# UNIVERSITÉ DE STRASBOURG

## ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ – ED 414

# THÈSE

**Présentée par**


## Stefano MORO


**Soutenue le : 06 avril 2022**

**Pour obtenir le grade de : Docteur de l'Université de Strasbourg**

**Discipline : Sciences de la Vie**

**Spécialité : Aspects moléculaires et cellulaires de la biologie**

# Caractérisation des mécanismes d'interaction de la kinase IKK avec ses substrats cellulaires


**Unité de Recherche : UMR 7242 (Illkirch-Graffenstaden)**
**Équipe Signalisation Cellulaire et Cancer**


**Thèse dirigée par :**

| | |
|---|---|
| **Dr. Katia ZANIER** | Chargé de recherches, Biotechnologie et Signalisation Cellulaire (UMR 7242) |

**Rapporteurs :**

| | |
|---|---|
| **Dr. Fabrice AGOU** | Directeur de recherches, Institut Pasteur (CNRS UMR 3523) |
| **Dr. Nicolas WOLFF** | Directeur de recherches, Institut Pasteur (CNRS UMR 3571) |

**Autres membres du jury :**

| | |
|---|---|
| **Dr. Mariel DONZEAU** | MCF - MCUPH, Biotechnologie et Signalisation Cellulaire (UMR 7242) |

**Invité :**

| | |
|---|---|
| **Dr. Matteo NEGRONI** | Directeur de recherches, Institut de Biologie Moléculaire et Cellulaire (CNRS UPR 9002) |

# Acknowledgments

I wish to express my gratitude to the members of the jury who accepted to evaluate my thesis: Marielle Donzeau, Fabrice Argou, Matteo Negroni and Nicolas Wolff.

This project would never be as good as I think it is without the precious help of our collaborators, who provided us with great professionality and seriousness, without whom this project would not be as good as I hope it is. I wish to thank Alastair McEwen, Pierre Poussin-Courmontagne, Professor Valdimir Torbeev, Professor Alain Chariot, Juri Rappsilber and Norman Davey (and of course all their amazing collaborators).

Of course, my gratitude goes to the Fondation pour la Recherce Médicale which granted the first three years of my PhD and to the Ligue Contre le Cancer that financed the six last additional months.

Personally, I wish to thank all my short and long-term colleagues, in particular Changqing who more than anyone else helped me to go through the many difficulties encountered in the laboratory and Hanna, a precious friend and confidant. Huge thanks go to Georges who helped me to write this thesis, giving me precious advices. I thank Murielle, Diane and Marie-Laurie for their kindness and for their support.

My sincere gratitude goes to Katia, who initially gave me the chance to come here and start this incredible journey and who checked every single step I took, leading me to the very end of the PhD.

Thanks to mom, dad and all my family. They have been always supporting me. I hope they will also forgive my absence for such a long time.

I've never been particularly sentimental so I will just make a short list of those friends who have always been present with affection, sharing good and bad times (also because they will never read these lines): Arianna, Rassa, L, Teo, Stu, Erika.

Last, I want to thanks to myself. Not so much time ago, I was a simple waiter working in a humble bar lost in the Lombard Lowlands, while now, after almost 6 years of hard work and sacrifices, I am about to finish a PhD.

This time, I think I did quite a good job.

