with histone modifications and chromatin remodelling activities, that embody multiple binding sites for distinct transcription factors, that posteriorly will convene in a specific and exceptional way, forming a three-dimensional architecture (Dolfini et al., 2020; Nardone et al., 2017). The coordinated activity of the transcription factors is essential during development, for the appropriate spatial-temporal expression of the genes, as for example during myogenesis (Zinzen et al., 2009).

The TF-binding sites are usually small and flexible which indicates that a typical human gene encompasses numerous possible binding sites for most TFs (Lambert et al., 2018). TFs can contain either one or more than one DNA-binding domains, that bind to the specific DNA sequences, while the other regions of the protein are responsible for the stimulatory or inhibitory factors on transcription (Mitchell & Tjian, 1989).

TFs used to be categorized as either activators or repressors, however is not the most accurate categorization to be made, since it has been described TFs that act as activator in a certain type of cell and as a repressor in a different one (Lambert et al., 2018). Many of these proteins act out as "master regulators" and as "selector genes" controlling processes that led to the specify of certain typed of cells and also commanding specific pathways, as for example immune responses (Lambert et al., 2018). TFs can be distinguished in two different functional categories, they can either be considered pioneers or activating TFs (Dolfini et al., 2020). Pioneers TFs are the ones that have present an intrinsic capacity to chromatin association, through binding to the corresponding motifs directly on a nucleosome (Cirillo et al., 2002; Zaret & Mango, 2016), on the other hand, activating TFs, when it binds to sites in a context predisposed by enzymatic machines, affecting the chromatin accessibility (Dolfini et al., 2020; Zambelli & Pavesi, 2017). Besides these general categories, TFs are also divided into different families, according to their recognizable conserved domains and homology regions. Transcription factor families are often composed by different members encoded by different genes, and further expansion is generated by differential splicing events which have allowed specialised functions, which enable precise gene regulation during development and in different cell types. The TFs belonging to a specific family may present similar protein sequences, which implies that the TF family at stake may have similar or overlapping biological functions in different organisms, or tissues. This presents as an advantage to analyse and identify unknown proteins in other different species (M. Li et al., 2019).

Studies involving eukaryotic and bacterial TFs, have shown that the formation of multiprotein complexes and the interaction with DNA of the TFs can be versatile in a substantial way (Nardone et al., 2017). GTFs distinguish from the "common" TFs by the fact that whereas the first ones bind to core promoters that are situated near to the transcription start site (TSS) and assemble in a specific manner with RNA polymerase, in order to form a functional pre-initiation complex, the second ones recognize and bind to short regulatory DNA sequences in promoters and enhancers to regulate chromatin accessibility to GTFs, and RNA polymerase activity (Thomas & Chiang, 2006).

It is important to understand that transcription factors are important and essential to all living organisms, in many cellular processes, playing different and vital roles in the regulation of different physiological pathways. Subsequently, mutations in TFs and the respective binding sites may lead to major issues, such as human diseases including developmental disorders, disorders of hormone responses and cancer, among others (Lambert et al., 2018; Latchman, 1996).

The regulatory pathways under the responsibility of TFs can be dynamic in the same organism, since the same TF can control different genes, when located in distinct cell types (Lambert et al., 2018). Many of the transcription factors are in fact crucial to the control of the cell growth, thus it is clear how, in case of alterations in their structure or expression, this might lead to tumorigenesis (Dolfini et al., 2019). The better understanding on how TFs bind to a specific DNA sequence and how the gene control is made are imperative to the development of knowledge on a large number of common disease-associated genetic variations (Lambert et al., 2018).

5.2 NF-Y

The Nuclear Factor Y alongside the element that it binds to, the CCAAT box (figure 1) – a major regulatory element (Dolfini et al., 2019) –, were the first cis-elements and trans-acting factors to be identified (Dolfini, Gatta, et al., 2012). The binding of NF-Y with the CCAAT box occurs with a high specificity in the promoter regions of the genes that are transcribed by the RNA polymerase II (X. Y. Li et al., 1992). This interaction is mandatory for the transcriptional activation or repression of numerous eukaryotic genes (Dolfini, Gatta, et al., 2012; M. Li et al., 2019). Despite the specific mechanisms that make NF-Y control the gene expression remain unclear, it is well known that this nuclear factor has a specific role in gene regulation (Oldfield et al., 2019). Systematic in-vivo analyses performed by the ENCODE consortium project – an encyclopaedia of DNA elements – show that 25% of this TF sites are located in promoters, and a similar amount are established at tissue specific enhancers (Zambelli & Pavesi, 2017).

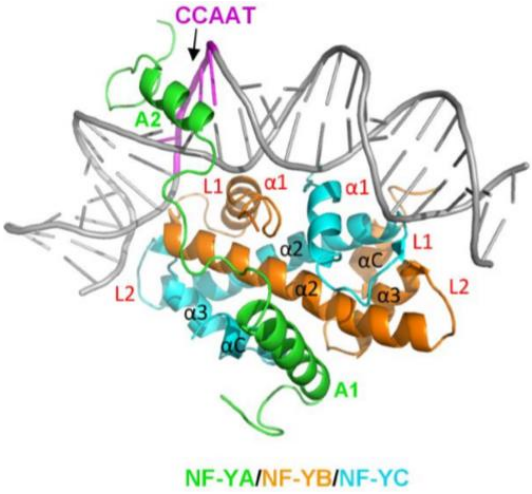

Figure 1. Three-dimensional representation of the NF-Y trimer and the binded CCAAT box, with respective secondary elements (Nardone et al., 2017).

NF-Y has three evolutionarily preserved (from yeast to human (Oldfield et al., 2014)) different subunits – NF-YA, NF-YB and NF-YC – (Ly et al., 2013; Roberto Mantovani, 1999; Nardone et al., 2017), encoded by independent genes , which are pivotal for DNA binding (Ceribelli et al., 2009; Nardone et al., 2017). It is also responsible for performing a major role in the differentiation and proliferation of cells (Bezzecchi et al., 2019), regulating the majority of cell cycle and promoters of growth-promoting genes (Bernardini et al., 2019; Dolfini et al., 2020; Oldfield et al., 2019).

The NF-Y regulatory function can be facilitated by its own capability of interaction with or even recruit various pathway-specific promoter-binding factors (Sankar N Maity, 2017). Different results and researches suggest that this TF plays a unique role in the transcriptional activation, also linking NF-Y to cellular transformation (Bernardini et al., 2019; Dolfini et al., 2020; Oldfield et al., 2019).

It is now general knowledge that this trimeric transcription factor, also called as heme-activated protein (HAP) or CCAAT binding factor (CBF), is categorized as a pioneer transcription factor in both mammals and plants (Dolfini et al., 2020; Libetti et al., 2020) having a corresponding task, consisting in setting the chromatin stage in order to recruit other transcription factors and coactivators (Dolfini et al., 2019; Oldfield et al., 2014). Besides that, it is also considered an ubiquitous TF, since it is not known any cell line that lacks this TF (Dolfini et al., 2020), and its levels and DNA-binding activity differ in little portions in distinct cell lines (Bezzecchi et al., 2019). This nuclear factor is also important for the activation of genes present in the biosynthesis of lipids, glycolysis, such as in the production of determined aminoacids (Dolfini et al., 2020).

Surprisingly, it has been shown that the scenario of the *in vivo* bindings of NF-Y is more complex than what was anticipated, due to two different reasons. First, owing to the fact that the biggest part of the NF-Y binding sites does not correspond to the acknowledged promoter regions (Ceribelli et al., 2009). Secondly, NF-Y is also bound to DNA regions which comprehend negative histone marks, as it is the case with the epigenetic H4K20me3 and H3K27me3 marks. Nevertheless, this nuclear factor behaves as a bi-functional TF, since it activates or represses the transcription depending on the cellular and chromatic contexts (Ceribelli et al., 2009).

5.3 THE CCAAT BOX ELEMENT

It is of an extreme relevance to understand the DNA sequence (frequently mentioned as motifs) that a specific TF binds to, in order to comprehend the TF itself (Lambert et al., 2018). CCAAT box is an important and well described cis-acting element – molecular switch that takes part in the transcriptional regulation of genes activities monitoring diverse biological processes (Yamaguchi-Shinozaki & Shinozaki, 2005) –, being one of the most common sequence found in the promoter and enhancer regions of several genes in eukaryotes (X. Y. Li et al., 1992; S N Maity & de Crombrugghe, 1998). When this regulatory element is localized in the promoters, it is usually situated, at the −60/−100 bp location from the transcription start site (TSS) (Dolfini et al., 2009; Dolfini, Minuzzo, et al., 2012; Roberto Mantovani, 1999), being essential to a high-level expression of the targeted genes (Bezzecchi et al., 2019, 2020).

The CCAAT box is the sequence recognized and regulated by the Nuclear Factor Y and performs a valuable role in the activation of 25% – 30% of mammalian genes (Libetti et al., 2020). This DNA element can be found in 30% of eukaryotic promoters, and have an identical incidence as the TATA box, since the CCAAT box is generally present even in the promoters absent of TATA element, among others pro-survival and cell-cycle promoting genes and also the ones involved in the metabolism (Dolfini, Gatta, et al., 2012).

This element, bound by the Nuclear Factor Y, has been considered as a promotor event practically exclusive, since NF-Y functions as a vital promoter organizer embroiled in the recruitment of the polymerase II, and surrounding transcription factors (Ceribelli et al., 2009; Oldfield et al., 2014), like growth-controlling and oncogenic TFs (Dolfini et al., 2016). The minor groove regions of the CCAAT box are intersected by the NF-YA in order to bind to it, however, major groove regions do not experience the same binding, meaning that they are available for further interactions and possible binding to other TFs (Nardini et al., 2013; Nardone et al., 2017).

The CCAAT box is included in a large spectrum of promoters regulating several genes that are, for instance, overexpressed in several kinds of cancer (Dolfini et al., 2016, 2019; Ly et al., 2013; Zambelli & Pavesi, 2017). NF-Y is the only protein demonstrated to require all of the five nucleotides, despite the fact that there are several other DNA binding proteins that interact with this sequence (Ly et al., 2013). Due to its importance, the knockout of the NF-Y TF shows to be lethal during the early embryonic development (Bhattacharya et al., 2003; Sankar N Maity, 2017).

The binding of NF-Y to the CCAAT box is guaranteed by the recognition made by the NF-YA subunit, in particular by the conserved A2 helix and the Gly-loop (please see chapter 5.6 - NF-YA). This combination detects the proper sequence of pyrimidine or purine bases that allow the hydrogen bond interactions and disfavours steric hindrance (Bi et al., 1997; Nardini et al., 2013). Importantly, the A2 helix alongside with the Gly-loop are necessary to achieve this specificity, as all of the NF-YA residues that are involved in the sequence-specific recognition are conserved, having in

consideration the earlier mutagenesis data and also with the matrix of DNA specificity (Huber et al., 2012; Nardone et al., 2017).

The current model on the mechanism of molecular recognition of the CCAAT box along the DNA molecule by the NF-Y protein envisions the possibility of the A2 helix of NF-YA sliding across the DNA, and in this way, playing a role of sequence sensor, in order to verify the interactions made, until reaching the suitable one, with DNA bases at the minor groove (Huber et al., 2012; Nardini et al., 2013). This supports the fact that the A2 helix of NF-YA shows a disordered structure when is lacking the bound DNA (Huber et al., 2012).

5.4 Subunit structure and splicing of NF-Y (regulation of the subunits)

The three subunits which compose this conserved TF can be described and characterized according to their sequence homology domains and lengths. It can be generalized that two of the subunits (NF-YB and NF-YC) display a Histone-Fold Motif (HFM), being NF-YC protein longer than NF-YB, as it also includes a transactivation domain. The regulatory NF-YA subunit, along with the conserved DNA-binding and trimerization domains, also displays a glu-rich transactivation domain, located in the N-terminal portion of the protein. (Ly et al., 2013).

Therefore, two of the subunits accommodate a core region, consisting of a histone-fold domain (HFD), conserved throughout evolution, and are tightly associated to form a stable heterodimer, that is NF-YB/NF-YC. (Bernardini et al., 2019; Dolfini et al., 2020; Dolfini, Gatta, et al., 2012). The remaining subunit NF-YA, includes the DNA binding domain, being responsible to confer the DNA sequence-specificity to the complex (Bernardini et al., 2019; Ceribelli et al., 2009; Dolfini et al., 2020; Dolfini, Gatta, et al., 2012; Ly et al., 2013).

The dimerization of the NF-YB and NF-YC subunits is responsible for the formation of a negatively charged surface that provides anchoring sites that make possible NF-YA association and, through trimerization, elevated affinity and specificity DNA binding (Ceribelli et al., 2009; Nardone et al., 2017). The HFD is also accountable for being involved in the non-specific DNA-binding (Nardone et al., 2017).

To obtain the final formation of the DNA-bound NF-Y complex NF-YB and NF-YC interact in the cytoplasm through their HFMs in order to form a stable heterodimer (M. Li et al., 2019; Ly et al., 2013; Sankar N Maity, 2017). This heterodimer formed will be subsequently transferred to the nucleus (M. Li et al., 2019) where it will interact with NF-YA leading to the assembly and the establishment of the heterotrimeric complex that is the transcription factor NF-Y, as we can observe on figure 2 (M. Li et al., 2019; Ly et al., 2013; Sankar N Maity, 2017). This final complex can communicate with DNA by sequence-specific interactions (Ceribelli et al., 2009), such as the CCAAT

motif (Sankar N Maity, 2017). CCAAT interactions mostly occur through the subunit, NF-YA, but several non-sequence-specific contacts are also formed (Ceribelli et al., 2009).

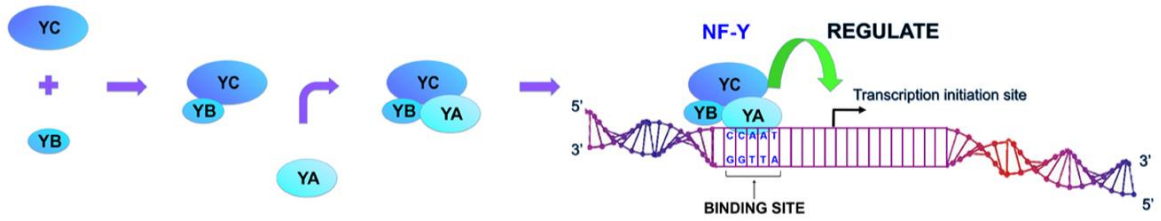

Figure 2. Formation of the NF-Y complex, with the interaction of the three subunits and the CCAAT box (Ly et al., 2013).

As mentioned, the sequence specific subunit, NF-YA, along with NF-YC, have essential characteristics to the transactivation functions of the nuclear factor Y. Namely by hosting a Q-rich transactivation domain (TAD), which consists in a high density of Gln and hydrophobic residues. (Ceribelli et al., 2009; Roberto Mantovani, 1999; Nardone et al., 2017) The harbouring of the Gln-rich TAD occurs at the N-terminal region in NF-YA, and at the C-terminal, in the case of the NF-YC (Roberto Mantovani, 1999; Nardone et al., 2017). These two subunits are also characteristic by suffering alternative splicing (AS) events (Dolfini et al., 2019) – process that generates variably spliced mRNAs, through the selection of diverse splice sites within a pre-mRNA (Greenberg & Soreq, 2013) –, and being, by that, produced different isoforms: two in the NF-YA case and multiple for NF-YC (Bezzecchi et al., 2019, 2020). These isoforms maintain similar features, in terms of HFD and DNA-binding, and show divergent levels of expression in different types of tissues and cell lines, but despite the similarities, they might display distinct activation potentials (Bezzecchi et al., 2020).

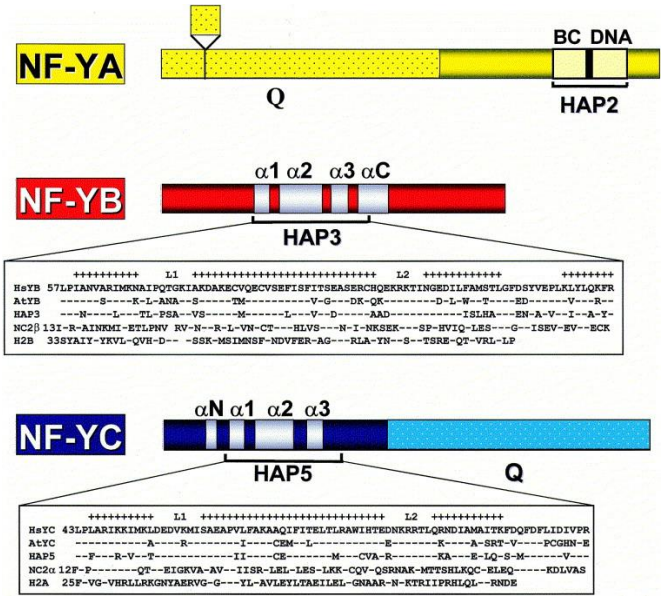

Figure 3. Schematic representation of the NF-Y genes (Roberto Mantovani, 1999).

Schematic representation where each NF-Y subunit is represented with a bar. Core regions are indicated (HAP2, HAP3, HAP5) and conserved alpha helices are represented as bright blocks. 'Q' stands for Q-rich domain.

5.5 HFD: NF-YB AND NF-YC

The HFD structure (Figure 4) is composed by the NF-YB and YC subunits within the NF-Y heterodimer (Nardone et al., 2017). Generally speaking, the HFD displays a minimum of 3 helices, α1, α2 and α3, that are separated by two loops, L1 and L2. CCAAT binding interactions are made through the DNA binding surface provided by basic residues located onto the L1-L2 loops of the NF-YB/NF-YC subunits dimer (Ceribelli et al., 2009). Through analysis of their structure, it was also possible to recognize that both α1 and α3 connect to the α2 in an almost perfect 90° angle (Nardone et al., 2017).