***Per aspera ad astra***

# List of abbreviations

**A**

ARD = Ankyrin repeat domain

AT = Atherosclerosis

ATP = Adenosine triphosphate

β-TRCP = β-transducin repeat containing subunit

**B**

BAFF = B cell activating factor

BCR = B cell receptor

BDNF = Brain development neurotrofic factor

BEVS = Baculovirus expression system vector

Bis-Tris = Bis-Tris(hydroxymethyl)aminomethane

**C**

CAF = Cancer associated fibroblast

CAT = Catalase

CBM = CARD11-BCL10-MALT1

CBP = CREB binding protein

CC = Coiled-coil domain

CDD = Common docking domain

CDK = Cyclin-dependent kinase

CNS = Central nervous system

Co-IP = Co-immunoprecipitation

COX = Cyclooxygenase

CREB = cAMP response element-binding protein

CSR = Class switch recombination

**D**

DAMP = Damage associated molecular pattern

DMEM = Dulbecco's modified essential medium

DMSO = Dimethyl sulfoxide

DR = Death receptor

DTT = 1,4-dithiotheritol

**E**

EAE = Experimental autoimmune encephalomyelitis

EDA = Ectodysplasin A

EDA-ID = Anhidrotic ectodermal dysplasia with immunodeficiency

EDTA = Ethylenediamenetetraacetic acid calcium disodium salt

**F**

FADD = Fas associated death domain

FoB = Follicular B cell

**G**

GC = Germinal center

GFP = Green fluorescent protein

GPCA = *Gaussia princeps* complementation assay

**H**

HDAC = Histone deacetylase

HGF = Hepatocyte growth factor

**I**

IBD = Inflammatory bowel disease

IGF = Insulin-like growth factor

IκB = Inhibitor of κB

IKK = Inhibitor of kB kinase

IL = Interleukin

ILR = Interleukin receptor

IP = Incontinentia pigmenti

IPTG = Isopropyl b-D-1-thiogalactopyranoside

IRAK = IL-1 receptor associated kinase

ITC = Isothermal titration calorimetry

**K**

KBD = Kinase binding domain

KD = Kinase domain

**L**

LB = Luria broth

LBP = Lipopolysaccharide binding protein

LDL = low-density lipoprotein

LPS = Lipopolysaccharide

LTD = Long term depression

LTP = Long term potentiation

LZ = Leucine zipper

**M**

MAPK = Mitogen activated protein kinase

MBP = Maltose binding protein

MHC = Major histocompatibility complex

MOI = Multiplicity of infection

MS = Multiple sclerosis

MZB = Marginal zone B cell

**N**

NBD = NEMO binding domain

NEMO = NF-κB essential modulator

NF-κB = Nuclear factor κB

NGF = Nerve growth factor

NIK = NF-kB inducing kinase

NLR = NOD-like receptor / Normalized luminescence ration

NLS = Nuclear localization sequence

NOA = NEMO-optineurin-ABIN domain

**O**

OD600 = Optical density at 600 nm

**P**

PAMP = Pathogen associated molecular pattern

PBS = Phosphate buffer saline

PFD = Pore forming domain

PIC = Protease inhibitors cocktail

PKA = Protein kinase A

PKB = Protein kinase B

PPI = Protein-protein interaction

PRR = Pathogen recognition receptor

PSD = Post synaptic density

PTM = Post translational modification

**R**

RA = Rheumatoid arthritis

RANK = Receptor activator of NF-kB

RD = Regulatory domain

RHD = Rel Homology domain

RIP1 = Receptor interacting protein 1

ROS = Reactive oxigen species

RSS = Recombination signal sequence

**S**

SCF-bTrCP = Skp1-Culin-Roc1/Rbx1/Hrt-1-F-box

SCID = Severe combined immunodeficiency

SDD = Scaqffold/dimerization domain

SH = Src homology domain

SLiM = Short linear motif

SOC = Super optimal broth with catabolite repression

SOD = Superoxide dismutase

SRR = Signal responsive region

**T**

TAB = TAK binding protein

TAD = Transactivation domain

TAK = TGFb-activated kinase

TBD = TANK binding domain

TBK = TANK binding protein kinase

TBP = TATA binding protein

TCR = T cell receptor

TD = Thymus dependent

TEV = Tobacco etch virus protease

TF = Transcription factor

TGF = Transforming growth factor

TI = Thymus independent

TLR =Toll-like receptor

TNF = Tumor necrosis factor

TNFR = Tumor necrosis factor receptor

TRADD = TNF-receptor associated death domain

TRAF = TNF-receptor-associated factor

**U**

UBAN = Ubiquitin-binding in ABIN and NEMO domain

UBD = Ubiquitin binding domain

ULD = Ubiquitin-like domain

**X, Z**

XIAP = X-linked inhibitor of apoptosis protein

ZNF = Zinc finger domain

# Index

**Abstract** **4**

**Résumé de thèse en Français** **6**

**Part 1 – Introduction** **21**

**Chapter 1 - Overview on the structural features and mechanisms of protein kinases** **22**
1.1 Kinase families according to the human kinome classification 22
1.2 Structure of the kinase domain and activation mechanisms 24
1.2.1 The N- and C-lobes. 25
1.2.2 Transition between inactive and active states of the kinase domain 27
1.3 Kinase-substrate interaction 27
1.3.1 Short linear motifs within the catalytic region determine a first selectivity criterium 28
1.3.2 Docking interactions enhance substrate affinity 29
1.3.3 Docking interaction mediated by ligand binding domains 32
1.4 Regulation of kinase activity 34

**Chapter 2 - The IKK protein family in the NF-κB signaling pathway and beyond** **35**
2.1 Overview on the molecular aspects of the NF-κB signaling 35
2.1.1 NF-κB protein family 35
2.1.2 IκB protein family 36
2.1.3 Canonical, alternative and atypical NF-κB pathways 37
2.1.4 Other functions of IKK1 and IKK2 in NF-κB signaling 40
2.2 IKK functions beyond NF-κB signaling 40
2.3 Structure of IKK1, IKK2 and NEMO proteins 41
2.3.1 The catalytic IKK1 and IKK2 subunits 41
2.3.2 Structural and functional properties of NEMO 43
2.4 IKK-related kinase TBK1 and IKKE 45
2.4.1 IKK-related kinases TBK1 and IKKE in the NF-κB and INF pathway 45
2.4.2 IKK-related kinases structural features 46

**Chapter 3 - Physiopathology of the NF-κB signaling pathway** **47**
3.1 NF-κB in inflammation and immune response 47
3.1.1 Functions of NF-κB in innate immune cells 47
3.1.2 Functions of NF-κB in B cells 49
3.1.3 Functions of NF-κB in T cells 50
3.2 NF-κB and apoptosis 50
3.2.1 Anti-apoptotic role of NF-κB in the innate immune response 50
3.2.2 Anti-apoptotic roles of NF-κB in the adaptive immune system 51
3.2.3 Crosstalk between NF-κB and tumor suppressors 42
3.3 Pathologies associated with deregulation of NF-κB signaling 52
3.3.1 Self-immune and inflammatory diseases 52
3.3.2 SCID 54
3.3.3 Incontinentia pigmenti 54
3.3.4 NF-κB-mediated viral pathogenicity 54
3.3.5 NF-κB and cancer 57
3.3.6 NF-κB in neurological diseases 59
3.4 Inhibitors of IKK 60
3.4.1 The NF-κB as a promising pharmacological target 60
3.4.2 IKK inhibitors clinical trials 62

**Part 2 - Methods** **64**

**Chapter 4 - Materials and methods** **65**
4.1 DNA cloning 65
4.1.1 Classical restriction cloning 65
4.1.2 Gateway cloning 66
4.1.3 Site-directed Mutagenesis 68
4.2 Protein expression 68
4.2.1 Protein expression in *E. coli* 68
4.2.2 Protein expression in insect cells 70
4.3 Protein purification 71
4.4 Protein interaction methods 73
4.5 Kinase activity assays 75
4.6 Crystallization 76