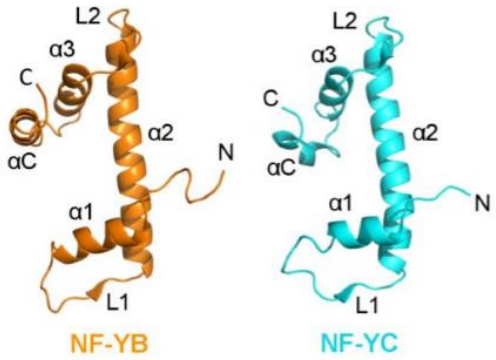

Figure 4. Three-dimensional structure of both NF-YB (orange) and NF-YC (blue) subunits and respective loops and helixes (Nardone et al., 2017).

The two dimeric subunits of the NF-Y trimer that form the HFD are homologous to core histones, in structure and sequence, being that NF-YB is related to H2B and NF-YC to H2A (Ceribelli et al., 2009; Dolfini, Gatta, et al., 2012; Nardone et al., 2017). Besides the elements already mentioned, there is other conserved and secondary structure elements, an additional αC helix can be detected at the C-termini in the NF-YB. On the other hand, in NF-YC we can observe a motif of loop-short helix loop. One of the few differences that can be found between the HFD of the NF-Y and these core histones is the absence of tails in the N-terminal regions of the HFD (Nardone et al., 2017; Romier et al., 2003).

Despite the structural similarities to core histones, there are certain characteristics that belong exclusively to NF-Y subunits. These particular features can be described as the existence of an intra-chain Arg-Asp bidentate pair linking the loop L2 to the helix α3 and that in NF-YC can be detected a presence of a totally conserved Trp at the termination of the helix α2, between the NF-YC loop L2 and the NF-YB loop L1 (Nardone et al., 2017; Romier et al., 2003).

Associating in a head-to-tail fashion, the HFD components of NF-Y display a quaternary structure, that is identical to the core histones (H2A/H2B), within the nucleosome (Nardone et al., 2017; Romier et al., 2003). This domain establishes a two-fold almost symmetric axis between the polypeptide chains, where it is possible to recognize an antiparallel association between the loops of each subunit, the L1 of NF-YB and L2 of NF-YC (Nardone et al., 2017). The outstanding stability of this complex is due, in part, to the interactions that are established between the loop regions and hydrophobic packing of residues that reside in the long α2 helixes of both NF-Y subunits (figure 5) (Nardone et al., 2017).

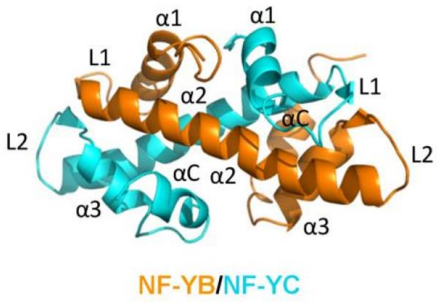

Figure 5. Three-dimensional of the NF-YB/NF-YC HFD dimer with helices (α) and loops (L) represented (Nardone et al., 2017).

Another characteristic of HFD, and in particular the L1-L2 pairs, is the formation of a molecular platform necessary to the binding and bending of the DNA (Nardone et al., 2017). With the calculation of the electrostatic potential of the upper portion of the heterodimer is possible to appreciate that this TF can provide an extremely basic surface which sustains polar and van der Waals interactions with the negatively charged phosphodiester backbone of the DNA (Nardone et al., 2017).

Due to the remarkable similarities, it is perceptible that this specific HFD can be compared to other histone-like proteins, however its wide surface groove, that is negatively charged, accounts for the majority of their specific singularity (Nardone et al., 2017). This negatively charged surface is predominantly built from helix α1 and α2 of NF-YC subunit and the helix α2 of NF-YB residues, which on their opposite side are also responsible for heterotrimerization, and allows NF-YA binding to the dimer, essentially resulting through the interaction of the basic A1 NF-YA helix and the NF-YB/NF-YC domain acidic surface groove (figure 6) (Nardini et al., 2013; Nardone et al., 2017).

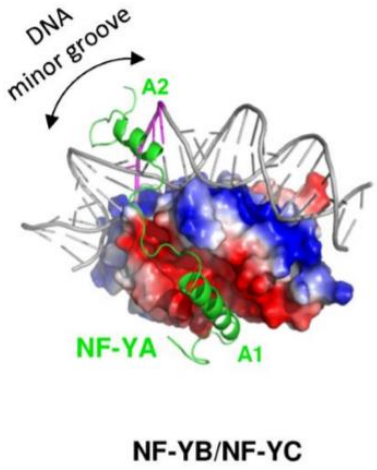

Figure 6. Electrostatic surface of NF-YB/NF-YC HFD dimer with NF-YA (green) and DNA (grey) represented (Nardone et al., 2017).

The αC helices, from both NF-YB and NF-YC, have an extremely important task in the recognition and binding of the NF-YA subunit to the dimer, however that is not the only function that these helices are accounted for in other HFD proteins, since it is known that they are also relevant in the TATA binding protein (TBP) association to NC2 (Nardone et al., 2017). Within this structurally homologous complex, they act as a scaffold to allow association of TBP to permit the assembly of the transcriptional preinitiation complex at the TATA box (Nardone et al., 2017). It is also known that these αC helices mediate interactions with other TFs, both in presence and absence of DNA, since the helices are not directly connected in DNA bonds (Nardone et al., 2017). Additionally, the αC helix on the NF-YC subunit is responsible for the interaction with cell cycle controlling proteins (as it is the case with the proto-oncogene c-myc and the tumor suppressor p53 (Imbriano et al., 2012)), which can translate the importance of these helices in eukaryotes (Nardone et al., 2017).

5.6 NF-YA

NF-YA, a subunit of the nuclear factor Y, is composed by 347 aminoacids and displays a conserved alpha-helical core domain of 53 aa composed by two helices, A1 and A2 (Bernardini et al., 2019; M. Li et al., 2019). This subunit is believed to be the regulatory and limiting subunit (Dolfini, Minuzzo, et al., 2012), which functional regulation has been described during cell cycle progression. It was also shown its downregulation during terminal differentiation in specific types of cells, including the C2C12 cell line (Basile et al., 2016). In fact, this specific subunit of NF-Y is absent from specific postmitotic cells or tissues (Dolfini, Minuzzo, et al., 2012).

Within the conserved core domain, the A1 helix contains a N-terminal region (Bernardini et al., 2019; Nardone et al., 2017) with numerous polar residues that are mainly positively charged (Nardone et al., 2017), which are involved in the interaction and binding of the HFD (NF-YB/NF-YC)

(Bernardini et al., 2019; Nardone et al., 2017). The mentioned association is mediated through hydrogen bonds and salt bridges, formed between the basic residues in the A1 helix, and the residues located in the surface groove of the HFD dimer that is negatively charged (Nardone et al., 2017).

In the C-terminal region of the core domain it is found the A2 helix responsible for the DNA binding, and responsible for the direct recognition of the CCAAT box (Bernardini et al., 2019; Nardone et al., 2017). This helix together with the Gly-X-Gly-Arg-Phe (where X can correspond to any residue) loop motif – located after the A2 helix – are accountable to deliver the sequence specific contacts needed to make the interactions with the CCAAT sequence, through the infiltration into the DNA minor groove. The process of DNA binding of the trimer is executed by the Gly-loop, while the backbone phosphates are contacted by the HFD subunits and CCAAT nucleotides by NF-YA residues (Bernardini et al., 2019; Nardini et al., 2013; Nardone et al., 2017). The Gly rich-loop contacts are specifically made by the juxtaposition presented by the protein backbone towards the two Gly-Gly carbonyl O atoms and the bases of the CCAAT complementary DNA strand (GG). On the other side, the bases positioned on the CCAAT DNA strand (AAT) are hydrogen-bonded onto the adjacent chains of Arg and His residues, included in the A2 helix of NF-YA (Nardone et al., 2017).

It is known that the NF-YA subunit is phosphorylated at 2 serines (Ser320 and Ser326) by the cyclin-dependent kinase 2 (CDK2) (Bernardini et al., 2019). These serines are located on the C-terminal side of the conserved (HAP2) domain and are extremely relevant to the cell-cycle progression. One of these residues (Ser320) directly contacts the DNA backbone at the CCAAT bases, providing further stabilization to the DNA-bound NF-Y complex (Bernardini et al., 2019).

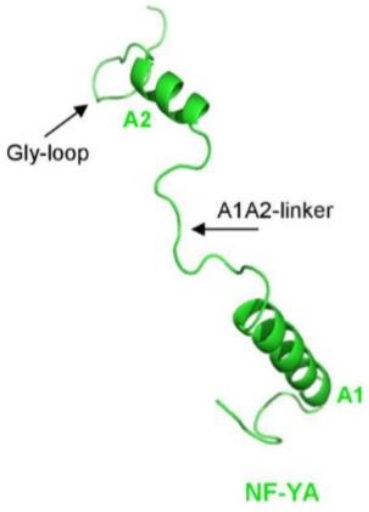

Figure 7. Three-dimensional structure with A1A2-linker, both helixes (A1 and A2) and the Gly-loop (Nardone et al., 2017).

Besides the structural elements already described, NF-YA core domain is provided with another relevant feature, the A1A2-linker (figure 7), that is equally essential to the maintenance of the subunit. This linker is the responsible to uphold the conformational flexibility that is required so that the NF-YA A2 helix can be directed upon the DNA, while at the same time, the A1 helix

interaction stability is guaranteed, when it is associated to the NF-YB/NF-YC interface (Nardone et al., 2017).

5.6.1 NF-YA isoforms: long and short

NF-YA, the regulatory subunit gene, encodes for two major protein isoforms generated by its alternative splicing regulation: the short – NF-YAs – and the long isoform – NF-YAl – with supplementary micro-heterogeneity available in a splicing acceptor site (Basile et al., 2016; Bezzecchi et al., 2020; Ceribelli et al., 2009; Dolfini et al., 2020; Dolfini, Minuzzo, et al., 2012; Libetti et al., 2020). This occurs depending whether the splicing of Exon 3, that codes for 28 aminoacids, occurs, or not (Basile et al., 2016; Dolfini et al., 2020; Dolfini, Gatta, et al., 2012; Dolfini, Minuzzo, et al., 2012; Libetti et al., 2020). As mentioned before, this stretch locates at the N-terminus of the protein, more specifically within the Gln-rich transactivation domain (Basile et al., 2016; Bezzecchi et al., 2019; Ceribelli et al., 2009; Dolfini et al., 2019; Dolfini, Minuzzo, et al., 2012; Libetti et al., 2020).

In this way, the short isoform is the one that emerges from alternative splicing, absenting the 28 amino acids that encode the exon of the NF-YA Q-rich domain (Basile et al., 2016; Ceribelli et al., 2009; Dolfini, Minuzzo, et al., 2012). The long isoform represents the full form of this subunit of NF-Y (Ceribelli et al., 2009), being the isoform with the longer transactivation domain (this isoform can be found, for instance, in muscle cells) (Libetti et al., 2020). The DNA binding activity, which is provided by the C-terminal HAP homology domain, is maintained intact in either isoforms (Sankar N Maity, 2017).

Despite the (short) N-terminal sequence differences, the two isoforms have identical subunits-interactions and DNA-binding properties *in vitro*. *In vivo*, the expression of the isoforms is subject to variations in expression levels and in relative abundance of the protein products in different tissues and cell lines (Libetti et al., 2020). It has been noticed that the levels of the short isoform of NF-YA are elevated in embryonic stem cells but decrease during differentiation of embryoid bodies (Ceribelli et al., 2009). This isoform is specifically required for stemness maintenance of the hematopoietic system, supporting the idea of functional specification of different NF-YA isoforms (Ceribelli et al., 2009).

It is important to note that NF-YA is responsible for harboring both DNA-binding and transactivation domains. This subunit provides the sequence-specific DNA contacts, while the HFD, composed by the other two subunits, interacts with the DNA via nonspecific HFD-DNA contacts (Oldfield et al., 2019). Furthermore, depletion of the NF-YA protein leads to the accumulation of ectopic nucleosomes over the TSS (transcription start site), reducing promoter accessibility (Oldfield et al., 2019).

The importance of this TFs underlies the fact that no cell line has yet been described lacking this nuclear factor activity, such as NF-YA or even the HFD. Additionally, the inactivation of the NF-YA, through RNAi, was shown to be fatal to cells, since it leads to cell-cycle arrest and apoptosis in different cellular contexts (Bezzecchi et al., 2019; Libetti et al., 2020).

5.7 NF-Y IN DIFFERENT CELL TYPES

The NF-Y TF, together with the CCAAT motif, have shown to play a major role in transcription in certain cell types, that can be regulated by distinct cellular signalling events or even pathogenic conditions (Sankar N Maity, 2017).

Some of the most relevant cellular signalling events mentioned to be regulated by NF-Y can be described as the mechanical stress, endoplasmic reticulum stress, cholesterol and fatty acid metabolism, interferon gamma responses and cell cycle progression. On the other side, the pathogenic conditions include DNA damage response, neurodegenerative diseases and cancer (Sankar N Maity, 2017).

5.7.1 Fibroblasts and type I collagen

Studies showed that NF-Y is required for the fibroblast proliferation in culture. It is also known that this TF is the one binding to and activating the transcription of the promotors of both subunits of the collagen type I genes (COL1A1 and COL1A2) (Sankar N Maity, 2017). The CCAAT motif of the COL1A1 gene can be responsible for generating excessive fibrosis in tissues and fibroproliferative lesions in the small blood vessels (Sankar N Maity, 2017). Further research lead to the understanding that the TGF-β improved the binding between NF-Y and the COL1A1 promoter in cardiac fibroblasts in rats (Sankar N Maity, 2017). It is also believed that TGF-β is responsible to increment the nuclear localization of the NF-YA subunit, involving mitogen-activated protein kinase cascades (Sankar N Maity, 2017). All this together shows the importance that NF-Y plays in the proliferation of fibroblast and in the expression of collagen type I, under diverse conditions controlled by the TGF-β signalling (Sankar N Maity, 2017).

5.7.2 Odontoblasts and chondrocytes

BMP2 is responsible to mediate NF-Y regulation of specific genes in odontoblasts and chondrocytes (Sankar N Maity, 2017). This protein will activate the expression of the DSPP (Dentin sialophosphoprotein) gene in odontoblasts, that is also mediated by the binding of NF-Y in the DSPP promoter (Sankar N Maity, 2017). BMP2 is also accountable to activate the transcription factor Sox9, during chondrocyte differentiation, which occurs concomitantly with the presence of NF-Y in the Sox9 promoter (Sankar N Maity, 2017). Both situations show the significant role played by NF-Y in odontoblasts and chondrocytes (Sankar N Maity, 2017).

5.7.3 Muscle cells

The adult muscle lacks NF-YA (Basile et al., 2016; Sankar N Maity, 2017) . A loss of activity of NF-Y during muscle differentiation results in the downregulation of diverse cell cycle genes (Sankar N Maity, 2017). Additionally it was noticed that the long isoform of NF-YA (NF-YAl) plays a specific role in the muscle-specific genes expression (Sankar N Maity, 2017). Ultimately, it is perceived that the activity of NF-Y in the muscle cell differentiation program is crucial for it, also due to the fact that the NF-YA expression inactivation culminated in the inhibition of proliferation and differentiation of muscle cells (Sankar N Maity, 2017).

5.7.4 Neuronal cells

In the neuronal cells lineage, it is possible to find NF-Y expressed in mature cells, and being deregulated when in the presence of hereditary neurodegenerative disorders, like polyglutamine expansion diseases (Sankar N Maity, 2017). Studies in a mouse model with spinal and bulbar muscular atrophy, expressing the androgen receptor gene, showed that TGF-β receptor type II gene, which is regulated by NF-Y, presents a diminished expression and also an interruption of its signal in the spinal motor neurons (Sankar N Maity, 2017). In this pathogenic atrophy the sequestration of NF-Y is triggered in molecular aggregates and an inhibition of the NF-Y mediated transcription is detected (Sankar N Maity, 2017).

NF-Y regulated chaperon genes have a repressed expression in the spinocerebellar ataxia type 17 disorder, that has a wide-range of symptoms, including seizures, cognitive dysfunction psychiatric symptoms and Parkinsonism (Sankar N Maity, 2017). TATA box-binding protein mutant is expressed in this disorder, interacting with NF-YA, constraining in that way the binding of NF-Y to the cellular HSPA5 promoter (Sankar N Maity, 2017).

Htt protein has a crucial role in the Huntington's disease and its mutant interacts with NF-Y, restraining it recruitment by the HSP70 promoter and inhibiting their expression (Sankar N Maity, 2017). This jointly with the SCA17 studies make it plausible that the inactivation of NF-Y can be one of the possible causes of both diseases, since it leads to restraining of different chaperone proteins possibly conducting to protein misfolding or aggregation and neuronal apoptosis (Sankar N Maity, 2017).

5.7.5 Stem cells

It is possible to find in bone marrow hematopoietic stem cells (HSC) expression of NF-Y, which regulates the homeobox gene family factor HOXB4 (Sankar N Maity, 2017). HOXB4 plays a vital role in HSC (Dolfini, Minuzzo, et al., 2012; Sankar N Maity, 2017) and is induced by two major NF-Y binding sites, one in the promoter and the other in an intronic enhancer (Dolfini, Minuzzo, et al., 2012). It was previously shown that, when NF-YA is overexpressed, primitive hematopoietic stem cells have the aptitude to stimulate the expression of various key genes, that are responsible for the regulation of HSC self-renewal (Sankar N Maity, 2017).

Studies have shown that NF-Y is essential for the development and differentiation of HSC, also in immune response homeostasis, through its role in the major histocompatibility complex class II genes and in the transcriptional circuitry for regulatory Treg cells' lineage specification (Sankar N Maity, 2017).

Experiments in transcriptional activation were performed and indicated that the two distinct isoforms presented a similar activation potential (Basile et al., 2016). Nonetheless, the short isoform is expressed in high levels in hematopoietic stem cells, decreasing during differentiation (Basile et al., 2016). The overexpression in the HSC boosts their self-renewal potential, enhancing, in this way, bone marrow transplantations in *ex vivo* experiments (Basile et al., 2016; Zhu et al., 2005).

NF-Y plays a major role in embryotic stem cells (ESC), since it is required for their proliferation. NF-Y is also responsible for the expression of various pluripotency-associated genes, such as Oct4, Nanog and Prdm14 (Sankar N Maity, 2017).

Experiments on this subject presented an understanding, in which, other transcription factors – that are important for ESC identity or neuronal differentiation – together with NF-Y are responsible for the regulation of cell type-specific functions (Sankar N Maity, 2017).

5.8 NF-Y IN DIFFERENT ORGANISMS

In yeast, NF-Y is evolutionary conserved and represented in a distinct way with the HAP genes. Instead of the common NF-YA, NF-YB and NF-YC of higher eukaryotes, in fungi it is represented by the HAP2, HAP3 and HAP5 gene products, respectively (Sankar N Maity, 2017; Nardone et al., 2017). The yeast counterpart of NF-Y is specifically responsible for the expression and regulation of nuclear genes that are part of the mitochondrial respiration chain (Sankar N Maity, 2017) and lacks the Q-region core domains. The activation is instead encoded by another subunit, the HAP4, that is also included in NF-Y complex in fungi (Nardone et al., 2017).

In plants, the nuclear factor Y also has three subunits, but instead of being encoded by one gene, it is instead by a family, that are differentially expressed in several tissues (Petroni et al., 2012). This difference causes an extensive diversity of the NF-Y complexes, due to the multitude of possible subunit combinations (Laloum et al., 2013; Nardone et al., 2017). This emerges as an advantage to plants, being able to face all of the numerous environmental conditions experienced (Nardone et al., 2017).

5.9 NF-Y AND ITS TARGET GENES

It is known that NF-Y is responsible for the regulation of the transcription of various genes that are related to the cell cycle and numerous cellular processes involved in human diseases (Ly et al., 2013).

NF-Y started by being considered as a protein binding only to the major histocompatibility complex (MHC) class II conserved Y box. Over the time, identification of other roles in biological pathways have been made, together with the identification of different target genes (Ly et al., 2013). It was possible to separate the target genes in two distinct categories: cell cycle-related genes and human disease-related genes (Ly et al., 2013).

5.9.1 Cell cycle related genes

In the category of the cell cycle-related genes, NF-Y has control over the expression of numerous key regulators of the cell cycle, as it is the case of topo IIα, cyclins and cdc25C genes (Ly et al., 2013).

The promoters of diverse cyclins and cdc25C genes contain CCAAT boxes, relating to the relevance of NF-Y in moduation of their activity, during cell cycle (Manni et al., 2001). Previous studies show that NF-Y mediate transcriptional inhibition of those genes, during p53-dependent G2 arrest induced by DNA damage, implying NF-Y with a role in transcriptional regulation during G2 checkpoint (Ly et al., 2013). Additionally, NF-Y is also accountable for the prevalent activation of G2/M and anti-apoptotic genes (Benatti et al., 2008; Ly et al., 2013).

5.9.2 Human disease related genes

Once the heterodimer is localized in the nucleus it activates the expression of several chaperone genes that will associate with pathogenic conditions (in particular including protein misfolding diseases). Examples of diseases that are impacted by the NF-Y are neurodegenerative diseases and different metabolic disorders, such as diabetes and non-alcoholic hepatic steatosis (Sankar N Maity, 2017).

On the other hand, it has been shown that during sterol depletion, NF-Y will associate with SREBP1, increasing cholesterol or fatty acid metabolism genes, that consecutively associate with the metabolic syndrome (Sankar N Maity, 2017).

This transcription factor is also recently emerging as one of the regulatory factors of many overexpressed genes in several distinct types of cancer (Ly et al., 2013). Besides the characteristics already cited, NF-Y can also regulate the transcription of numerous human diseases related genes (Ly et al., 2013). Examples include gamma-globin, Hoxb4, MHC class II, transforming growth factor beta type II receptor and the SRY-related HMG-box (Sox) family (Ly et al., 2013).

Although NF-Y genes are infrequently mutated or amplified in cell lines or cancer specimens (Bezzecchi et al., 2019, 2020; Dolfini et al., 2019), over the time, the overexpression of the NF-YA

subunit has been associated with the increase proliferation of cancer cells and/or its aggressiveness and help in the progression of tumours (Bernardini et al., 2019; Bezzecchi et al., 2019). It was previously verified an over-representation of CCAAT in the promoters of cancer "signature" genes, as Sox1, KLF4 among others, and TFBSs (Transcription Factor Binding Sites), particularly in breast cancer (Bezzecchi et al., 2019; Dolfini et al., 2019).

Accordingly, it seems that the tumours depend on the binding of the CCAAT box to activate a considerable number of genes requested for cancer progression (Bezzecchi et al., 2019; Dolfini et al., 2019), and a transient functional inactivation of NF-YA induces the inhibition of cell-cycle and apoptosis in diverse cellular contexts (Bezzecchi et al., 2020).

Different NF-Y elements have been shown to be expressed in different types of cancer (Bezzecchi et al., 2020; Dolfini et al., 2019), co-binding with oncogenic transcription factors, like E2F, FOS and MYC (Bezzecchi et al., 2019, 2020). High levels of NF-YA appears to give a poorer prognosis. Elevated levels of NF-YA, specially of its short isoform, are noticed in cells of the epithelial ovarian cancer (Bezzecchi et al., 2019; Dolfini et al., 2019), and in the most of the tumours with epithelial origin (Bezzecchi et al., 2019). Higher expression of the short isoform, apace with other TFs were described in triple negative breast cancers (Bezzecchi et al., 2019; Dolfini et al., 2019). Elevated expression of NF-YA mRNA were reported in the "diffuse" type of gastric cancer (Dolfini et al., 2019). NF-YB acts as promoter in genes of the diffuse large B-cell lymphoma (Bisikirska et al., 2016). Lastly, in gliomas, colon adenocarcinomas and choroid plexus carcinomas was noted higher levels of the NF-YC (Bezzecchi et al., 2019; Dolfini et al., 2019).

6 AIMS

Before the start of our project, two rabbits were immunized by injection of the full-length NF-YA long, with the general aim of obtaining *sera* recognising the N-Terminal portion of NF-YAl.

The main goal of our research was to obtain reagents that are specific to the long isoform of the nuclear factor YA, including the exon 3 of this isoform. These antibody reagents will turn out valuable in the detection of this specific isoform in Western blot experiments, as well as in Immunofluorescence and Immuno-Histochemistry.

To achieve the principal aim some different steps were required, starting by producing and purifying the recombinant NF-YAl exon3 protein (TRX-6his-YA333) in order to be used as bait for affinity chromatography purification of anti-YAl antibodies, together with a non-relevant control protein (TRX-6his).

After the previous step is achieved it is important to test rabbit 1 and rabbit 2 specificity for YA333 recognition as well as to covalently couple the recombinant proteins to the solid matrix.

Lastly, we could proceed to the purification of the YA333-specific antibodies by two step affinity chromatography and also evaluate these purified antibodies for their YAl specificity.

This project will permit to obtain specific reagents responsible to distinguish both isoforms of the NF-YA subunit. This specific Ab will allow the design of new experiments in order to improve the knowledge of the mechanisms behind the different isoforms.

CHAPTER 2

MATERIAL AND METHODS

7 MATERIAL AND METHODS

7.1 BACTERIAL GROWTH MEDIA

For the necessary bacterial growth media, it was prepared three different types of plates, always under sterile conditions.

Selective plates prepared:

- LB medium and 100µg/ml of Ampicillin;
- LB medium, 100µg/ml of Ampicillin and 35µg/ml of Chloramphenicol;
- LB medium, 100µg/ml of Ampicillin, 35µg/ml of Chloramphenicol and 0,5% Glucose;

LB medium: 1% tryptone, 0.5% yeast extract, 1% NaCl (sterilized).

LB agar plates: 1% tryptone, 1% NaCl, 0.5% yeast extract and 1.5% agar.

Antibiotics stock:

- 50mg/ml of Ampicillin (Amp);
- 35mg/ml of Chloramphenicol (Cm) in EtOH;

Glucose stock: 40%

7.1.1 Selective plates procedure

For the preparation of the selective plates we followed the procedure:

1. Melt the LB medium at 100 °C water bath (if stored solid);
2. Let it cool and added the antibiotics and glucose the way desired:
   - 100 µg/ml of Ampicillin;
   - 35 µg/ml of Chloramphenicol;
   - 0,5 % Glucose.
3. Mixed it well and pour ~ 20 ml in each bacteriology plate;
4. Let it solidify before using or storing it at 4 °C.

7.2 BACTERIAL STRAIN

In this project, it was utilized *Escherichia coli* cells in the BL21 (DE3) strain, that are characterized by being chemically competent cells, prepared by the method described hereafter (co-transformation of calcium chloride competent cells). This strain is used for protein expression of T7 RNA Polymerase-based systems, being suitable for plasmid transformation and protein expression and co-expression. All of this is possible due to the ease of its manipulation and its rapid growth. It is also characterized by:

- Transformation efficiency of 1–5 x 10^7 cfu/µg pUC19 DNA;
- Derived from B strain;
- Deficient in proteases Lon and OmpT, which had the potential to degrade the proteins expressed;
- Resistant to phage T1 (fhuA2).

7.3 DNA VECTOR

pSB7 harbouring the genes encoding for YA333 and TRX proteins. The exon 3 of NF-YAs was cloned in a pET-32a(+) plasmid. The proteins encoded by this plasmid are:

- ➪ YA333 (266 aa and molecular weight of 28,1 kDa);
- ➪ TRX (189 aa and molecular weight of 20,36 kDa).

This pET-32a vector has the following features (figure 8): An origin of replication pBR322, T7 promoter, Lac operator region, several cloning sites, His-Tag and S-Tag sequences and 5900 bp.

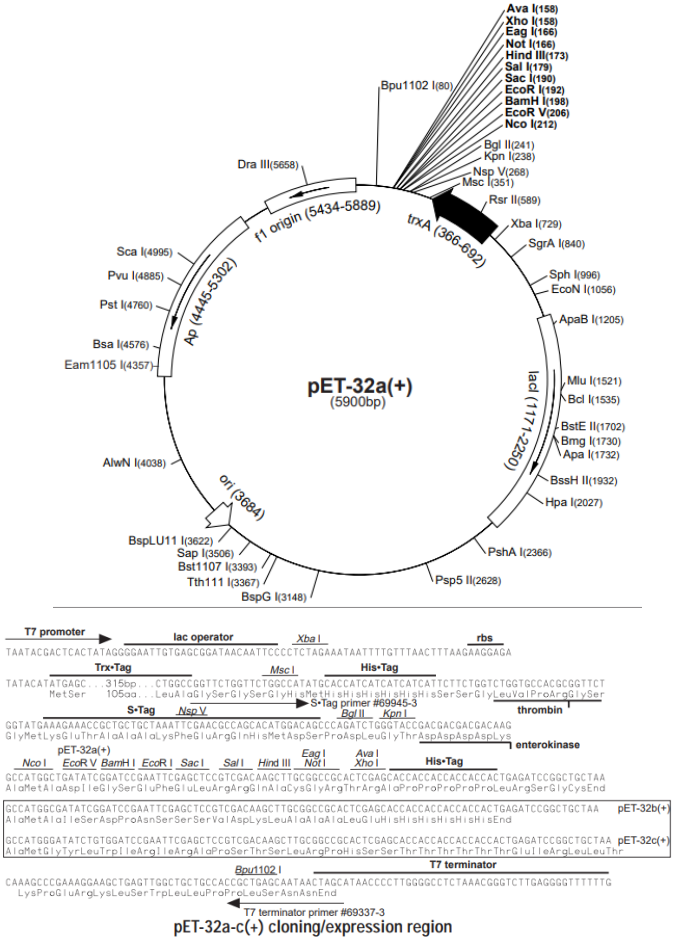

Figure 8. pEt32a vector map (Novagen).

7.4 CO-TRANSFORMATION OF CALCIUM CHLORIDE COMPETENT CELLS

In 1970, Mandel and Higa described a procedure where *E. coli* cells treated with calcium chloride can become competent cells to the DNA transformation (Dagert & Ehrlich, 1979).

Cells with this characteristic are cells who have the ability to income extracellular DNA, from its environment. The bacterial strain used in this experiment, the *E. coli* BL21(DE3), is not naturally competent so, in order to achieve this characteristic, it is necessary to transform these cells through $CaCl_2$. This procedure will generate pores in the membrane of the cell, which will allow the genetic material to pass through it, conferring, in this way, the competence to the cell.

7.4.1 Material

- 3 pre-chilled Eppendorf tubes;
- SOC medium;
- 4 selective LB-agar plates.

7.4.2 Procedure

Before starting the procedure, prepare the necessary Eppendorf tubes and let them on ice.

1. Thaw the competent cells on ice;
2. Gently mix the cells by flicking the tube;
3. Aliquot 100-150 µl of the competent cells suspension in the pre-chilled Eppendorf tubes, and reserve one Eppendorf to a NO DNA sample – to test for strain sensitivity to selective agent;
4. Add 10 pg – 150 ng DNA (or ligation mix) and mix gently by flicking the tube;
5. Incubate the tubes on ice for 30 minutes;
6. Heat pulse the tubes in a 42 °C water bath, for 45 seconds – the length of the pulse is critical for highest transformation efficiencies;
7. Immediately transfer the tubes to ice and incubate them for 2 minutes;
8. Transfer the tubes to RT and add 1 ml of SOC medium in each tube;
9. Incubate the tubes, under agitation, for 1 hour, at 37 °C;
10. Spin for 1 minute, at 10.000 rpm at RT;
11. Aspirate and discard 0,9 ml of supernatant and resuspend the cell pellet in the remaining medium;
12. Spread 1/10 and 9/10 of the transformed cells culture on selective LB-agar plates;
13. Incubate overnight (ON), at 37 °C.

7.5 Glycerol Stocks preparation

The following preparation must occur under sterile conditions.

1. Transfer 1 ml of overnight culture (at 37 °C) of co-transformed cells into a sterile tube and centrifuge it for 1 minute at 13,000 rpm RT.
2. Remove the supernatant and resuspend the cells in 500 µl of selective medium (LB + Amp/Cm) and 500 µl of 2x glycerol stock solution (65 % glycerol, 0,1 M $MgSO_4$, 25 mM Tris-HCl pH 8,0).
3. Preserve the preparation at -80 °C.

7.6 Buffers and Reagents

- → Buffer A (BA): dH_2O, pH8 Tris 10 mM, NaCl 400 mM, $MgCl_2$ 2 mM;
- → Buffer A2 (BA2): dH_2O, pH8 Tris 10 mM, NaCl 400 mM, $MgCl_2$ 2 mM, Imidazole (important to interfere with the binding of other proteins and therefore elute them) 2 mM.
- → Buffer A100 (BA100): dH_2O, pH8 Tris 10 mM, NaCl 400 mM, $MgCl_2$ 2 mM, Imidazole 100 mM, 10% glycerol;
- → Buffer A250 (BA250): dH_2O, pH8 Tris 10 mM, NaCl 400 mM, $MgCl_2$ 2 mM, Imidazole 250 mM, 10% glycerol;
- → Protein Assay Dye Reagent Concentrate (BioRad): 20 %.

7.7 Recombinant His-Tag Protein Expression and Purification

7.7.1 Induction

Day 1:

Streak the glycerol stock of transformed cells on selective-agar plates.

Incubate overnight at 37 °C.

Day 2:

Preparation of the starting culture.

- → In a 250 ml flask, pick 1-3 colonies of individual clone and inoculate in 50 ml of LB to which are also added ampicillin and chloramphenicol antibiotics.
- → Keep the culture under agitation ON, at 37 °C.

Day 3:

Large-scale induction.

1. Prepare a flask with the desired amount of LB medium (corresponding to 1/3 of the total volume of the flask) and the antibiotics;
2. Inoculate the starting culture (LB V=1/50 of the final volume) in the pre-warmed medium;

3. Put the culture under agitation (190 rpm) at 37 °C, and monitor the OD_{600} from t=0 until it reaches OD_{600} ~ 0,5 (generally takes 75 minutes);
4. When the desired value is reached, lower the temperature to 30 °C for 15 minutes;
5. Take 1 ml of the preparation to an Eppendorf tube, for the Non Induced sample – to be later used in the SDS-PAGE;
6. Induce the preparation with IPTG (0,4 mM) and maintain under agitation 190 rpm, for 3 hours at 30 °C;
7. Take 1 ml of the preparation to an Eppendorf tube, for the Induced sample;
8. Centrifuge at 4.400 rpm for 30 minutes at 4 °C;
9. Discard the supernatant and weight the pellets;
10. Freeze the samples in liquid nitrogen and store them at -80 °C.

7.8 Induction and Purification Test

This test is used in order to verify the efficiency of the induction previously done and if the induced proteins are soluble. It can also be conducted to determine the best conditions for large-scale purification and to estimate the overall purification yield.

7.8.1 Procedure

1. Start by thawing the pellet obtained from the 50 ml of the induced culture;
2. Cell lysis in native state: Resuspend the pellet carefully in 5 ml of lysis buffer BA2 for each gram (1 g) of pellet – do the procedure on ice;
3. Sonicate for 5 cycles of 30 seconds on ice and 30 seconds pause;
4. Centrifuge the lysate for 30 mins at 13.000 rpm, at 4°C;
5. Place the supernatant in a new tube;
6. Take 50 µl of the preparation to an Eppendorf tube, for the Soluble Fraction sample – to be later used in the SDS-PAGE;
7. Discard the obtained pellet – the protein is soluble;
8. Nickel resin equilibration:
   - Take required volume of the 50% His-Select® Nickel Affinity Gel (Sigma-Aldrich) in ethanol in an Eppendorf tube to obtain 25 µl of bed volume for each purification test – This nickel resin has a capacity of 15 mg of bound protein/ml;
   - Resuspend the resin in > 5 Column Volume (CV) of H_2O and centrifuge for 30 seconds at 4.000 rpm at 4°C;
   - Aspirate, using a syringe, and discard the supernatant;
   - Repeat this procedure with H_2O and BA;
   - Resuspend the resin in 5 CV of BA2 and centrifuge at 4.000 rpm for 30 seconds;
   - Collect the equilibrated beads: aspirate and discard the supernatant;

> Load 1/4 (500 µl) of the cell lysis soluble fraction volume on the resin;

9. Place the mixture under rotation, at 4°C, for 40-60 minutes – In this phase His-tag fusion proteins will interact and bind to the resin exploiting the natural affinity of poly-histidine tag with nickel ions.
10. After, centrifuge for 30 seconds at 4.000 rpm;
11. Take 50 µl of the supernatant to an Eppendorf tube, for the Flow Through sample (unbound sample);
12. Wash twice the pellet which represents the bound fraction with BA2. In each time resuspend the resin pellet with 500 µl of BA2, centrifuge for 30 seconds at 4.000 rpm and discard the supernatant;
13. Dry the resin and add 50 µl of Sample Buffer 2x to obtain the bound sample;
14. Do the preparation of the previous samples and store them at -20 °C, until needed.