4.7 Structure determination of IKK2 (1-669) EE bound to the IκBα docking peptide by X-ray crystallography 77

**Part 3 - Results and discussion** **79**

**Objectives** **80**

**Chapter 5 - Identification and characterization of the IKK1 and IKK2 interactions with a novel docking motif** **81**

5.1 Results 82

5.1.1 Identification of a docking motif within the canonical IκBα substrate 82

5.1.2 Activity and functional analyses on the YDDΦXΦ docking interaction 84

5.1.3 Structural studies on the IKK2/IκBα peptide complex 86

5.1.4 Identification of other substrates interacting with IKK1 or IKK2 via the YDDΦXΦ motif 88

5.2 Discussion 90

**Chapter 6 - Characterization of the interaction between NEMO and IκBα** **96**

6.1 Results 96

6.1.1 The NEMO binding region within IκBα involves the YDDΦXΦ motif 96

6.1.2 Characterization of the IκBα binding region within NEMO 98

6.2 Discussion 98

**Conclusions and perspectives** **101**

**Chapter 7 - Side project** **103**

**Bibliography** **105**

**Annex - Manuscript** **126**

## Abstract (English)

The inhibitor of κB kinase (IKK) is the master regulator of NF-κB signalling and plays a fundamental role in the activation and regulation of inflammatory processes and innate immune response. In the absence of stimuli, NF-κB transcription factors are retained in the cytoplasm, by interaction with inhibitor of κB (IκB) proteins or in the form of precursor proteins. Upon receptor stimulation, IKK phosphorylates IκB proteins, inducing their poly-ubiquitination and degradation (canonical pathway), or the p100 precursor, promoting its maturation into the p52 NF-κB subunit (alternative pathway). In both cases, NF-κB signalling results in the translocation of NF-κB transcription factors into the nucleus where they regulate gene expression.

IKK is an enzymatic complex with variable subunit composition. The IKK species regulating NF-κB signaling are the canonical IKK complex, consisting of two homologous catalytic subunits (IKK1 and IKK2) and the regulatory scaffold NEMO, and the IKK1 homodimer. Although structural information is available for the IKK subunits, detailed knowledge of the mechanisms of IKK-substrate interactions is lacking.

In this work, we have identified a novel YDDΦXΦ docking motif that is present in the disordered C-terminus of IκBα. This motif is necessary and sufficient for the IκBα interactions with both the IKK1 and IKK2 catalytic subunits. A direct correlation between the affinity of the YDDΦXΦ docking interaction and IKK2 kinase activity is observed. We also provide a 3D structural model of this interaction based on x-ray diffraction and cross-linking mass spectrometry data. In this model, the SDD domains of the IKK2 homodimer form a pocket for the YDDΦXΦ motif. The YDDΦXΦ docking motif is also present in other IKK substrates, such as IκBβ, p100 and IRF7. Moreover, our preliminary analyses indicate that this motif is also recognized by the NEMO regulatory subunit. Finally, synthetic peptides containing the YDDΦXΦ motif are able to inhibit substrate phosphorylation by IKK2, opening the way to future development of a new class of selective IKK inhibitors.

## Résumé (Français)

L'inhibiteur de la kinase κB (IKK) est le régulateur principal de la signalisation du NF-κB et joue un rôle fondamental dans l'activation et la régulation des processus inflammatoires et de la réponse immunitaire innée. En l'absence de stimuli, les facteurs de transcription NF-κB sont retenus dans le cytoplasme par interaction avec les protéines inhibitrices de κB (IκB) ou sous forme de protéines précurseurs. Lors de la stimulation des récepteurs, IKK phosphoryle les protéines IκB, induisant leur poly-ubiquitination et leur dégradation (voie canonique), ou le précurseur p100, favorisant sa maturation en sous-unité p52 du NF-κB (voie alternative). Dans les deux cas, la signalisation NF-κB entraîne la translocation des facteurs de transcription NF-κB dans le noyau où ils régulent l'expression des gènes.

IKK est un complexe enzymatique dont la composition des sous-unités est variable. Les espèces d'IKK régulant la signalisation du NF-κB sont le complexe IKK canonique, composé de deux sous-unités catalytiques homologues (IKK1 et IKK2) et l'échafaudage régulateur NEMO, et l'homodimère IKK1. Bien que des informations sur la structure soient disponibles pour les sous-unités IKK, une connaissance détaillée des mécanismes d'interactions IKK-substrat fait défaut.

Dans ce travail, nous avons identifié un nouveau motif d'amarrage, ou docking, qui est basé sur le consensus YDDΦXΦ et qui est présent dans l'extrémité C-terminale désordonnée de IκBα. Ce motif est nécessaire et suffisant pour les interactions de IκBα avec les sous-unités catalytiques IKK1 et IKK2. Une corrélation directe entre l'affinité de l'interaction de docking et l'activité kinase d' IKK2 est observée. Nous fournissons également un modèle structural 3D de cette interaction basé sur la diffraction aux rayons X et des données de spectrométrie de masse avec agents de pontage chimique. Dans ce modèle, les domaines SDD de l'homodimère IKK2 forment une poche pour le motif YDDΦXΦ. Le motif d'amarrage YDDΦXΦ est également présent dans d'autres substrats d'IKK, tels que IκBβ, p100 et IRF7. De plus, nos analyses préliminaires indiquent que ce motif est également reconnu par la sous-unité régulatrice NEMO. Enfin, les peptides synthétiques contenant le motif YDDΦXΦ sont capables d'inhiber la phosphorylation du substrat par IKK2, ouvrant la voie au développement futur d'une nouvelle classe d'inhibiteurs sélectifs d'IKK.

# Résumé de la thèse

## PARTIE I - Introduction

**Chapitre 1 : Aperçu des caractéristiques structurelles et des mécanismes des protéines kinases**

Dans ce chapitre, je passe en revue les principales caractéristiques structurelles des protéines kinases et leurs mécanismes d'action, en mettant l'accent sur le domaine kinase.

Les protéines kinases sont des enzymes qui appartiennent au groupe des phosphotransférases, catalysant le transfert d'un groupe phosphate d'une molécule donneuse (donatrice) de haute énergie (généralement l'ATP ou le GTP) à une protéine substrat au niveau d'un résidu spécifique. Dans le cas des sérine/thréonine kinases, ce résidu est soit (indifféremment) un acide aminé de type sérine ou thréonine, alors que dans le cas des tyrosine kinases, il s'agit d'une tyrosine. En outre, un petit groupe de protéines kinases, appelées kinases à double spécificité, peuvent phosphoryler les résidus tyrosine ou sérine/thréonine [1]. Les résidus phospho-accepteurs sont généralement entourés d'une séquence consensus typique. Dans le génome humain, environ 500 gènes codant pour des kinases et des pseudo-kinases sont identifiés, ce qui constitue environ 2 % de tous les gènes.

La classification des kinases en différentes familles, proposée à l'origine par Hanks et Hunter [2], se fonde principalement sur des analyses de séquence du domaine kinase, sur la structure des domaines extérieurs au domaine kinases et sur des fonctions biologiques apparentées. Cette classification décrit neuf groupes (protéine kinase dépendante de l'AMPc/protéine kinase G/protéine kinase C famille étendue ou ACG kinases, Calcium/calmodulin regulated kinases ou MKs, CDK, MAP kinase, glycogène synthase kinase et CDK-like ou CMGCs, protéine tyrosine kinases ou PTKs homologues de STE11 et STE20 ou STEs, caséine kinases ou CK1s et tyrosine kinase-like ou TKLs et récepteur guanylate cyclase kinases ou RGCs), 90 familles et 145 sous-familles [3].

Bien que les kinases présentent une grande hétérogénéité dans leurs contributions fonctionnelles, l'architecture du domaine kinase et les changements structurels requis lors de la transition de l'état inactif à l'état actif sont hautement conservés[4]. Le domaine kinase est organisé en deux lobes globulaires distincts, conventionnellement appelés lobe N et lobe C. Ces lobes sont assemblés par une série de séries d'étapes qui permettent d'assurer la continuité de l'activité de la kinase. Ces lobes sont assemblés par une série de α-hélices et de feuillets-β, profondément interconnectés d'un point de vue structurel et fonctionnel. Les lobes abritent des éléments pivots pour l'activation et le fonctionnement de la kinase, tels que les boucles d'activation et catalytiques et le site de liaison à l'ATP. Ces boucles contiennent des motifs linéaires hautement conservés qui participent au

mécanisme d'activation, à la stabilisation de la poche de liaison à l'ATP et à l'architecture globale du domain kinase.

Les kinases sont généralement maintenues inactives par la boucle d'activation, qui se replie sur la boucle catalytique, empêchant toute interaction avec les substrats. Le déplacement de la boucle d'activation est obtenu par cis- ou trans-phosphorylation de résidus conservés au sein de la boucle d'activation. La transition d'un état basal inactif à une configuration catalytiquement compétente (active) est réalisée par l'assemblage de deux motifs spatiaux : l'épine régulatrice (R-spine) et l'épine catalytique (C-spine). À de rares exceptions près, ces épines hydrophobes sont hautement conservées et omniprésentes dans toutes les kinases actives, alors qu'on ne les trouve jamais dans les kinases inactives.

Les kinases sont très sélectives en ce qui concerne les substrats. Même dans le cas de kinases étroitement apparentées, ces enzymes sont capables de reconnaître et de phosphoryler des substrats très différents. En général, la sélectivité peut être considérée comme une série de niveaux successifs et progressivement rigoureux qui partent de la discrimination primaire du type de résidu phospho-accepteur et s'étendent pour inclure les interactions de docking (amarrage) en dehors de la région catalytique, les mécanismes de régulation allostérique et la physiologie cellulaire, telle que l'augmentation de la concentration du substrat ou la sous-localisation cellulaire.

La sélectivité d'une kinase est déterminée par les interactions protéine-protéine médiées par le domaine kinase ou par des domaines supplémentaires de liaison au ligand (tels que les domaines SH2, SH3, PTB) présents dans la protéine kinase et impliquant de larges surfaces d'interaction. Souvent, ces interactions de docking sont médiées par des motifs linéaires courts (SLiM) au sein des protéines substrats, qui sont des segments courts conservés, de trois à 11 résidus d'acides aminés. Les SLiM ont évolué pour servir de médiateurs aux interactions protéine-protéine et se trouvent fréquemment dans des régions intrinsèquement désordonnées, où ils apparaissent comme des îlots de conservation dans des régions non conservées. Les interactions médiées par les SLiMs sont faibles, transitoires et réversibles, ce qui en fait un outil optimal pour la modulation de différents processus dynamiques. Les interactions dépendantes des SLiMs peuvent être régulées par des modifications post-traductionnelles.

Le réglage fin de l'activité de la kinase est réalisé par des mécanismes de régulation allostérique. Les interactions protéine-protéine, les modifications post-traductionnelles, la liaison avec de petites molécules ou des effecteurs et les variations des acides aminés éloignés participent à la régulation de l'activité d'un enzyme en induisant des réarrangements structurels et en altérant les propriétés dynamiques du domaine kinase. L'assemblage de ces structures fonctionnelles transitoires est une

étape critique qui amène la kinase dans un état dynamique engagé [4] [5]. Dans cette configuration, on observe un autre mécanisme subtil de régulation, caractérisé par de petits mouvements sous-structurels, des vibrations intramoléculaires qui définissent des communautés dynamiques virtuelles qui modifient les caractéristiques de l'activité kinase, sans altérer la structure du domaine kinase [6].

**Chapitre 2 : La famille des protéines IKK dans la signalisation NF-κB et au-delà**

Dans ce chapitre, je passe en revue les caractéristiques structurelles et fonctionnelles de la famille des inhibiteurs de la kinase κB (IKK), en mettant particulièrement l'accent sur leur contribution à la régulation de la signalisation NF-κB.

La famille IKK se compose de cinq protéines, IKK1 (ou IKKα), IKK2 (ou IKKβ), IKKE, TBK1, et NEMO (NF-κB essential modulator ou IKKγ). Les protéines IKK1, IKK2 et NEMO s'assemblent en différents complexes enzymatiques (appelés collectivement complexe IKK) dont la stœchiométrie et la composition des sous-unités sont variables. Le complexe IKK est le régulateur principal de la signalisation NF-κB, bien qu'il joue également un rôle clé dans d'autres voies de signalisation et de régulation transcriptionnelle importantes, alors que IKKE et TBK1 jouent un rôle fondamental dans la voie de l'interféron.