7.9 LARGE-SCALE PURIFICATION

The large-scale purification is performed in a similar way to the Purification Test. However, it is conducted with the larger pellet from the induction (> 50 ml), and completed with a Bradford Test.

7.9.1 Procedure

1. The large-scale purification is carried out as the small-scale purification but having in mind that a larger amount of starting culture must be used.
2. The cultures are lysed in native conditions and centrifuged in order to eliminate the lysate;
3. The resulting supernatant is then incubated with the pre-equilibrated resin in lysis buffer for 1h at 4 °C and shaken mildly to favour the protein-resin interaction.
4. Take a small disposable Bio-Rad column and wash it with full volumes of H_2O and BA2, by spinning it for 1 minute at 4.000 rpm and discard it;
5. After the rotation is completed, spin for one minute at 4.000 rpm;
6. Add 500 µl of BA2 to the resin, resuspend it and load it into the pre-equilibrated column;
7. Proceed with a Bradford Test, by performing a colorimetric dosage (using Bradford reagent) of the eluates and comparing the intensity of the eluate colour with the one of BSA solution of known concentration. Utilize 10 µl of each sample (or Buffer A for control) and 150 µl of Bio-Rad reagent;
8. Start doing the washes, until reach the same colour as the control (usually 4 washes). Do each wash with 500 µl of BA2 in the column. Wait for the buffer to go through the column, test it in the Bradford Test. Repeat these steps to each wash;
9. Start the elution steps: The elutions are made with two different buffers, each one until it reaches the same colour as the control (usually 3 elutions A and 2 B):

   ⇨ Elutions A: 250 µl of BA100;

- ⇨ Elutions B: 250 µl of BA250;
- ⇨ After the elution is made, take 20 µl of each sample to a new tube;

10. Wash the column twice with full column volume with BA2, add 500 µl of BA, close the column and store it at 4 °C, for further use.

7.10 Samples Preparation

Before proceeding to the analysis of the proteins in SDS-PAGE, it is necessary to prepare each sample, from the Induction and Purifications procedures, accordingly. In each sample will be added a specific volume of Sample Buffer 2x (SB 2x).

→ SB 2x: 2% SDS, 100 mM Dithiothreitol (DTT), 10% Glycerol, 20 mM Tris-HCl pH 6,8, 0,1% Bromophenol blue.

1. For the Non-Induced and Induced samples, you must start by spinning the 1 ml samples, at 13.400 rpm for 1 minute and discard the supernatant. Add 100 µl of SB 2x to the pellet and resuspend it (up and down).
2. To the Soluble Fraction, Unbound and Resin (bound) samples add 50 µl of SB 2x, and to each elution sample add 20 µl of SB 2x.
3. Boil the samples for 5 minutes, at 95 °C;
4. Finally, spin the samples at 13.400 rpm, for 1 minute, and store them at -20 ºC, until needed, or load them in a SDS-PAGE.

7.11 Protein Analysis

7.11.1 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-Polyacrylamide Gel Electrophoresis is a useful technique in the separation of proteins by size. The Sodium Dodecyl Sulfate (SDS) is a detergent responsible to denature secondary and nondisulfide-linked tertiary structure, and posteriorly coating them with a negative charge, accordingly with their length, in order to be possible to estimate the molecular weights (Brunelle & Green, 2014).

This system entails a discontinuous gel, consisting in two different polyacrylamide concentrations and pH gel. The final polyacrylamide concentration is decided based on the size range of the involved molecules in the specific sample. These two phases of the gel are named by Stacking and Running gels (Brunelle & Green, 2014).

Stacking gel is the upper part of the gel, with a lower concentration of polyacrylamide – 4% in this case –, making it possible for the proteins to move in a quick way through it and stack into a

thigh band before passing to the next phase of the gel, the one with a higher percentage of polyacrylamide, that will allow the separation of the molecules. The running gel makes it also possible for the smaller proteins to move faster through it than the larger ones (Brunelle & Green, 2014).

7.11.1.1 Material

- 2 glass plates;
- 3 thigh gap springs;
- A rubber;
- An electrophoresis chamber and power supply;
- Comb (with 10 or 14 wells, accordingly to the number of samples);
- 2 spacers.

7.11.1.2 Reagents

- Milli-Q water;
- Tris – pH 8,8;
- Tris – pH 6,8;
- SDS;
- Acrylamide;
- APS (Ammonium persulfate);
- TEMED (Tetramethylethylenediamine);
- Isopropanol;

7.11.1.3 Solutions

→ Running Buffer 1x: 0,1% SDS, 1x TG (Tris-Glycine), dH_2O;

→ Coomassie-Brilliant Blue solution: 25% Coomassie Brilliant Blue R-250, 50% ethanol, 10% acetic acid.

This dye is utilized as staining of the proteins, after its separation by electrophoresis. In this way, after the binding of the protein by this solution, a negative charge is transferred by it to the protein, making it possible to the proteins separation to occur, through the polyacrylamide gel electrophoresis under native conditions (Arndt et al., 2012).

→ Distaining solution: dH2O, 40% EtOH, 7% Acetic Acid.

7.11.1.4 Procedure

1. Begin by assembling the material;
2. Make a pre-mixed reaction with the reagents on the following table, without APS and TEMED;

Table 1. Necessary solutions for SDS-PAGE and respective amounts.

Component	Running Gel (12%)		Stacking Gel (4%)	
	Concentration	Amount (stock solutions)	Concentration	Amount (stock solutions)
MQ H_2O	–	3,44 ml	–	2,6 ml
Tris	375 mM	2 ml (pH 8,8 1,5 M)	125 mM	1 ml (pH 6,8 0,5 M))
SDS	0,1%	80 µl (10%)	0,1%	40 µl
Acrylamide	12%	2,4 ml (40%)	4%	400 µl
APS	0,1%	80 µl (10%)	0,1%	40 µl
TEMED	0,01%	3,2 µl	0,01%	4 µl

3. Continue the mix of the running gel;
4. Transfer the final solution to the glass plate sandwich previously assembled, leaving around 1,5 cm from the top;
5. Cover the top of the gel with isopropanol in the glass superficies, and wait until the running gel solidifies;
6. After the polymerization is completed (~ 20 minutes), discard the isopropanol and wash the top of the gel with distilled water;
7. Add the APS and TEMED to the stacking gel;
8. Pour the stacking gel solution into the sandwich and assemble the comb until it solidifies;
9. Pour the running buffer into the electrophoresis chamber, and add the glass sandwich, avoiding bubbles formation;
10. Once everything is correctly assembled, remove the comb carefully;
11. Wash gently each well with the running buffer and fill the rest of the chambers with the same buffer;
12. Lay each pre-prepared sample to analyse and the protein ladder into each well;
13. Assemble the power supply, in order to activate the electric field;
14. Run at a constant current of 16 mA, until a blue tracking dye reaches the bottom of the gel, about 2 hours and 15 minutes after;
15. After the running is over, turn off the apparatus and remove the gel carefully from the glass plate sandwich;
16. Pour the gel in a box filled with Coomassie Blue, in a way to stain the gel, for at least 2 hours, under agitation, at room temperature;
17. When the staining is finished, exchange the Coomassie solution in the box for the distain solution, until the gel is correctly distained.

7.11.2 Bradford Method – Protein Assay

The Bradford Protein Assay is one of the methods utilized in the measuring of a protein concentration in solution.

A BSA curve can be obtained by interpolating the measured absorbance to the curve, in order to estimate the concentration of the desired unknown proteins.

1. Posteriorly to complete the procedure of SDS-PAGE, analyse the gel in order to understand which eluates are worth keeping;
2. Centrifuge, for 5 minutes, 13.000 rpm and at 4 °C, and mix those eluates;
3. Add 10% of glycerol to the new elution tube, and centrifuge it in the same conditions as before;
4. Start the Protein Assay by preparing the necessary cuvettes to analyse the desired samples, and make a BSA curve, as the following table:

Table 2. Necessary samples for the protein assay.

	Elution	BA2	**BSA (0,2 mg/ml)**	BioRad
Blank	–	50 µl	–	950 µl
3 µg/ml	–	40 µl	10 µl	
4 µg/ml	–	30 µl	20 µl	
6 µg/ml	–	20 µl	30 µl	
8 µg/ml	–	10 µl	40 µl	
Elution A	10 µl	40 µl	–	
Elution B	20 µl	50 µl	–	

5. Being BioRad the last reagent to add to the cuvettes, wait 5 minutes after adding it;
6. After that time is passed, mix each cuvette and start measuring the OD of each sample at 595 nm, in the NanoDrop Microvolume Spectrophotometers and Fluorometer system;
7. With the data obtained, you are able to construct the BSA Curve and estimate the concentrations wanted.

7.11.3 Western Blot

The western blot technique occurs after the separation of the proteins by the electrophorese gel, by their size, transfer and posterior immobilization to a support membrane and its selective detection, utilizing for instance, a detection system mediated by antibodies. Usually, this technique is utilized in order to identify a specific protein and supply information about the protein's molecular weight, especially those in a low abundance (Kurien & Scofield, 2003).

This technique is done by different steps, starting with a sample preparation for the gel electrophoresis (SDS-PAGE), following a membrane transfer and blocking saturation, continuing with primary antibody incubation and respective washes, secondary antibody incubation and also the respective washes, ending with the analysis of the results.

The blocking step is made in order to avoid the binding of the Abs to the membrane non-specifically. This can be done with two different blocking saturations, 5% milk or BSA (Mahmood & Yang, 2012). In this project both solutions were used, however the BSA solution was only used in one of the experiments.

Antibodies are responsible to detect the target proteins in this procedure. The primary Ab recognizes and binds the antigen, while the secondary Ab conjugated with fluorescent enzymes binds to the primary one. The conjugation of the secondary Ab with the fluorescent enzyme, triggers a reaction, when combined with the reagent (enhancer and peroxide solution), emitting light that will enable the detection. The washes, in the appropriate duration of time, between the primary antibody incubations is made in order to minimize the background signal and to remove the unbound antibodies (Mahmood & Yang, 2012).

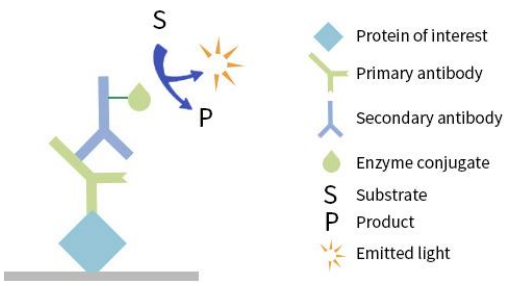

Figure 9. Representation of chemiluminescence detection of protein on membrane (SinoBiological (2020). Chemiluminescent Western Blotting Detection. Accessed in 02 of December 2020, at (https://www.sinobiological.com/category/chemiluminescence-wb).

7.11.3.1 Solutions

Before starting the procedure, prepare the following solutions:

- TBS-T (1x): 90% dH_2O, 10% TBS 10x and 0,2% Tween20 (50%).
- Transfer Buffer: 70% dH_2O, 10% TG 10x and 20% methanol (100%).
- Milk 5% (Blocking/ saturation solution): 0,05 g/ml milk powder and TBS-T (1x).
- BSA 3% (Blocking/ saturation solution): 3% BSA, 0,1% Tween, TBS 1x and dH_2O.
- Enhancer solution.
- Peroxide solution.

7.11.3.2 Procedure

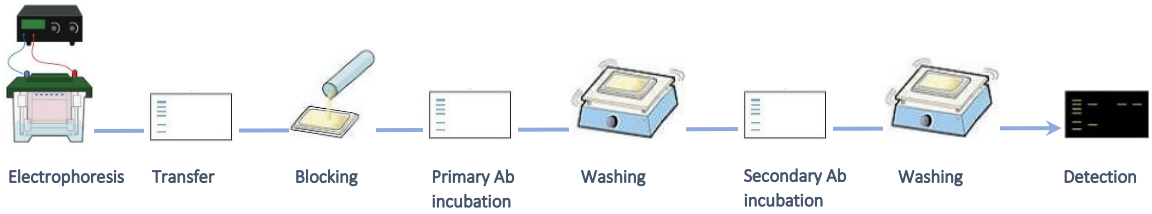

Figure 10. Schematic representation of the western blot procedure.

Day 1

1. Prepare a SDS-PAGE (12%) gel and load it with the desired samples;
2. Run the gel at constant mA (16 mA/gel) and maximum voltage (300V);
3. Prepare 1 membrane (6,5 cm x 9 cm) and 2 filter papers (7,5 cm x 9,5 cm);
4. 5 minutes before the run of gel finish you must:
   - → Immerge the membranes in dH_2O and then transfer it to transfer buffer;
   - → Immerge the sponges and the filter papers in the transfer buffer;
5. After electrophoresis, immerge the gel in the transfer buffer, for at least 5 minutes;
6. Transfer to a nitrocellulose membrane Sandwich, accordingly to the disposition presented in the image (figure 11). In between every step add TB and pass a roller carefully, to remove any bubble formed;

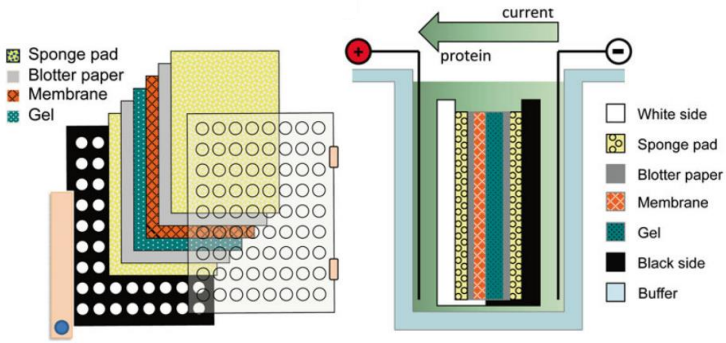

Figure 11. Schematic representation of the blot components in the transfer cassette (Hnasko & Hnasko, 2015).

7. Prepare the tank and add a magnet and some transfer buffer;
8. Place the sandwich with the black face oriented towards the black face of the tank;
9. Before closing the tank fill it with the transfer buffer and add an ice block;
10. Run for 2h at constant mA (350mA) and maximum voltage (300V);
11. Star the process of blocking of non-specific binding;
12. Prepare different boxes for each membrane to be tested differently with 7ml of blocking solution;
    - → Take off the membranes from the sandwich carefully, prepare them the way desired and place them in the respective boxes;
    - → Let it in the blocking solution, under agitation for 1 hour;
    - → Discard the blocking solution and place the respective antibody prepared in 5 ml of the blocking solution;
13. Incubation with the primary antibody overnight, under agitation at 4 °C.

7.11.3.3 Antibodies

Antibodies and respective dilution used in the Western Blot experiments:

Table 3. Antibodies utilized and corresponding dilutions.

Antibody	Dilution		Antibody	Dilution
α-Histidine	1:1000		Elutions 2+3 pH2 concentrated in BSA 3%	1:250
α-NF-YA	1:500			
α-Tubuline	1:2000		Flowthrough NF-YA	1:100
Elution 1 pH2 concentrated	1:500		Flowthrough Thioredoxin	1:100
Elution 2 pH2	1:500		Input Thioredoxin	1:100
Elution 2 pH2,7	1:500		NF-YA	1:500
Elution 2 pH4	1:500		Pre immunization Rabbit 1	1:50
Elutions 2+3 pH2 concentrated	1:500		Rabbit 1	1:50
Elutions 2+3 pH2 concentrated	1:250		Rabbit 2	1:50
Elutions 2+3 pH2,7	1:250		Thioredoxin	1:2000

7.11.3.4 Day 2

14. Store each antibody dilution, start doing the washes, by covering each membrane with TBS-T and let it under agitation for 10 minutes; Discard and repeat twice;
15. Add 10 ml of the blocking solution in each box supplemented with 1 µl of the respective secondary Ab;
16. Let it stay under agitation for 1 hour.
17. Do 3 washes with TBS-T (the same way as the previous ones);
18. Prepare the reagent with 4 ml of enhancer solution and 4 ml of peroxide solution;
19. Place the membranes in the reagent for 1 minute;
20. Proceed to the analysis on the Imaging system with a Chemidoc MP Biorad, in the signal accumulation mode;
21. Store the membranes with TBS-T at 4 °C.

7.12 Dialysis of Protein Solution

Dialysis is the method by whom small molecules are separated from larger ones, through diffusion, using a selectively permeable membrane exclusive to the small molecules. In this process it is possible to change the salt composition of the desired solution (Andrew et al., 2002).

The permeable membrane utilized has a capacity of 1,8 ml/cm (corresponding to 14 kDa) and it requires an addition of an extra 10 cm to each sample.

During the procedure it is recommended to use gloves, in order to not transfer cellulolytic microorganisms to the dialysis membrane.

7.12.1 Procedure

1. Start by cutting the membranes in the respective measures;
2. Immerge the membranes in dH_2O for about 30 minutes, exchanging the water around 3 or 4 times;
3. Discard the water, replacing it by coupling buffer. Close one side of each membrane with clams;
4. Transfer the protein solutions to the respectively membrane, closing the other side of the membrane, carefully;
5. In different beakers add 500 ml of coupling buffer, inserting there the membranes;
6. Let it under agitation, for 2 hours, at 4 °C;
7. After the 2 hours, exchange the buffer for a new coupling buffer, and let it overnight, under the same conditions as before;
8. In the day after, the protein solutions are ready to be transferred to new tubes and be used when necessary.