Le facteur nucléaire kappa-light-chain-enhancer des cellules B activées (NF-κB) est une famille de facteurs de transcription nucléaires inductibles, qui régulent l'expression des groupes spécifiques de gènes impliqués dans plusieurs fonctions physiologiques, notamment les réponses immunitaires et inflammatoires. La famille NF-κB comprend cinq membres, regroupés en deux sous-familles. La sous-famille Rel comprend RelA (p65), RelB et c-Rel, tandis que la sous-famille NF-κB comprend NF-κB1 (p105/p50) et NF-κB2 (p100/p52). Tous les membres partagent un domaine de liaison/dimérisation de l'ADN hautement conservé (domaine d'homologie REL, RHD) par lequel ils forment des homo- et hétéro-dimères. Les protéines de la sous-famille Rel sont des facteurs de transcription actifs, retenus dans le cytoplasme par l'interaction directe avec les protéines inhibitrices du facteur nucléaire κB (IκB). La famille IκB est un groupe de cinq protéines (IκBα, IκBβ, IκBε, IκBζ et Bcl-3), qui se lient aux dimères Rel de NF-κB, masquant leur séquence de localisation nucléaire (NLS) et provoquant ainsi leur rétention dans le cytoplasme. Toutes les protéines IκB sont caractérisées par la présence d'un domaine répété d'ankyrine (ARD) et de régions N- et C-terminales désordonnées. En revanche, les protéines NF-κB p105 et p100 sont présentes dans le cytoplasme sous forme de précurseurs inactifs.

L'activation de la signalisation NF-κB se réalise de trois manières : la voie canonique, la voie alternative et la voie atypique. Dans la voie canonique, un complexe IKK (appelé complexe IKK canonique) composé d'un noyau constitué de IKK1, IKK2 et NEMO phosphoryle les protéines IκBα

ou IκBβ. La phosphorylation à des résidus sérine spécifiques génère un motif « phosphodégron » qui induit la reconnaissance par la ligase E3 Skp1-Culin-Roc1/Rbx1/Hrt-1-F-box (SCF-bTrCP) et la dégradation ultérieure des protéines IκB *via* le protéasome. Dans la voie alternative, le complexe IKK est constitué d'un homodimère de IKK1, qui cible p100 en favorisant la formation de la forme active p52 *via* une protéolyse partielle médiée par l'ubiquitine ligase SCF-bTrCP. En revanche, la voie atypique repose sur des mécanismes indépendants d'IKK.

De plus, le complexe IKK est impliqué dans un nombre considérable de voies de manière indépendante de NF-κB. Par exemple, IKK1 participe au remodelage de la chromatine, en ciblant le co-répresseur transcriptionnel SMRT ou d'autres facteurs tels que p65 et CBP, qui régulent l'activité de HDAC3. IKK1 active aussi plusieurs facteurs, y compris IRF7, et la β-caténine. Notamment, la β-caténine est ciblée par IKK2, qui induit sa dégradation *via* la création d'un motif « phosphodégron ».

IKK1 et IKK2 sont les sous-unités catalytiques du complexe IKK. IKK1 et IKK2 ont un haut degré d'homologie, avec 51% d'identité de séquence et des structures très similaires [7, 8, 9]. Les deux protéines comprennent un domaine kinase N-terminal (KD), un domaine de type ubiquitine (ULD), un domaine charpente/dimérisation (SDD) composé de 3 hélices antiparallèles et un domaine C-terminal de liaison à NEMO (NBD). IKK1 et IKK2 existent tous deux sous forme d'homo- et d'hétérodimères. La dimérisation est médiée par les domaines SDD, *via* une large interface. Plusieurs contacts résidus -spécifiques existent entre les domaines KD, ULD et SDD qui sont essentiels à l'architecture globale de IKK1 et IKK2, qui a une forme caractéristique en ciseaux avec un certain degré de flexibilité.

NEMO est au contraire la charpente et la protéine régulatrice du complexe canonique IKK. NEMO est nécessaire pour une activité correcte de l'IKK, comme elle est nécessaire pour cibler les protéines IκB par IKK2 in vivo [10]. Dans le complexe canonique, un hétérodimère IKK1/2 se lie à deux copies de NEMO. NEMO semble être une protéine avec une conformation « molen globule » composée d'un noyau d'hélices entrecoupées de régions désordonnées. Plusieurs structures de fragments hélicoïdaux de NEMO seul [11, 12, 13]ou liée à l'ubiquitine [14, 15] ont été determineés. La région N-terminale de NEMO contient le domaine de liaison à la kinase (KBD) nécessaire à la liaison à IKK1 et IKK2, suivi d'un premier domaine à structure «coiled-coil» (CC1). La partie centrale de NEMO contient le deuxième domaine hélicoïdal (HLX2), qui comprend le domaine de liaison TANK (TBD) et qui sert vraisemblablement de médiateur à l'interaction avec d'autres partenaires de liaison et protéines régulatrices. Le TBD est suivi d'un deuxième domaine « coiled-coil » (CC2), du domaine NEMO-optineurin-ABIN (NOA) et d'un domaine « leucin-zipper ». Ces trois domaines abritent (portent, renferment, forment) le domaine de liaison à l'ubiquitine dans ABIN et NEMO (UBAN) nécessaire à l'interaction avec la poly-ubiquitine. Enfin, la région C-terminale de NEMO contient un domaine à doigt de zinc (ZF) [12].

La structure de la TANK-binding kinase 1 (TBK1) a également été décrite (résolue) [16]. TBK1 et IKKE sont deux kinases liées à l'IKK, qui exercent le rôle important de coordonner la voie du NF-κB et la voie de l'interféron (INF).

**Chapitre 3 : Physiopathologie de la voie de signalisation NF-κB**

Compte tenu du rôle central du complexe IKK dans la régulation de la signalisation NF-κB, je n'ai sélectionné dans ce chapitre que les fonctions physiologiques les plus significatives de la signalisation NF-κB. En particulier, je me concentre sur l'activation et la régulation de la réponse immunitaire et des processus inflammatoires, en décrivant les mécanismes moléculaires qui contrôlent l'activation, la maturation, la prolifération, la polarisation et la mort cellulaire. Je décris également certaines maladies marquantes liées à des déficiences du NF-κB, en donnant un aperçu des inhibiteurs pharmacologiques testés pour IKK [17].

La signalisation NF-κB contrôle l'activation, la maturation et la destinée des cellules immunitaires innées et adaptatives. Les macrophages, les cellules dendritiques et les neutrophiles détectent plusieurs éléments exogènes, appelés motifs moléculaires associés aux agents pathogènes (PAMP) et motifs moléculaires associés aux dommages (DAMP), par l'intermédiaire de leurs récepteurs primitifs de reconnaissance des motifs (PRR). Dans les macrophages, l'interaction des DAMPs et PAMPs avec les récepteurs PRRs active la voie canonique NF-κB, induisant l'expression de molécules pro-inflammatoires telles que les cytokines, les chimiokines et autres médiateurs inflammatoires [18]. De cette façon, NF-κB induit la polarisation des macrophages vers le phénotype pro-inflammatoire M1 et les recrute pour la pyroptose, un programme de mort cellulaire basé sur l'inflammasome, qui renforce la réponse inflammatoire.

Selon le type cellulaire considéré, la maturation et les conditions environnementales, la signalisation NF-κB contrecarre ou favorise l'apoptose. Par exemple, dans le système immunitaire inné, NF-κB induit l'activation et la maturation des cellules, stimulant les programmes de survie et de prolifération et, en parallèle, génère des signaux de mort cellulaire pour la pyroptose ou l'apoptose. De cette façon, NF-κB détermine la durée de vie des cellules et les fonctions des cellules immunitaires innées. En outre, NF-κB est également impliqué dans la sélection et la maturation du système immunitaire adaptatif (cellules B et T). Étant donné que NF-κB régule les processus d'inflammation ainsi que le comportement des cellules, il joue également un rôle dans le contrôle des cellules aberrantes en interagissant avec plusieurs suppresseurs de tumeurs tels que p53 et ARF.

En raison du rôle critique de NF-κB dans la régulation du système immunitaire et des processus d'inflammation, sa dérégulation est souvent observée dans différents types de cancers solides et elle est également liée à une série de maladies auto-immunes et inflammatoires (par ex. la polyarthrite

rhumatoïde, les maladies inflammatoires de l'intestin, la sclérose en plaques et l'athérosclérose). Ces pathologies sont caractérisées par une activation constante de NF-κB qui génère une inflammation chronique, altérant l'homéostasie de différents tissus et organes. Des mutations dans le gène IKBKG codant pour la protéine NEMO conduisent au syndrome multi-systémique Incontinentia pigmenti (IP) et aux formes spécifiques de dysplasie ectodermique anhydrotique avec immunodéficience (EDA-ID).

Étant donné le rôle du NF-κB dans la régulation des processus inflammatoires et du comportement cellulaire, des altérations de la signalisation sont fortement liées au cancer. Dans des conditions physiologiques, le NF-κB autorégule son activité par des rétroactions négatives induites par l'expression des protéines IκB et d'autres facteurs qui inhibent l'activation de la kinase IKK, comme l'A20 [19]. La dérégulation du NF-κB peut augmenter le risque de cancer par différents mécanismes. L'activation chronique de NF-κB augmente le potentiel pro-tumoral des cellules car elle induit la prolifération et favorise la formation de métastases en modifiant l'expression des métalloprotéases et des gènes d'adhésion cellulaire. Il stimule également l'angiogenèse et augmente la résistance des cellules à l'hypoxie [20]. Indirectement, le NF-κB induit la transformation tumorale *via* l'activation des cellules immunitaires, des lymphocytes et en induisant la polarisation des macrophages [21].

Par ailleurs, les virus ont développé des mécanismes qui ciblent la signalisation du NF-κB. En modifiant le comportement cellulaire, certains virus peuvent augmenter la résistance à la mort cellulaire immunodépendante (par exemple le VIH, le HSV et le HBV) des cellules infectées ou réduire la réponse immunitaire en induisant l'apoptose des cellules immunitaires (virus de la vaccine). D'autres virus, tels que le SARS-CoV-2, ont montré leur capacité à optimiser leur virulence en affaiblissant la réponse immunitaire et en augmentant simultanément la survie des cellules infectées en modulant la signalisation NF-κB [22].

Le NF-κB est également impliqué dans le développement, la protection et l'homéostasie du système nerveux central. Une augmentation de l'inflammation dans le cerveau causée par des altérations de la microglie, participe à la génération de syndromes neurologiques, tels que la dépression et les maladies neurodégénératives. De même, l'hyperactivation de la microglie pendant et immédiatement après un accident vasculaire cérébral ischémique peut accroître les dommages neurologiques [23].

L'inhibition sélective de la signalisation NF-κB en présence d'une inflammation chronique est considérée comme une stratégie pharmacologique prometteuse pour réduire l'hyperactivation néfaste du système immunitaire. Jusqu'à présent, plusieurs inhibiteurs ciblant directement le complexe IKK ont été découverts (par exemple, des mimétiques de l'ATP, des inhibiteurs allostériques, des composés thiol-réactifs). Malheureusement, la majorité d'entre eux ont échoué aux essais cliniques

en raison des multiples fonctions de la kinase IKK dans la cellule, ce qui entraîne de graves effets secondaires.

# PARTIE II – Méthodes

## Chapitre 4: Matériels and méthodes

Dans ce chapitre, je décris les techniques que j'ai utilisé pour mon projet de doctorat.