7.13 Antibody purification: Affinity Chromatography

Affinity chromatography is described by being one of the methods of liquid chromatography. This approach is characterized by the binding of the protein and its cognate ligand being reversible and specific. This specificity of the interactions immobilizes one of the interacting agents – the affinity ligand – to the column (Arora et al., 2017).

Before obtaining the desired target, a buffer is applied, that will elute other components, while the target stays on the column. The elution buffer is utilized in order to do the dissociation of the target from the affinity ligand (Rodriguez et al., 2020). To promote target elution an alteration of the mobile phase composition is necessary, and this is achieved by altering the pH, as it was done, from a higher to a lower pH (Hage et al., 2012). After this step, the target is released and collected for the posterior analysis.

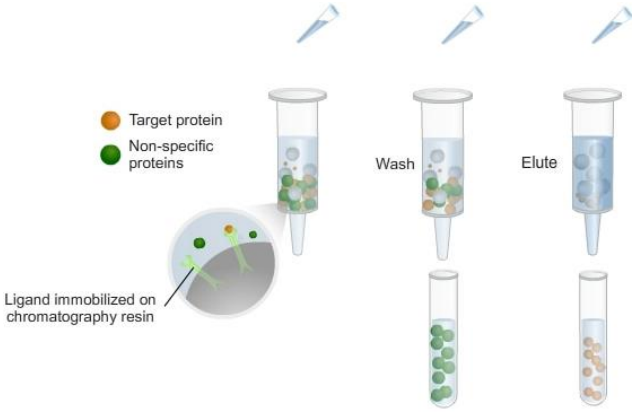

Figure 12. Schematic representation of affinity chromatography (Thermo Fisher (2020). Overview of Affinity Purification. Accessed in 03 of December 2020, at https://www.thermofisher.com/pt/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-learning-center/protein-biology-esource-library/pierce-protein-methods/overview-affinity-purification.html).

7.13.1 Solutions and buffers

Solutions and buffers needed to be prepared, before starting the protocol:

- HCl: pH 3,5, 1 mM.
- Coupling Buffer (CB): pH 7,85 NaP (0,5M sodium phosphate dibasic, 0,5M sodium phosphate monobasic and dH_2O), 5M NaCl, dH_2O.
- Acetate Buffer (AB): 0,1M Sodium Acetate, dH_2O, 0,1M Acetic Acid.
- Elution Buffers: 1M Glycine pH 7,57, dH_2O.
  - pH 2;
  - pH 2,7;
  - pH 4.
- Washing Buffer: 50 mM NaP, 500 mM NaCl, MQ dH_2O.
- PBS: 150 mM sodium chloride, 150 mM sodium phosphate.
- Tris: 1M, pH 8;
- NaCl: 1M;
- Glycine: 1M, pH 7,5;

7.13.2 Procedure

1. Before starting, it is needed to assemble the columns and wash them with 5 CV of water (with aspiration);
2. Pour resin with HCl in the column and aspirate the HCl, add water once again and aspirate;
3. Add 2 CV of coupling buffer in the column and resuspend it in the resin;
4. Leave the columns under rotation, at 4 °C, overnight, with the protein, in order to favour the protein-resin interaction;
5. Centrifuge the proteins for 5 minutes, at 2.300 rpm and 4 °C – protein-resin suspension;
6. Transfer the SN to a new falcon – unbound fraction – and do a Bradford test (in order to validate the percentage of protein that is bound/unbound after coupling);

7. Wash the column with 6 ml of coupling buffer, resuspending the resin and transferring it to a new 50 ml falcon;
8. Centrifuge it for 6 minutes, at 2.300 rpm and 4 °C;
9. Discard the SN and add 25 ml of glycine and leave it under agitation for 4 hours, at 4 °C;
10. Transfer it back to the columns, add 50 ml of CB and aspirate it;
11. Proceed to the following washes, doing them 5 times each, alternately:
    - → Add 13 ml of AB, and let it flow through the resin, by aspiration;
    - → Add 13 ml of CB, and let it flow through the resin, by aspiration;
12. Don't aspirate the last wash and resuspend the resin in the buffer, transferring it to a new column;
13. Wash the column with coupling buffer as many times as needed, until there is no more resin;
14. Leave it still for around 30 minutes, in order to the resin to settle;
15. Let the supernatant (CB) flow throughout the column and afterwards add 2,5 ml of NaCl (1M);
16. Let the NaCl pass through the resin, close the bottom of the column, and add once again 2,5 ml of NaCl;
17. Close carefully and tightly the columns and store them at 4 °C, during the night;
18. Assemble the columns;
19. Add water until the top of the column and let it dry;
20. Wash the column with water for 5 minutes, and after with PBS, also for 5 minutes;
21. Pour the rabbit serum and let it aspirate it until the end (around 90 minutes);
22. Wash the column with PBS, until the UV (absorbance) is approximately 0 (around 5 minutes and 10 CV);
23. Let the resin dry and add 2,5 ml of washing buffer;
24. Wash with the resin on the column with 2,5 ml of PBS;
25. Start doing the elution with the different pH, accordingly to the following table:

Table 4. Different elution buffer solutions and the respective Tris to be added.

Buffer	Dilution	Antibody
EB pH 4	2,5 ml	7,5 µl/ml
EB pH 2,7		7,5 µl/ml
EB pH 2		150 µl/ml

26. At the end of the elution, add the corresponding amount of Tris to each sample, to neutralize the solutions;
27. Wash the column with 5 CV (2,5 ml) of washing buffer;
28. Wash it again with PBS until the pH is at least 7;
29. Add 5 CV (2,5 ml) of NaCl (1 M) and close the column.

7.14 Protein Concentration

In order to enhance the performance of the procedure, an Amicon® Ultra-0.5 centrifugal filter device was used. This device is utilized to proceed to the ultra-filtration of the sample, with a recovery of the concentration from the dilute.

In this case, the objective is to concentrate the elution 2 and 3 together and the elution 1, both from pH 2, from the antibody purification.

1. Before starting the filtration, is necessary to mix the elution 2 and 3 together;
2. Add 495 µl of the mix (elution 2 and 3) and of the elution 1 in the Amicon® centrifugal filters;
3. Centrifuge both tubes with the centrifugal filters, for 10 minutes, at 14.000 rpm and 4 °C;
4. Transfer the flow through to a new tube;
5. Repeat the previous steps until obtaining 2,5 kDa measured in the filter;
6. Once the 2,5 kDa goal is reached, place the filter upside down, inside the tube, centrifuge it, in the same conditions as before, for about 3 minutes, so that the protein can be stored in the tube.

CHAPTER 3

RESULTS AND DISCUSSION

8 RESULTS AND DISCUSSION

8.1 PRODUCTION AND ANALYSIS OF PURIFIED RECOMBINANT PROTEINS

The first part of the project was dedicated to the production and purification of recombinant proteins to be used for affinity purification of antibodies. In order to be able to isolate antibodies specific to YAl, the YA333 recombinant protein was designed including a tandem repeat of the NF-YA exon-3 sequence, positioned downstream of the Thioredoxin tag, which improves protein solubility, and a 6-his tag for purification (see figure 13).

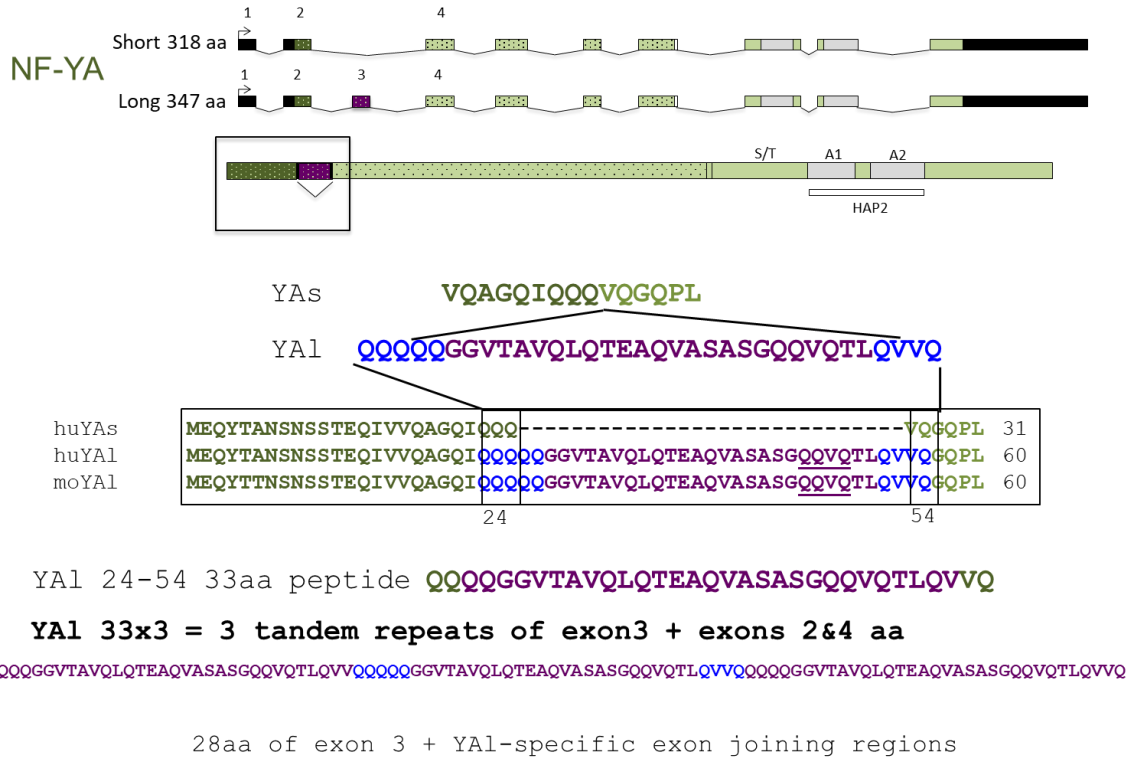

Figure 13. Schematic representation of the design of the YA333 recombinant protein.

Thioredoxin was utilized here as it is one of the most frequently used proteins, responsible for increasing the solubility of fusion proteins, having as advantage its small size which permits a good percentage of fusion. Also 6-His tag is used, as it binds with high affinity to the nickel ion in the columns of affinity chromatography (figure 14).

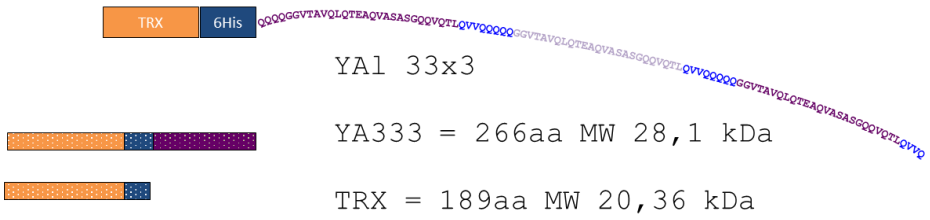

Figure 14. Schematic representation of the design of the YA333 recombinant protein (purple), with thioredoxin (orange) and 6-His tag (blue).

Several induction and purification procedures were made, in order to obtain the desired amount of protein necessary to proceed to the preparation of the affinity matrix. These included the test of different *E.coli* strains, growth media, and temperature and timing of protein expression induction. The YA333 protein, together with the TRX-tag control, was then purified using the Nickel-ion affinity chromatography method (see methods section). Subsequent to the induction and purification steps, the purified samples were analysed by SDS-PAGE, to verify the quality and quantity of proteins obtained, according to the concentration determined with the Bradford method.

In the following figures, it is possible to observe two SDS-PAGE, which show the samples corresponding to the total preparations of purified proteins (figures 16 and 17).

Before the SDS-PAGE experiment it was executed a protein quantification assay, using the Bradford Method, determining, in this way, the final concentration of the protein extracts. After preparing the different solutions according to the presented table 5 (Elution (protein sample), BA2 (buffer), BSA (protein concentration standard) and BioRad (Bradford reagent) columns – see materials and methods), samples were analysed at 595 nm to obtain the respective OD measure, as presented in the table. Lastly, the protein concentration was calculated and presented in the last four columns of the table. In 8 different purifications, the total amount of protein obtained was 5,63 µg (6,7 ml of sample at 0,84 µg/µl) for the YA333-trx protein, and 3,00 µg (2,7 ml of sample at 1,11 µg/µl) for Trx protein.

Table 5. Bradford Method for protein quantification.

	Elution	BA2	BSA	BioRad	OD (595 nm)	µg total	µg/µl	µg/µl final	volume (ml)	mg total
Blank	–	50 µl	–	950 µl	0	–				
3 µg/ml	–	40 µl	10 µl		0,133					
4 µg/ml	–	30 µl	20 µl		0,254					
6 µg/ml	–	20 µl	30 µl		0,384					
8 µg/ml	–	10 µl	40 µl		0,470					
Trx	5 µl	45 µl	–		0,395	6,45	1,29	1,11	2,7	3,00
	10 µl	40 µl	–		0,561	9,23	0,92			
YA333	5 µl	45 µl	–		0,284	4,60	0,92	0,84	6,7	5,63
	10 µl	40 µl	–		0,458	7,51	0,75			

After the protein quantification assay, and before the SDS-PAGE, two tubes were prepared in order to have a concentration of 0,2 mg/ml in a final volume of 80 µl. The following image (figure 15) describes the two different tubes of each protein in detail.

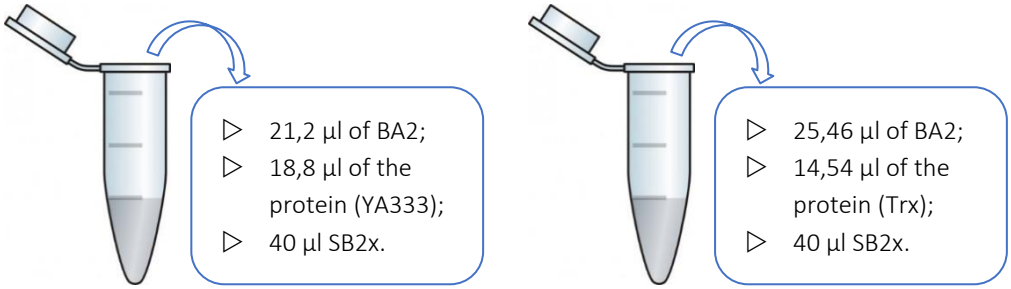

Figure 15. Description of the purified recombinant solutions analysed in the SDS-PAGE.

To visualise the purified proteins obtained, a SDS-PAGE gel was loaded with 7,5 µl and 15 µl of each protein (NF-YAl exon3 protein (TRX-6his-YA333) and the non-relevant control protein (TRX-6his)) in each lane, in order to obtain 1,5 and 3 ug of protein per lane, respectively (Figure 15). This SDS-PAGE was useful to control the concentration, quality and purity of each protein in the protein preparation.

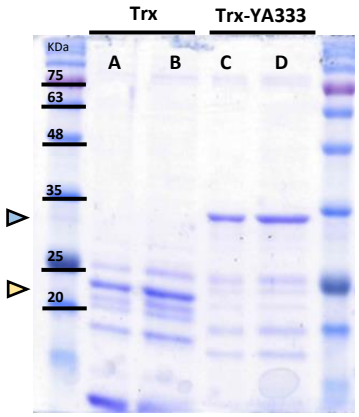

Figure 16. SDS-PAGE with NF-YAl exon3 (blue arrow) protein and thioredoxin (yellow arrow) control protein.

SDS-PAGE gel was loaded with protein ladder (MW indicated in kDa) in both extremities, 7,5 µl of the (A), and 15 µl (B) of the Trx (thioredoxin) control protein, 7,5 µl (C) and 15 µl (D) of the Trx-YA333 (NF-YAl exon3) recombinant protein (1,5 and 3 µg of purified protein per lanes A, C and B, D, respectively). Protein bands were revealed by Coomassie staining.

The results of the SDS-PAGE (figure 16) show that both protein preparations samples contain the desired recombinant protein, near to the expected molecular weight (20 kDa for Trx and 28 kDa for Trx-YA333). Although the Trx control displays different degradation products, the protein preparations were satisfactory, as the desired YA333 recombinant protein was obtained predominantly in its full-length form, with a good purity. For this, we esteemed that the purified samples were of good enough quality to proceed with the following step, consisting in the preparation of the matrix.

Another and last SDS-PAGE was executed, as a final test, with all of the four proteins relevant to the project, including the full-length NF-YA long and short isoform, the purified Trx-YA333 which includes the NF-YA exon 3 fragment, and the thioredoxin control (figure 17).

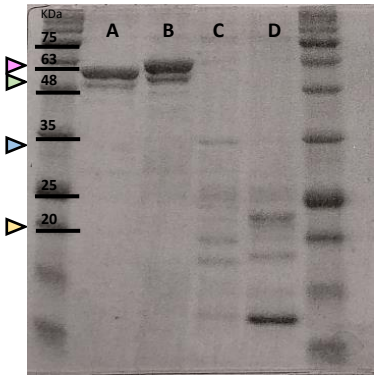

Figure 17. SDS-PAGE of NF-YAl exon3 protein and thioredoxin control protein purified samples to be used for Western Blotting.

SDS-PAGE gel loaded with protein ladder (expressed in kDa) in both extremities, and 3 µg of each protein the full-length short isoform of NF-YA (lacking exon 3) (A – green arrow), and long isoform of NF-YA (B – pink arrow), together with the Trx-YA333 (exon 3 of NF-YA) (C – blue arrow) and thioredoxin (D – yellow arrow) protein samples.

Although for some technical reason the recombinant proteins were not evenly stained, the SDS-PAGE came out with the expected results regarding the size of loaded proteins, indicating in this way, the identity and quality of the purified proteins to be used for Western blotting.