*Clonage d'ADN*. Deux méthodes de clonage ont été utilisées, à savoir le clonage de restriction classique et l'approche Gateway. Le clonage classique a été utilisé pour les fusions 6xHis-IκBα, 6xHis-NEMO, MBP-IκBα et MBP-peptides qui ont servie pour les expériences de pulldown et les dosages d'activité kinase. Ces constructions ont été clonées dans des vecteurs de la série pETM, à savoir pETM-10 (pour les fusions 6xHis) et pETM-40 (fusions MBP), pour l'expression des protéines dans *E. coli*. L'approche Gateway a été utilisée pour créer des constructions pour l'expression dans des cellules eucaryotes. Plus précisément, les fusions Gluc1 (G1) et Gluc2 (G2) d' IKK1 et IKK2, NEMO et IκBα, utilisées pour les expériences d'interaction protéine-protéine basées sur l'approche GPCA (voir ci-dessous), ont été créés en utilisant les vecteurs de destination pSPICA-N1 et pSPICA-N2. Les constructions 6xHis-IKK1(10-667) EE et 6xHis-IKK2(1-669) EE pour l'expression dans des cellules d'insecte ont été clonées dans le vecteur de transfert pour le système baculovirus pBAKPac. La mutagenèse dirigée de ces constructions a été réalisée par PCR à cycle continu.

*Expression des protéines.* Deux systèmes d'expression ont été utilisés, à savoir *E. coli* et des cellules d'insecte. Les constructions 6xHis-IκBα, 6xHis-NEMO, MBP-IκBα et MBP-peptides ont été produites dans les souches *E. coli BL21 (DE3)* ou *Rosetta*. Les constructions 6xHis-IKK1 (10-667) EE et 6xHis-IKK2 (1-669) EE ont été produites dans les cellules d'insecte Spodoptera frugiperda Sf9 par le système Baculovirus.

*Purification des protéines.* Toutes protéines recombinantes ont été purifiées par chromatographie d'affinité (résine d'amylose pour les constructions MBP et résine $Ni^{2+}$-NTA pour les constructions 6xHis), généralement suivie d'une chromatographie par filtration sur gel à l'aide de colonnes Superdex 75 ou Superdex 200, selon la taille des protéines. Pour les études structurales par cristallographie aux rayons X, l'étiquette de fusion a été clivée par digestion à la protéase TEV (Tobacco Etch Virus).

*Tests d'interaction protéine-protéine*. Les interactions protéine-protéine ont été étudiées *in vitro* à l'aide d'un test de type pulldown et *in cellulo* par l'approche GPCA (*Gaussia princeps* complementation assay). Les expériences de pulldown ont été effectuées afin d' analyser les interactions entre les protéines recombinantes purifiées. Dans ces expériences, la protéine appât est

fusionnée à l'étiquette MBP et immobilisée par capture d'affinité à la résine d'amylose. Les résines sont incubées avec la protéine proie purifiée. Après l'incubation, la résine est lavée puis le complexe (appât + proie) élué par addition du compétiteur maltose. L'élution des complexes de la résine est cruciale afin d'éliminer le bruit de fond dérivé de l'absorption des protéines IKK1 et IKK2 sur la résine.

L'approche GPCA est un test de complémentation protéique utilisé pour étudier les interactions protéine-protéine dans la cellule. Ici, la luciférase est divisée en deux fragments (Gluc1 ou G1 et Gluc2 ou G2), qui sont fusionnés aux deux protéines d'intérêt. Si les protéines cibles interagissent, les fragments sont mis en contact étroit et la structure native de la luciférase est reconstituée. Dans ce cas, l'ajout du substrat coelentérazine génère un signal luminescent. L'intensité de l'interaction est extrapolée par le rapport de luminescence normalisé [25]. Toutes les expériences GPCA décrites dans cette thèse ont été réalisées par co-expression de protéines de fusion G1 et G2 dans des cellules de mammifère HEK293T.

*Dosages d'activité kinase.* Les expériences d'activité kinase décrites ici utilisent les protéines 6xHis-IκBα recombinantes purifiées et soit la construction 6xHis-IKK2 EE (1-699) EE purifiée, soit le complexe IKK endogène immunoprécipité à partir de cellules HEK293T à l'aide d'un anticorps anti-NEMO. Afin de pouvoir évaluer les différences dans les niveaux de phosphorylation des protéines IκBα de type sauvage et mutants, dans ces expériences, nous avons utilisé de faibles concentrations de kinase et de protéines substrat, ainsi que des fortes concentrations de BSA afin de simuler les conditions du milieu cytoplasmique. La phosphorylation spécifique d'IκBα a été détectée à l'aide d'un anticorps anti-phospho-IκBα.

*Cristallisation et détermination de la structure.* Les expériences de cristallisation ont été réalisées en collaboraration avec la plateforme de cristallisation macromoléculaire de l'IGBMC (Strasbourg). Dans ces expériences la construction IKK2 (1-669) EE a été mélangé avec un peptide synthétique comprenant les résidus 297-317 d'IκBα. Le criblage initial des conditions de cristallisation par la méthode de diffusion de vapeur en goutte assise a été réalisé à l'aide des kits commerciaux. Ensuite, les cristaux ont été optimisés par des sous-criblages optimisés autour des conditions initialement identifiés. La collecte des données de diffraction des rayons X et la détermination de la structure ont été réalisées par Alistair McEwen (plateforme de cristallographie macromoléculaire, IGBMC). La structure du complexe peptidique IKK2/IκBα a été résolue par remplacement moléculaire en utilisant un monomère extrait de la structure du dimère asymétrique de l'IKK2 humain comme modèle de recherche [7].

## Partie III – Résultats et discussion

La kinase IKK est un régulateur clé de la signalisation NF-κB, qui joue un rôle fondamental dans l'activation et la régulation des processus inflammatoires et de la réponse immunitaire innée. Dans la voie dite 'canonique' de NF-κB, lors de la stimulation du récepteur, IKK phosphoryle les protéines IκB, ce qui génère un motif phospho-degron et entraîne la dégradation complète d'IkB, conduisant ainsi à la translocation dans le noyau des facteurs NF-κB. Dans la voie dite 'non canonique' ou 'alternative' de NF-κB, IKK phosphoryle le précurseur p100, conduisant ainsi à sa dégradation partielle et donc à sa maturation en p52.