8.2 IMMUNOBLOTTING ANALYSES OF NF-YA POLYCLONAL ANTIBODIES

8.2.1 Analysis of rabbit sera for YAl -exon 3- specificity

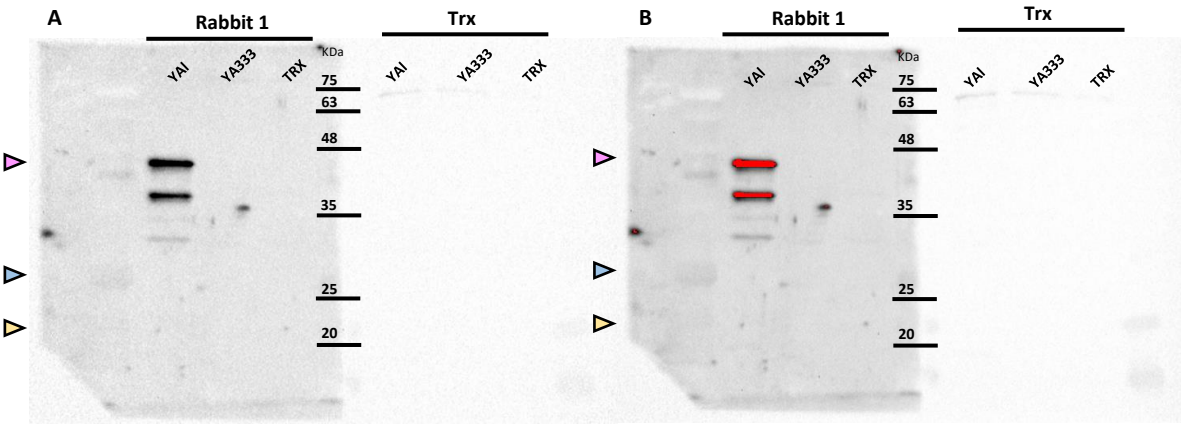

Figure 18. Western Blot of NF-YA long proteins with Rabbit 1 serum and the thioredoxin control Abs.
Recombinant protein samples were loaded on SDS-PAGE gels in duplicate lanes (From left to right, lanes corresponding to molecular weight, NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow)) and transferred to nitrocellulose membranes. The left part of the membrane was probed with Rabbit 1 anti-YA serum and the right part with control Abs to the Trx protein. Western Blot results at short exposure times (20,6 seconds) **(A)** and long exposure times (118,0 seconds) **(B)** are shown. Red coloring highlights saturated pixels in the long exposure image.

In the first Western Blot experiments, we started to analyse the specificity of the available sera from rabbit 1 and 2, which were immunised with the N-terminal region of YA, to the different types of proteins together with Antibodies raised to the control protein thioredoxin and control Antibodies to both isoforms of Nuclear Factor YA. With these analyses we wanted to evaluate if, by using the C-terminal portion of NF-YAl protein as antigen, the immunization of the two animals actually led to the production of antibodies specific to the exon 3 portion of the protein, as initially desired.

Figure 18 shows the results of the Western blot analysis performed with Rabbit 1 serum (left part) and control antibodies (right part). It is possible to observe that the presence of the long form of NF-YA (YAl) is evident in the membrane probed with anti-bodies from rabbit 1, as it was expected. In the same membrane, the presence of the exon 3 of NF-YA (YA333) recombinant protein is not as clear as it was predicted, however is observable a small band in the same localization as the expected MW. This might indicate that the Exon3 segment could be actually detected, but for some technical reason, a bubble was formed there, preventing the visualization of the full YA333 protein band. Relatively to the thioredoxin protein lane, it is not possible to visualise any signal. This might be due to the same technical reason as the one detailed previously.

On the membrane corresponding to the thioredoxin antibodies the overall signal is very low to non-existent. In this case, it was supposed to be detected only in the band of thioredoxin (TRX), and of the recombinant TRX-YA333. This did not happen possibly due to technical issues, like the available antibodies being not in the right storage solution conditions, making it lose activity.

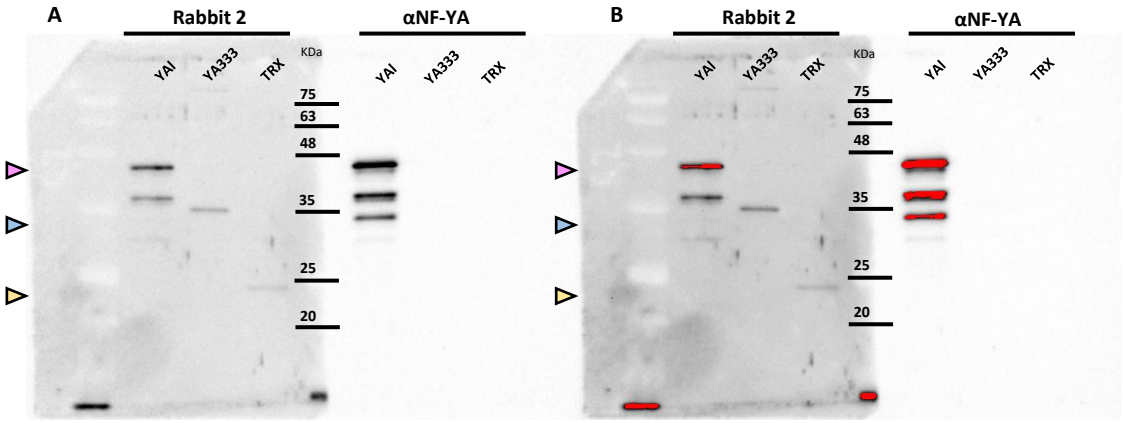

Figure 19. Western Blot of NF-YAl proteins with Rabbit 2 serum (on the left) and the short isoform of NF-YA control Abs (on the right).

Recombinant protein samples were loaded on SDS-PAGE gels in duplicate lanes (From left to right, bands corresponding to molecular weight, NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow)) and transferred to nitrocellulose membranes. Western Blot results at short exposure times (12,8 seconds of exposure) (A) and long exposure times (61,5 seconds of exposure) (B) are shown. Red colouring highlights saturated pixels in the long exposure image.

The results (figure 19) obtained with the rabbit 2 serum, and with NF-YAs specific Abs (also recognising the long isoform) are shown in Figure 12. The rabbit 2 Abs corresponding membrane showed the expected detection of NF-YAl, but on the other side, also presents a low-signal level of both exon 3 of NF-YA (Trx-YA333) and thioredoxin proteins. These evidences show a non-specific recognition of the exon 3 recombinant protein in the rabbit 2 model, possibly due to some Abs binding to the Trx-tag portion of the protein.

In the last membrane, corresponding to NF-YA Abs, it was expected to find a signal with the NF-YAl protein and not in the thioredoxin nor in the exon 3 YA lanes. It is observable that results corresponded to the expectations, as bands of NF-YAl were detected, and no signal was present for the thioredoxin and YA333 samples.

8.2.2 Validation of NF YAl exon 3 recognition by Rabbit sera

To verify the results obtained in the previous analyses, the following WB experiments NF-YA protein membranes were tested with the same conditions (equal samples – NF-YAl, exon 3 of NF-YA and thioredoxin – and western blot procedure). Additionally to rabbit 1 serum, was also tested a αHis Ab to detect the loaded recombinant proteins (figure 20).

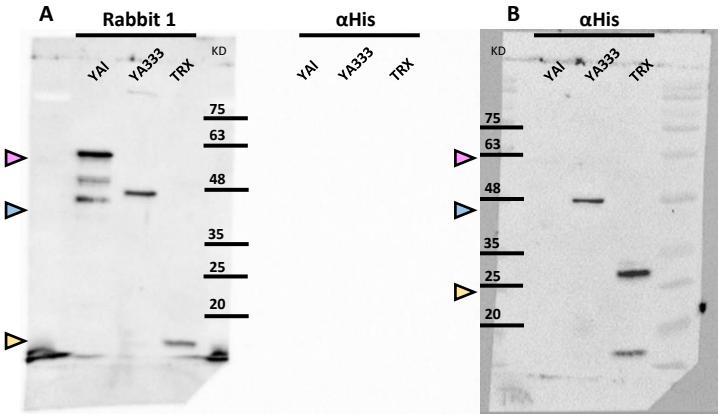

Figure 20. Western Blot analysis of NF-YA proteins with Rabbit 1 serum and αHis antibodies.

From left to right, lanes corresponding to molecular weight, NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow) are indicated. **(A)** *Western Blot with antibodies corresponding to the Rabbit 1 (on the left) and the* αHis *(on the right), at the time of 6,0 seconds of exposure*. **(B)** Western Blot *with antibodies corresponding only to the* αHis protein, at the time of 90,0 seconds of exposure, in a Multichannel image together with the molecular weight.

In the first membrane (decorated with rabbit 1 serum) NF-YAl is detected with comparable levels as the previous experiment, as expected. Relatively to the detection of exon 3 of NF-YA, while the recognition was not completely evident in the previous experiment, in this one YA333 was detected with a good recognition of the relative band. In the thioredoxin lane some signal was revealed, however at a lower MW.

αHis antibody detection (figure 20B) confirmed that the his-tagged proteins (YA333 and Trx) are present at similar levels, supporting the hypothesis that rabbit 1 serum can specifically recognise the YA exon 3 epitope. As expected, the YAl protein (untagged) is not recognised by anti-his tag Abs.

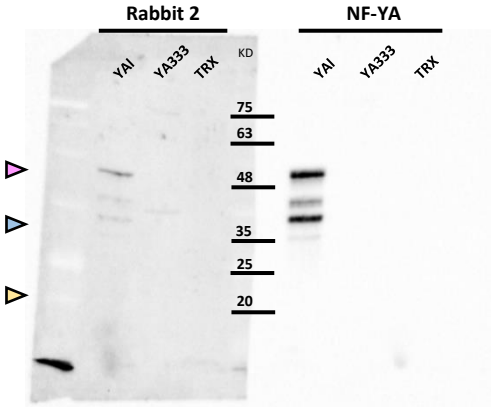

Figure 21. Western Blot analysis of NF-YA proteins with Rabbit 2 serum and NF-YA antibodies.

Western Blot analysis of YA recombinant proteins with Rabbit 2 serum antibodies (on the left) and the NF-YA (on the right), captured at the time of 6,0 seconds of exposure. From left to right, lanes corresponding to molecular weight, NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow) are indicated.

The rabbit 2 was also tested in the same conditions (figure 21) as the previous experiment, however not presenting conclusive results. Despite the low signal detected, it seems to have a recognition of both NF-YAl and the NF-YA exon 3 and not the thioredoxin, as predicted for the three bands, possibly indicating some specificity to YA exon 3, also for Rabbit 2 serum.

The membrane probed with anti-NF-YA antibodies confirmed the previous result with a strong signal corresponding to the NF-YAl band and no signal in the remaining two proteins lanes, as only NF-YAl was detected and not the exon 3 of NF-YA (YA333) or thioredoxin, which were expected to give no signal.

8.2.3 Analysis of Rabbit 1 affinity purified Abs

Following the previous western blot experiments, an affinity cromatography for antibody purification was made (see figure 22) starting with Rabbit 1 serum, and the respective fractions obtained were used to verify the Ab purification procedure.

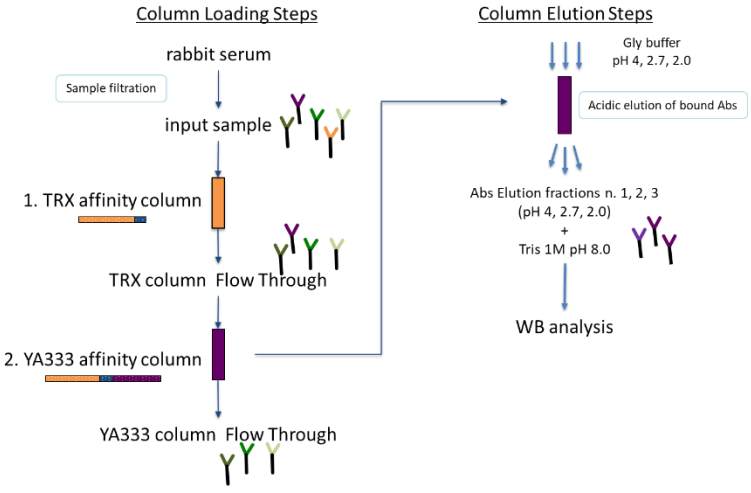

Figure 22. Purification strategy of NF-YA exon 3 antibodies.

Schematic representation of the affinity chromatography purification, with YA333 represented in purple, thioredoxin in orange and 6-His tag in blue.

Six membranes were prepared with samples from NF-YAl, exon 3 of NF-YA and thioredoxin proteins. The membranes were tested with different fractions obtained: the flow through of thioredoxin column (pre-clearing step) and of the YA333 column (capture step), with rabbit 1 serum (input sample), and the Ab elution fractions number 2 of each pH step (pH 4, 2.7 and 2) (Figure 23).

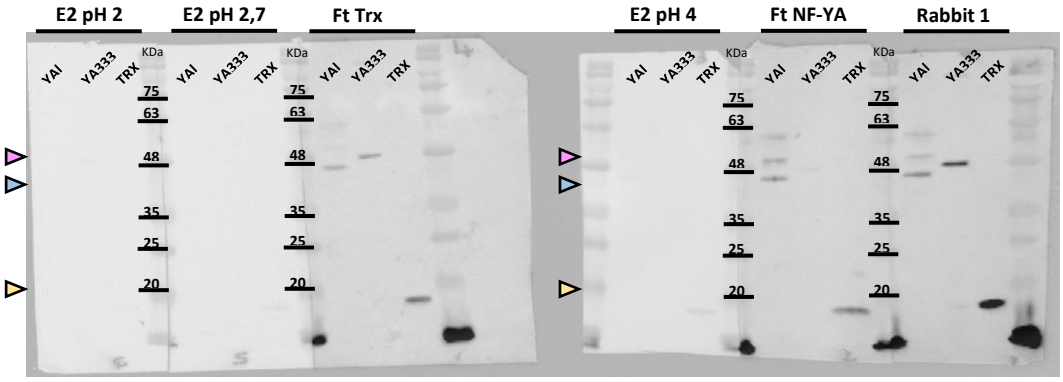

Figure 23. Western Blot with antibodies obtained in affinity chromatography purification of rabbit 1 serum, corresponding to the elution 2 of the pH 2 (E2 pH2), the elution 2 of the pH 2,7 (E2 pH 2,7), the flow through of the thioredoxin (Ft Trx), the elution 2 of the pH 4 (E2 pH4), flow through of the NF-YA (Ft NF-YA) and the Rabbit 1 (from left to right).

Western Blot images captured at the time of 60,0 seconds of exposure. From left to right, lanes corresponding to molecular weight, NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow) are shown for each membrane.

The mebrane decorated with the elution 2 of pH 2 (figure 23 membrane E2 pH 2) seems to be the one with best results of all, having enriched the serum fraction with Abs to this YA epitope. The desirable result would have been to detect the exon 3 of NF-YA (YA333 protein) and not the TRX, and at lower levels the YAl protein in remaining lanes. Despite the overall low signal, with this elution it is possible to detect a signal given by the exon 3 of NF-YA. Relatevily to the bands of NF-YAl and thioredoxin the opposite is observed, as desired, with no signal detected for Trx, and lower signal for full length YAl.

With elution 2 (at pH 2,7) the response starts to differ from the previous. Although a signal at the exon 3 of NF-YA is detected, a lower signal from NF-YAl, as expected, and a low signal at thioredoxin also appear, going against to what was predicted, indicating a non-specific component of the purified Abs (figure 23 membrane E2 pH 2,7).

In the flowthrough of thioredoxin was expected the relative lane band to disappear and not in the other two lanes relative to NF-YA. The signal from thioredoxin appears, but on the other hand, so does the remaining two, indicating that this step was not efficient for depleting non-specific antibodies as desired (figure 23 membrane Ft Trx).

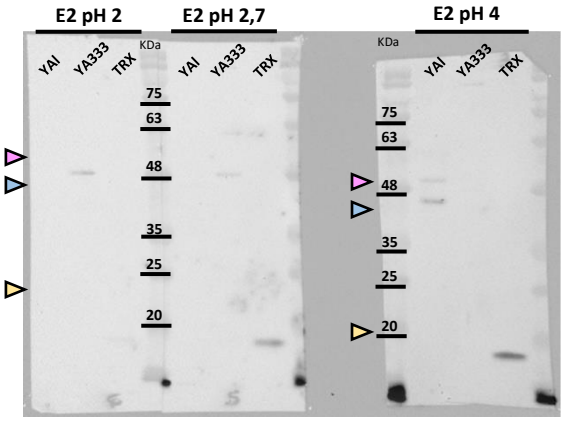

Figure 24. Western Blot with antibodies corresponding exclusively to the eluates.

Western Blot with antibodies corresponding, from left to right, to the elution 2 of the pH 2, the elution 2 of the pH 2,7, and the elution 2 of the pH 4, at the time of 300,0 seconds of exposure. From left to right, bands corresponding to NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow), thioredoxin (TRX – yellow arrow) and molecular weight.

Contrarily to what was expected and to the other 2 elutions (figure 24 membrane E2 pH 2 and E2 pH 2,7), the elution 2 of pH 4 (figure 24 membrane E2 pH 4) presents negative results. In summary, a low signal band of NF-YAl appears, a stronger signal of thioredoxin also surfaces and no signal of the exon 3 of NF-YA is detected.

The flowthrough of YA333 column (figure 23 membrane Ft NF-YA) is expected to have a signal in NF-YAl (for NF-YA Abs not recognising exon 3), lower levels (or absence) of the exon 3 of NF-YA band and no signal of thioredoxin. Even though the first two occur as predicted, a non desirable signal in thioredoxin is noticed.

In the rabbit 1 membrane (figure 23), the results were supposed to be the same as the expected for the flowthrough of NF-YA, positive for NF-YAl and exon 3 of NF-YA and negative to thioredoxin. Unfortuntely the results were the same as the previously obtained and not the expected.

8.2.4 WB analysis of affinity chromatography pH 2 elution fractions

Considering the results obtained with the different elution steps of the Ab purification procedure, the fractions 2 and 3 of pH 2 elution were mixed and concentrated together, since it seemed to have better and similar results. In this WB experiment pH 2 mixed fractions were tested comparing them to the rabbit 1 serum, the clarified input of thioredoxin column, the concentrated elution 2 and 3 of pH 2, the elution 2 of pH 4, αHisTag mouse monoclonal Abs and the flow through of the NF-YA333 column (Figure 25).

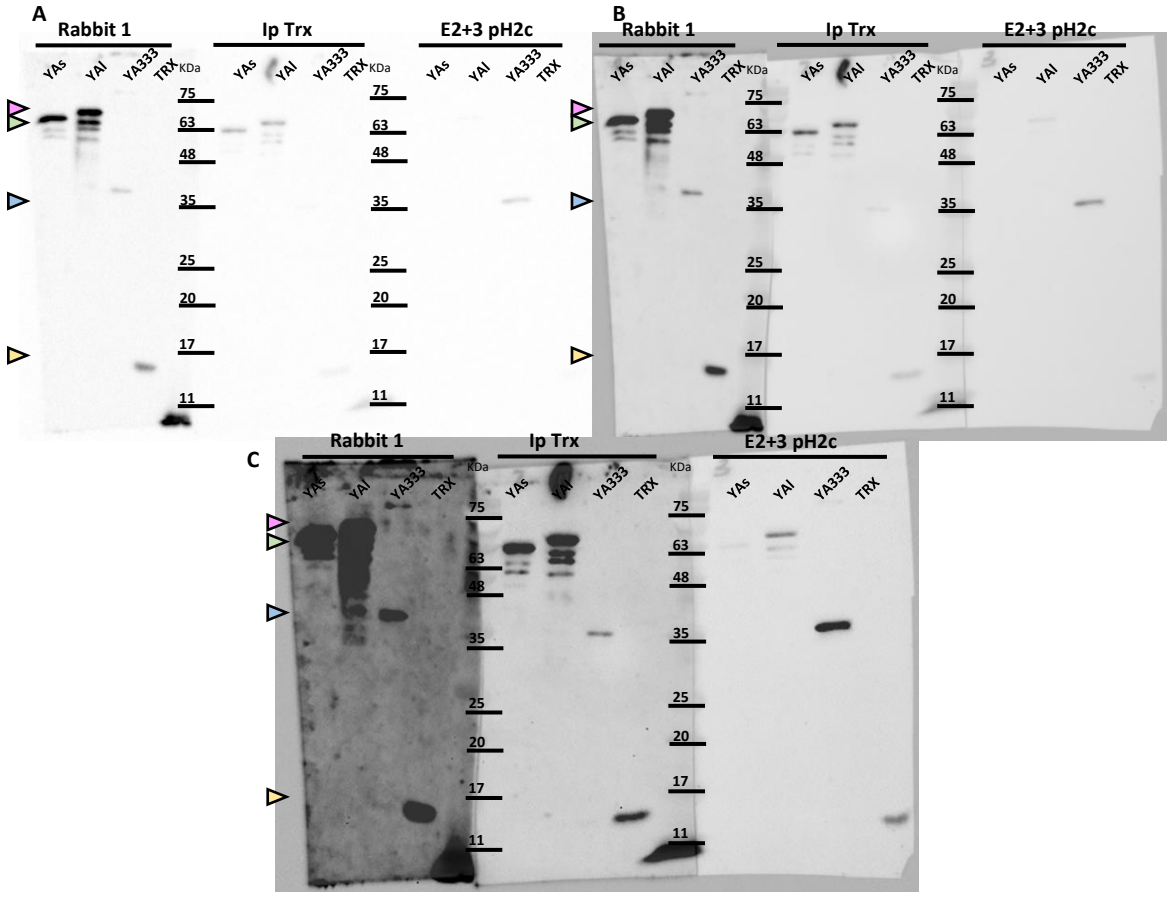

Figure 25. Western Blot decorated with antibodies corresponding (from left to right) to the rabbit 1 serum, input of thioredoxin column (ip Trx) and the concentrated the elutions 2 and 3 of pH 2 (E2+3 pH2c).

From left to right, bands corresponding to molecular weight, NF-YA short (YAs – green arrow), NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow). **(A)** Western Blot images captured at the time of 3,0 seconds of exposure. **(B)** Western Blot images captured at the time of 34,3 seconds of exposure, in a Multichannel image with the molecular weight. **(C)** Western Blot images captured at the time of 300,0 seconds of exposure, also in a Multichannel image with the molecular weight.

As described before, in the rabbit 1 serum decorated membrane is predicted to have a signal in the samples prevenient from NF-YA, being NF-YAs, NF-YAl and the exon 3 of NF-YA samples, and no signal of thioredoxin. Although the positives results appeared, the thioredoxin signal also did, although at lower MW.

In the thioredoxin column input membrane is predicted a similar signal as in the serum in the thioredoxin and no other signal in the remaining samples. Unexpectedly it was identified a signal in all samples, which translates in only one positive result, the signal in thioredoxin.

The sample correspondent to the concentrate of elution 2 and 3 of pH 2 is expected to detect NF-YAl and the exon 3 of NF-YA, and contrarily, to find no signal of NF-YAs and thioredoxin. The positive signal of the NF-YAl and exon 3 of NF-YA was successfully detected. It was not distinguished a signal of thioredoxin as expected. On the other hand, a non-desired low signal of NF-YAs appeared, indicating some cross-reactivity.

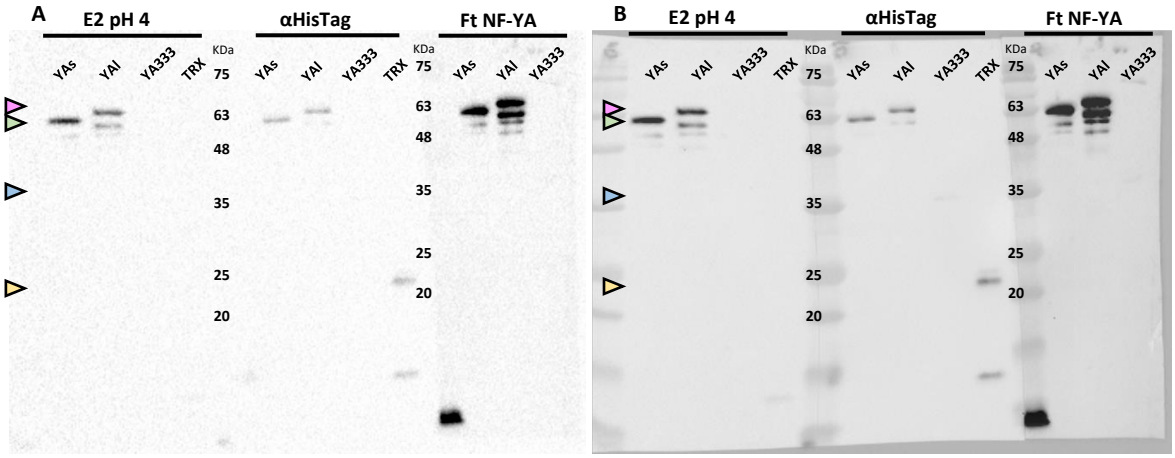

Figure 26. Western Blot analysis with antibodies corresponding (from lef to right) to the second elution of the pH 4 (E2 pH 4), mouse αHisTag Abs (αHisTag), and the flowthrough of the NF-YA column (ft NF-YA).

From left to right, lanes corresponding to molecular weight, 6His-NF-YA short (YAs – green arrow), 6His-NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow) samples. **(A)** Western Blot at the time of 3,0 seconds of exposure. **(B)** Western Blot at the time of 81,2 seconds of exposure, in a Multichannel image with the molecular weight.

The membrane of the second elution at pH 4 (figure 26) is, like the other elution fractions of the different pH, expected to present a signal in both NF-YAl and the exon 3 of NF-YA, and no signal in NF-YAs and thioredoxin. Some of the results expected were visible, like the absence of thioredoxin signal and the presence of NF-YAl. Discordantly, a signal of NF-YAs appeared but not the exon 3 of NF-YA. This indicates that, at this pH 4, it was not possible to obtain the antibodies correspondent to the exon 3 of NF-YA.

Probing of membranes was used to evaluate equivalent loading of protein samples, since in this case all proteins display the αHistidine-tag. As observed in part B of figure 26 (long exposure image), full length YA proteins and the TRX control were loaded with similar amounts, while the YA333 protein resulted to be less than expected.

As described before, the NF-YA flow through membrane is predicted to have a signal in all the bands, which correspond to NF-YA full length forms (NF-YAl, NF-YAs and not the exon 3 of NF-YA). This happens in the first two samples, but not with the exon 3 of NF-YA, confirming that the capture step on the YA333 column was achieved.

8.2.5 WB analysis of NF-YAl Ab affinity purification fractions

To conclude the analysis of fractions obtained with YAl affinity purification of rabbit 1 serum Abs, in this western blot experiment six different membranes were tested with the following fractions the rabbit 1 serum, again rabbit 1 serum but before immunization (pre-immune rabbit 1), the flow-through of NF-YA column, the elution 2 and 3 mixed together of pH 2,7, the elution 1 of pH 2 concentrated and the elution 2 and 3 (mixed and concentrated) of pH 2.

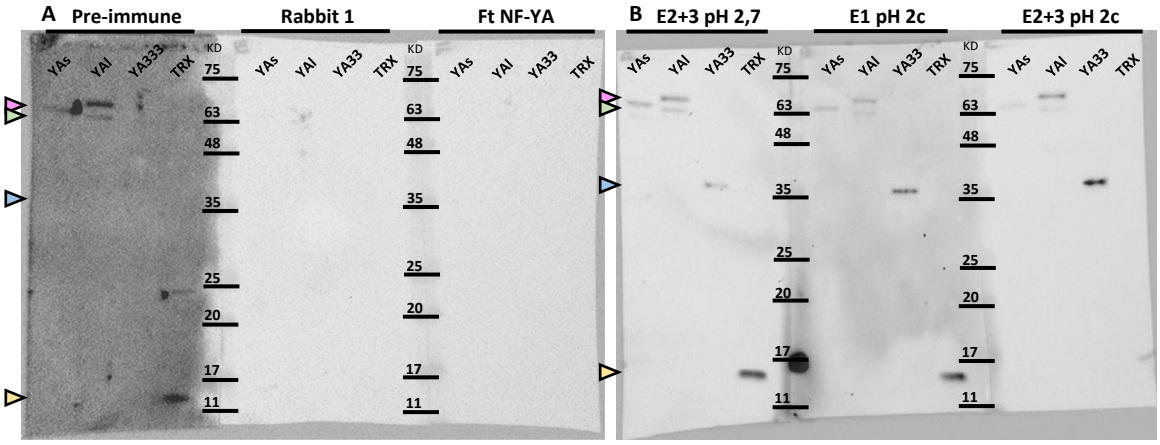

Figure 27. Western Blot with antibodies corresponding to the pre-immune rabbit 1, Rabbit 1 and flow through of NF-YA, and also to the 2nd and 3rd elution of the pH 2,7, the 1st elution of the pH 2 concentrated and the 2nd and 3rd elution of the pH 2 concentrated.

Western Blot with bands corresponding, from left to right, to molecular weight (in KDa), NF-YA short (YAs – green arrow), NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow) and thioredoxin (TRX – yellow arrow). **(A)** Western Blot with antibodies corresponding respectively to the pre-immune rabbit 1 (pre-immune), Rabbit 1 and flow through of NF-YA (ft NF-YA), at the time of 14,4 seconds of exposure. **(B)** Western Blot in a Multichannel image with the molecular weight, with antibodies corresponding respectively to the 2nd and 3rd elution of the pH 2,7 (E2+3 pH 2,7), the 1st elution of the pH 2 concentrated (E1 pH 2c) and the 2nd and 3rd elution of the pH 2 concentrated (E2+3 pH 2c) (at the time of 300,0 seconds of exposure).

The membrane of the pre-immune rabbit 1 is predicted to show no signal of NF-YAl and the exon 3 of NF-YA and to detect it from NF-YAs and thioredoxin. The membrane shows a signal in every sample, with an exception in the exon 3 of NF-YA, also showing high background, indicating non-specific binding (figure 27A).

In the next two membranes (rabbit 1 and flow through of NF-YA) the results expected are the same as in the previous experiments. However, there was no signal detected, in any of the samples in both membrane, this may indicate an over-use of the Abs solutions.

In the remaining membranes (figure 27B) the results expected are the same as before. It is hoped that the three membranes will detect the long isoform of NF-YA and its exon 3, but not the NF-YA short isoform and thioredoxin. These results are positive to the elution 2 and 3 of pH 2 with a good signal of NF-YAl and the exon 3 and no signal to the remaining samples. In the elution 2 and 3 of pH 2,7 membrane the results differ a little since the signal of NF-YAl and exon 3 are lower and a signal in both NF-YAs and thioredoxin is detected. The membrane with the elution 1 of pH 2, starts showing a stronger signal of NF-YAl and exon 3, and lower signal of NF-YAs and thioredoxin. Yet the

results of this last elution are not as good as the expected, and as the elution 2 and 3 of the same pH.

8.2.6 Analysis of NF-YA proteins expressed in mammalian cells with affinity purified Abs to YAl

In this experiment different samples were tested, in order to verify the results obtained until now, using instead full-length proteins as expressed in mammalian cells, such as endogenous proteins in samples of not transfected cells, and NF-YAs or NF-YAl overexpressed in transfected cells. These samples were tested in two membranes, one with αYA short isoform Abs and the other with the elution 2 and 3 of pH 2 concentrated. Additionally, another membrane (elution 2 and 3, pH 2, concentrated) was also tested with samples of NF-YA long and the NF-YA exon 3 recombinant proteins (figure 28).

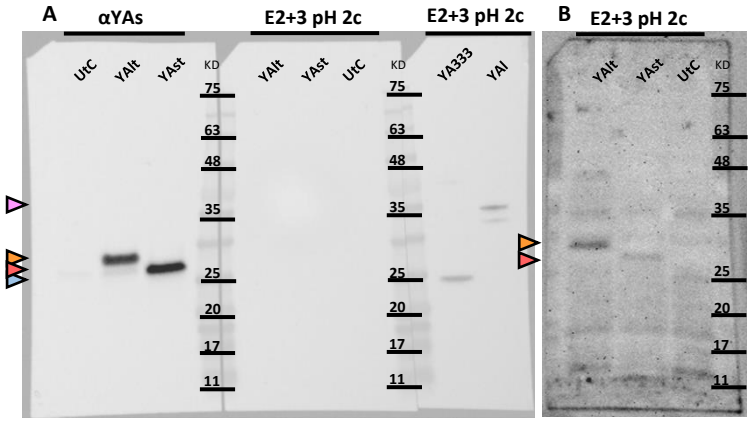

Figure 28. Western Blot analysis of mammalian cells extracts decorated with antibodies corresponding to the second and third elution of the pH 2 concentrated (E2+3 pH 2c), and αYA short isoform (αYAs).

Different bands corresponding to molecular weight (in KDa), NF-YA long (YAl – pink arrow), exon 3 of NF-YA (YA333 – blue arrow), NF-YA long of transfected cells (YAlt – orange arrow), NF-YA short of transfected cells (YAst – red arrow) and untransfected control cells (UtC). **(A)** Western Blot in a Multichannel image with the molecular weight at the time of 7,1 seconds of exposure. From left to right, 2nd and 3rd elution of the pH 2 concentrated, again the 2nd and 3rd elution of the pH 2 concentrated and lastly the αYA short isoform. **(B)** Western Blot corresponding exclusively to the αYA short isoform at the time of 1500,0 seconds of exposure, in a Multichannel image with the molecular weight.

In the membrane corresponding to the αYA short isoform it is supposed to see the signal of both transfected of NF-YAs and NF-YAl, and not observe any indication of not transfected cells. The result is the expected.

Relatively to the membrane of the elution (elution 2 and 3, at pH 2, concentrated), with the same samples as the previous membrane, it is expected a signal in the sample of NF-YAl transfected cells exclusively. Despite the low signal, comparatively with the previous membrane, the result was positive (see figure 28B).

In the last membrane, of the elution with different samples, the results awaited are the same as before, with a detection of signal in both samples (NF-YAl and NF-YA exon 3). Once again, the results were positive, corresponding to what was anticipated.

In addition, the first two membranes of the previous western blot were stained with *α tubulin*.

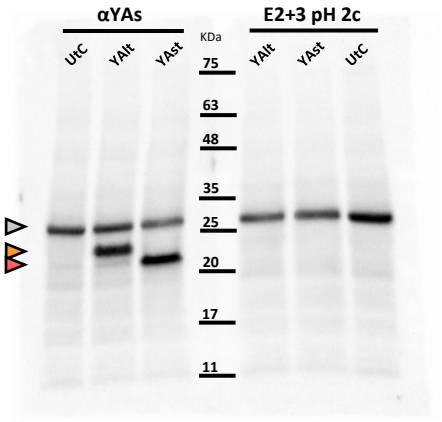

Figure 29. Western Blot of mammalian cells extracts stained with α tubulin.

Western Blot from figure 9 (A), stained with α tubulin (grey arrow) at the time of 17,7 seconds of exposure. Previously stained with antibodies corresponding to the 2nd and 3rd elution of the pH 2 concentrated. Different bands corresponding to molecular weight (ranging from 75 to 11 KDa), NF-YA long of transfected cells (YAlt – orange arrow), NF-YA short of transfected cells (YAst – red arrow) and untransfected control cells (UtC) lanes are indicated.

Seeing that the *α tubulin* is detected in both membranes in the same and correct levels (figure 29), it is possible to conclude that the low results acquired before were accurate, possibly indicating that low amounts of Abs were obtained with our purification procedure.

8.2.7 WB Analysis of different blocking solutions for affinity purified Abs to NF-YAl

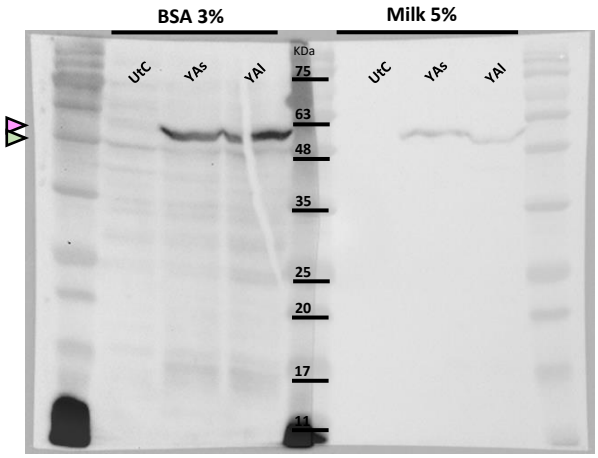

Figure 30. Western Blot with antibodies corresponding to the 2nd and 3rd elution of the pH 2 concentrated, with different blocking solutions.

Western Blot in a Multichannel image with antibodies corresponding to the second and third elution of the pH 2 concentrated, with different blocking solutions – BSA 3% at the left and Milk 5% at right– at the time of 95,7 seconds of exposure. Bands corresponding to molecular weight (in KDa), untransfected control cells (UtC), NF-YA short (YAs – green arrow) and NF-YA long (YAl – pink arrow).

As a last test, an experiment was made testing two different blocking solutions. As the previous experiments came out with low signal, and not always well perceived, a different blocking solution was tested in order to verify if the results were clearer. The two membranes were tested with the same samples, but while the first was blocked with BSA (3%), the other was blocked with milk (5%), as the previous experiments (figure 30).

As expected, the results where similar, although showing a better and more perceiving signal in the membrane blocked with BSA. This translates in the BSA being a better choice as blocking solution in these experiments, since there is no alteration of the results, but just an amplification of the signal.

CHAPTER 4

CONCLUSIONS

9 CONCLUSIONS

In this project, we analysed and purified polyclonal antibodies derived from rabbits immunised with the YAl isoform of NF-YA. During the beginning of the project, it was performed an affinity purification strategy, using a recombinant protein. This strategy led to a successful purification and production of the protein that included the protein portion of NF-YAl encoded by exon 3, typical of this subunit.

The different tests and 8 purifications done throughout the first part of the project permitted to optimize the protocol, in a way to turn it more efficient in the usage of time and solutions. However, there is still space for further optimizations, that could lead to obtain bigger volumes of the protein in less time.

The affinity purification, from the first part of the project, allowed us to isolate the Ab fractions that specifically recognise the YAl isoform of NF-YA, having the exon lacking on the NF-YAs isoform. The results showed the correct recognition of the isoform intended, despite the initial limited quantity.

Data from these experiments demonstrated that the project had a successful outcome and that with further optimizations it will be possible to isolate this valuable reagent with improved and sizeable yields. This reagent can be used in further experiments to analyse the splicing isoforms of NF-YA, facilitating the comprehension of the specific mechanisms behind the distinct functions of the different isoforms of this important transcription factor.

10 BIBLIOGRAPHY

Andrew, S. M., Titus, J. A., & Zumstein, L. (2002). Dialysis and concentration of protein solutions. *Current Protocols in Toxicology*, *Appendix 3*, A.3H.1-5. https://doi.org/10.1002/0471140856.txa03hs10

Arndt, C., Koristka, S., Feldmann, A., Bartsch, H., & Bachmann, M. (2012). Coomassie-Brilliant Blue staining of polyacrylamide gels. *Methods in Molecular Biology (Clifton, N.J.)*, *869*, 465–469. https://doi.org/10.1007/978-1-61779-821-4_40

Arnold, C. D., Nemčko, F., Woodfin, A. R., Wienerroither, S., Vlasova, A., Schleiffer, A., Pagani, M., Rath, M., & Stark, A. (2018). A high-throughput method to identify trans-activation domains within transcription factor sequences. *The EMBO Journal*, *37*(16), e98896. https://doi.org/10.15252/embj.201798896

Arora, S., Saxena, V., & Ayyar, B. V. (2017). Affinity chromatography: A versatile technique for antibody purification. *Methods (San Diego, Calif.)*, *116*, 84–94. https://doi.org/10.1016/j.ymeth.2016.12.010

Basile, V., Baruffaldi, F., Dolfini, D., Belluti, S., Benatti, P., Ricci, L., Artusi, V., Tagliafico, E., Mantovani, R., Molinari, S., & Imbriano, C. (2016). NF-YA splice variants have different roles on muscle differentiation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1859*(4), 627–638. https://doi.org/https://doi.org/10.1016/j.bbagrm.2016.02.011

Benatti, P., Basile, V., Merico, D., Fantoni, L. I., Tagliafico, E., & Imbriano, C. (2008). A balance between NF-Y and p53 governs the pro- and anti-apoptotic transcriptional response. *Nucleic Acids Research*, *36*(5), 1415–1428. https://doi.org/10.1093/nar/gkm1046

Bernardini, A., Lorenzo, M., Nardini, M., Mantovani, R., & Gnesutta, N. (2019). The phosphorylatable Ser320 of NF-YA is involved in DNA binding of the NF-Y trimer. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *33*(4), 4790–4801. https://doi.org/10.1096/fj.201801989R

Bezzecchi, E., Ronzio, M., Dolfini, D., & Mantovani, R. (2019). NF-YA Overexpression in Lung Cancer: LUSC. *Genes*, *10*(11), 937. https://doi.org/10.3390/genes10110937

Bezzecchi, E., Ronzio, M., Semeghini, V., Andrioletti, V., Mantovani, R., & Dolfini, D. (2020). NF-YA

Overexpression in Lung Cancer: LUAD. *Genes*, *11*(2), 198. https://doi.org/10.3390/genes11020198

Bhattacharya, A., Deng, J. M., Zhang, Z., Behringer, R., de Crombrugghe, B., & Maity, S. N. (2003). The B subunit of the CCAAT box binding transcription factor complex (CBF/NF-Y) is essential for early mouse development and cell proliferation. *Cancer Research*, *63*(23), 8167–8172.

Bi, W., Wu, L., Coustry, F., de Crombrugghe, B., & Maity, S. N. (1997). DNA binding specificity of the CCAAT-binding factor CBF/NF-Y. *The Journal of Biological Chemistry*, *272*(42), 26562–26572. https://doi.org/10.1074/jbc.272.42.26562

Bisikirska, B., Bansal, M., Shen, Y., Teruya-Feldstein, J., Chaganti, R., & Califano, A. (2016). Elucidation and Pharmacological Targeting of Novel Molecular Drivers of Follicular Lymphoma Progression. *Cancer Research*, *76*(3), 664–674. https://doi.org/10.1158/0008-5472.CAN-15-0828

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., & Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947–956. https://doi.org/10.1016/j.cell.2005.08.020

Brunelle, J. L., & Green, R. (2014). Chapter Twelve - One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). In J. B. T.-M. in E. Lorsch (Ed.), *Laboratory Methods in Enzymology: Protein Part C* (Vol. 541, pp. 151–159). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-420119-4.00012-4

Ceribelli, M., Benatti, P., Imbriano, C., & Mantovani, R. (2009). NF-YC Complexity Is Generated by Dual Promoters and Alternative Splicing. *The Journal of Biological Chemistry*, *284*, 34189–34200. https://doi.org/10.1074/jbc.M109.008417

Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M., & Zaret, K. S. (2002). Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. *Molecular Cell*, *9*(2), 279–289. https://doi.org/https://doi.org/10.1016/S1097-2765(02)00459-8

Dagert, M., & Ehrlich, S. D. (1979). Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene*, *6*(1), 23–28. https://doi.org/10.1016/0378-1119(79)90082-9

Dolfini, D., Andrioletti, V., & Mantovani, R. (2019). Overexpression and alternative splicing of NF-YA in breast cancer. *Scientific Reports*, *9*(1), 12955. https://doi.org/10.1038/s41598-019-49297-5

Dolfini, D., Gatta, R., & Mantovani, R. (2012). NF-Y and the transcriptional activation of CCAAT promoters. *Critical Reviews in Biochemistry and Molecular Biology*, *47*(1), 29–49. https://doi.org/10.3109/10409238.2011.628970

Dolfini, D., Minuzzo, M., Pavesi, G., & Mantovani, R. (2012). The short isoform of NF-YA belongs to the embryonic stem cell transcription factor circuitry. *Stem Cells (Dayton, Ohio)*, *30*(11), 2450–2459. https://doi.org/10.1002/stem.1232

Dolfini, D., Minuzzo, M., Sertic, S., & Mantovani, R. (2020). NF-YA overexpression protects from glutamine deprivation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1867*(2), 118571. https://doi.org/https://doi.org/10.1016/j.bbamcr.2019.118571

Dolfini, D., Zambelli, F., Pavesi, G., & Mantovani, R. (2009). A perspective of promoter architecture from the CCAAT box. *Cell Cycle (Georgetown, Tex.)*, *8*, 4127–4137. https://doi.org/10.4161/cc.8.24.10240

Dolfini, D., Zambelli, F., Pedrazzoli, M., Mantovani, R., & Pavesi, G. (2016). A high definition look at the NF-Y regulome reveals genome-wide associations with selected transcription factors. *Nucleic Acids Research*, *44*(10), 4684–4702. https://doi.org/10.1093/nar/gkw096

Fulton, D. L., Sundararajan, S., Badis, G., Hughes, T. R., Wasserman, W. W., Roach, J. C., & Sladek, R. (2009). TFCat: the curated catalog of mouse and human transcription factors. *Genome Biology*, *10*(3), R29. https://doi.org/10.1186/gb-2009-10-3-r29

Greenberg, D. S., & Soreq, H. (2013). *Alternative Splicing* (S. Maloy & K. B. T.-B. E. of G. (Second E. Hughes (eds.); pp. 97–98). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-374984-0.00043-7

Hage, D. S., Anguizola, J. A., Bi, C., Li, R., Matsuda, R., Papastavros, E., Pfaunmiller, E., Vargas, J., & Zheng, X. (2012). Pharmaceutical and biomedical applications of affinity chromatography: recent trends and developments. *Journal of Pharmaceutical and Biomedical Analysis*, *69*, 93–105. https://doi.org/10.1016/j.jpba.2012.01.004

Hnasko, T. S., & Hnasko, R. M. (2015). The Western Blot. *Methods in Molecular Biology (Clifton, N.J.)*, *1318*, 87–96. https://doi.org/10.1007/978-1-4939-2742-5_9

Huber, E. M., Scharf, D. H., Hortschansky, P., Groll, M., & Brakhage, A. A. (2012). DNA minor groove sensing and widening by the CCAAT-binding complex. *Structure (London, England : 1993)*, *20*(10), 1757–1768. https://doi.org/10.1016/j.str.2012.07.012

Imbriano, C., Gnesutta, N., & Mantovani, R. (2012). The NF-Y/p53 liaison: well beyond repression. *Biochimica et Biophysica Acta*, *1825*(2), 131–139. https://doi.org/10.1016/j.bbcan.2011.11.001

Kolomeisky, A. B. (2011). Physics of protein–DNA interactions: mechanisms of facilitated target search. *Physical Chemistry Chemical Physics*, *13*(6), 2088–2095. https://doi.org/10.1039/C0CP01966F

Kurien, B. T., & Scofield, R. H. (2003). Protein blotting: a review. *Journal of Immunological Methods*, *274*(1), 1–15. https://doi.org/https://doi.org/10.1016/S0022-1759(02)00523-9

Laloum, T., De Mita, S., Gamas, P., Baudin, M., & Niebel, A. (2013). CCAAT-box binding transcription factors in plants: Y so many? *Trends in Plant Science*, *18*(3), 157–166. https://doi.org/10.1016/j.tplants.2012.07.004

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, *172*(4), 650–665. https://doi.org/10.1016/j.cell.2018.01.029

Latchman, D. S. (1996). Transcription-Factor Mutations and Disease. *New England Journal of Medicine*, *334*(1), 28–33. https://doi.org/10.1056/NEJM199601043340108

Latchman, D. S. (1997). Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology*, *29*(12), 1305–1312. https://doi.org/https://doi.org/10.1016/S1357-2725(97)00085-X

Li, M., Li, G., Liu, W., Dong, X., & Zhang, A. (2019). Genome-wide analysis of the NF-Y gene family in peach (Prunus persica L.). *BMC Genomics*, *20*(1), 612. https://doi.org/10.1186/s12864-019-5968-7

Li, X. Y., Hooft van Huijsduijnen, R., Mantovani, R., Benoist, C., & Mathis, D. (1992). Intron-exon organization of the NF-Y genes. Tissue-specific splicing modifies an activation domain. *The Journal of Biological Chemistry*, *267*(13), 8984–8990.

Libetti, D., Bernardini, A., Sertic, S., Messina, G., Dolfini, D., & Mantovani, R. (2020). The Switch

from NF-YAl to NF-YAs Isoform Impairs Myotubes Formation. *Cells*, *9*(3). https://doi.org/10.3390/cells9030789

Ly, L. L., Yoshida, H., & Yamaguchi, M. (2013). Nuclear transcription factor Y and its roles in cellular processes related to human disease. *American Journal of Cancer Research*, *3*(4), 339–346. https://pubmed.ncbi.nlm.nih.gov/23977444

Mahmood, T., & Yang, P.-C. (2012). Western blot: technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, *4*(9), 429–434. https://doi.org/10.4103/1947-2714.100998

Maity, S N, & de Crombrugghe, B. (1998). Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends in Biochemical Sciences*, *23*(5), 174–178. https://doi.org/10.1016/s0968-0004(98)01201-8

Maity, Sankar N. (2017). NF-Y (CBF) regulation in specific cell types and mouse models. *Biochimica et Biophysica Acta. Gene Regulatory Mechanisms*, *1860*(5), 598–603. https://doi.org/10.1016/j.bbagrm.2016.10.014

Manni, I., Mazzaro, G., Gurtner, A., Mantovani, R., Haugwitz, U., Krause, K., Engeland, K., Sacchi, A., Soddu, S., & Piaggio, G. (2001). NF-Y mediates the transcriptional inhibition of the cyclin B1, cyclin B2, and cdc25C promoters upon induced G2 arrest. *The Journal of Biological Chemistry*, *276*(8), 5570–5576. https://doi.org/10.1074/jbc.M006052200

Mantovani, R. (1998). A survey of 178 NF-Y binding CCAAT boxes. *Nucleic Acids Research*, *26*(5), 1135–1143. https://doi.org/10.1093/nar/26.5.1135

Mantovani, Roberto. (1999). The molecular biology of the CCAAT-binding factor NF-Y. *Gene*, *239*(1), 15–27. https://doi.org/https://doi.org/10.1016/S0378-1119(99)00368-6

Mitchell, P. J., & Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science (New York, N.Y.)*, *245*(4916), 371–378. https://doi.org/10.1126/science.2667136

Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos P., G., & Eliopoulos, E. (2020). Transcription factors and evolution: An integral part of gene expression (Review). *World Acad Sci J*, *2*(1), 3–8. https://doi.org/10.3892/wasj.2020.32

Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D.,

Romier, C., Bolognesi, M., & Mantovani, R. (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. *Cell*, *152*(1–2), 132–143. https://doi.org/10.1016/j.cell.2012.11.047

Nardone, V., Chaves-Sanjuan, A., & Nardini, M. (2017). Structural determinants for NF-Y/DNA interaction at the CCAAT box. *Biochimica et Biophysica Acta. Gene Regulatory Mechanisms*, *1860*(5), 571–580. https://doi.org/10.1016/j.bbagrm.2016.09.006

Oestreich, K. J., & Weinmann, A. S. (2012). Master regulators or lineage-specifying? Changing views on CD4+ T cell transcription factors. In *Nature reviews. Immunology* (Vol. 12, Issue 11, pp. 799–804). https://doi.org/10.1038/nri3321

Oldfield, A. J., Henriques, T., Kumar, D., Burkholder, A. B., Cinghu, S., Paulet, D., Bennett, B. D., Yang, P., Scruggs, B. S., Lavender, C. A., Rivals, E., Adelman, K., & Jothi, R. (2019). NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. *Nature Communications*, *10*(1), 3072. https://doi.org/10.1038/s41467-019-10905-7

Oldfield, A. J., Yang, P., Conway, A. E., Cinghu, S., Freudenberg, J. M., Yellaboina, S., & Jothi, R. (2014). Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. *Molecular Cell*, *55*(5), 708–722. https://doi.org/10.1016/j.molcel.2014.07.005

Petroni, K., Kumimoto, R. W., Gnesutta, N., Calvenzani, V., Fornari, M., Tonelli, C., Holt 3rd, B. F., & Mantovani, R. (2012). The promiscuous life of plant NUCLEAR FACTOR Y transcription factors. *The Plant Cell*, *24*(12), 4777–4792. https://doi.org/10.1105/tpc.112.105734

Rodriguez, E. L., Poddar, S., Iftekhar, S., Suh, K., Woolfork, A. G., Ovbude, S., Pekarek, A., Walters, M., Lott, S., & Hage, D. S. (2020). Affinity chromatography: A review of trends and developments over the past 50 years. *Journal of Chromatography B*, *1157*, 122332. https://doi.org/https://doi.org/10.1016/j.jchromb.2020.122332

Romier, C., Cocchiarella, F., Mantovani, R., & Moras, D. (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *The Journal of Biological Chemistry*, *278*, 1336–1345. https://doi.org/10.1074/jbc.M209635200

Thomas, M. C., & Chiang, C.-M. (2006). The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology*, *41*(3), 105–178.

https://doi.org/10.1080/10409230600648736

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. In *Nature reviews. Genetics* (Vol. 10, Issue 4, pp. 252–263). https://doi.org/10.1038/nrg2538

Yamaguchi-Shinozaki, K., & Shinozaki, K. (2005). Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science*, *10*(2), 88–94. https://doi.org/https://doi.org/10.1016/j.tplants.2004.12.012

Zambelli, F., & Pavesi, G. (2017). Genome wide features, distribution and correlations of NF-Y binding sites. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1860*(5), 581–589. https://doi.org/https://doi.org/10.1016/j.bbagrm.2016.10.007

Zaret, K. S., & Mango, S. E. (2016). Pioneer transcription factors, chromatin dynamics, and cell fate control. *Current Opinion in Genetics & Development*, *37*, 76–81. https://doi.org/https://doi.org/10.1016/j.gde.2015.12.003

Zhu, J., Zhang, Y., Joe, G. J., Pompetti, R., & Emerson, S. G. (2005). NF-Ya activates multiple hematopoietic stem cell (HSC) regulatory genes and promotes HSC self-renewal. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(33), 11728 LP – 11733. https://doi.org/10.1073/pnas.0503405102

Zinzen, R. P., Girardot, C., Gagneur, J., Braun, M., & Furlong, E. E. M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature*, *462*(7269), 65–70. https://doi.org/10.1038/nature08531