IKK est un complexe enzymatique avec une composition variable de ses sous-unités. Le complexe d'IKK 'canonique' est constitué d'un noyau composé des deux sous-unités catalytiques IKK1 et IKK2, et de deux copies de la sous-unité régulatrice NEMO. En revanche, l'espèce d'IKK, qui régule la voie 'non canonique', est un homodimère d'IKK1.

Les structures des homodimères d'IKK1 et d'IKK2 et de différentes régions de la protéine NEMO ont déjà été déterminées. Cependant, à ce jour il n'existe aucune information sur les mécanismes moléculaires des interactions d'IKK avec ses substrats. L'objectif principal de cette thèse est donc de caractériser ces mécanismes. Dans ce but, nous avons utilisé différentes approches biochimiques, fonctionnelles et structurales.

### Chapitre 5: Identification et caractérisation des interactions d'IKK1 et d'IKK2 avec un nouveau motif de docking

Dans ce chapitre, je décris les principaux résultats de mon projet de recherche doctorale, qui sont liés à l'identification d'un nouveau motif de docking (amarrage, arrimage) et à la caractérisation structure-fonction de ses interactions avec les sous-unités catalytiques IKK1 et IKK2.

*Identification d'un motif d'amarrage (docking) dans le substrat IκBα.* Dans la voie canonique NF-κB, la sous-unité catalytique IKK2 phosphoryle les résidus Ser32 et Ser36 dans la région N-terminale désordonnée d'IκBα. Pour caractériser davantage l'interaction IKK2/IκBα, j'ai effectué des expériences de pulldown *in vitro* en utilisant la protéine IKK2 recombinante et purifiée, et un ensemble de constructions de délétion d'IκBα. La construction d'IKK2 utilisée, c'est-à-dire His-IKKβ (1-669) EE, correspond à la construction précédemment cristallisée [7] et est produite dans des cellules d'insectes. En revanche, les constructions d'IκBα sont produites dans *E. coli* sous forme de fusions N-terminales à l'étiquette MBP. Les résultats de ces expériences montrent qu' IKK2 interagit avec une séquence de vingt acides aminés située entre les résidus 301 et 317 dans l'extrémité C-

terminale désordonnée d'IκBα. Ensuite, afin d'identifier les résidus nécessaires à l'interaction, j'ai procédé au remplacement de toutes les positions du peptide IκBα (301-317) en sérines. Les résultats montrent que les résidus Y305, D306, D307, C308 et F310 sont essentiels pour l'interaction avec IKK2. De manière cohérente, ces mêmes résidus sont conservés dans les protéines IκBα de différentes espèces et définissent la séquence consensus YDDΦXΦ (où Y est une tyrosine, D est un acide aspartique, Φ est un acide aminé hydrophobe et X est n'importe quel acide aminé). Fait remarquable, l'introduction de la double mutation Y305S/D307S dans le contexte d'IκBα pleine longueur supprime l'interaction non seulement avec IKK2, mais aussi avec la sous-unité IKK1. Cette dernière expérience a été réalisée en utilisant la méthode GPCA (*Gaussia princeps* protein complémentation assay) qui permet la détection des interactions protéine-protéine *in cellulo* [25].

Les interactions IKK1 et IKK2 avec un peptide synthétique comprenant les résidus 297-317 d' IκBα ont été analysées par calorimétrie de titration isotherme (ITC). Ces expériences ont été réalisées par ma collègue Changqing Li (chercheur post-doctoral) en collaboration avec le Prof. Vladimir Torbeev (BSC-UMR7242, Strasbourg) et montrent que IKK1 se lie au motif docking IκBα avec une affinité légèrement supérieure (avec une constante d'affinité à l'équilibre, $K_D$, de 9 μM) par rapport à IKK2 ($K_D$=40 μM).

Afin de caractériser davantage le consensus YDDΦXΦ du motif, j'ai analysé quinze nouveaux mutants du peptide IκBα (trois mutations par position conservée du consensus) et évalué par pulldown leur impact sur les interactions avec les protéines IKK1 et IKK2 purifiées. La plupart des mutations ont un impact négatif sur les interactions avec IKK1 et IKK2. Conformément aux données d'affinité de liaison, alors qu' IKK2 affiche une sensibilité plus élevée aux mutations, IKK1 semble tolérer un plus grand nombre de substitutions au sein du motif. Surtout, j'ai pu identifier deux mutations à la première position hydrophobe (Φ) du motif, à savoir C308L et C308F, qui affichent une activité de liaison accrue envers IKK1 et IKK2. Les expériences d' ITC (réalisées par Changqing Li. Déjà noté plus haut) montrent que cela correspond à une augmentation de 30 fois de l'affinité pour IKK1 (KD=350 nM) et une augmentation de 10 fois de l'affinité pour IKK2 (KD=2.6 μM).

Dans l'ensemble, ces résultats montrent que :(i) les sous-unités catalytiques IKK1 et IKK2 contribuent aux interactions avec le substrat et (ii) ces interactions sont médiées par la reconnaissance du nouveau motif de docking YDDΦXΦ.

De plus, nous avons pu identifier des mutations au sein de ce motif qui modulent les affinités des interactions avec IKK1 et IKK2 à la fois de manière positive et négative.

*Analyses d'activité et fonctionnelles sur les interactions avec le motif YDDΦXΦ.* Pour déterminer l'impact des interactions de docking sur l'activité catalytique d'IKK2, j'ai exprimé et

purifié des protéines IκBα pleine longueur, soit de type sauvage, soit contenant la double mutation suppressive Y305S/Y307S ou la mutation C308L qui augmente l'affinité de liaison. Ces trois protéines IκBα ont été évaluées pour la phosphorylation au niveau des résidus Ser32 et Ser36 par un test kinase qui utilise la construction IKK2 (1-669) EE recombinante et purifiée. Les résultats montrent que IκBα C308L présente les niveaux de phosphorylation les plus élevés. En revanche, le niveau de phosphorylation du mutant Y305AS/Y307S est inférieur à celui d'IκBα de type sauvage. Les expériences d'activité kinase ont été répétées en remplaçant l'IKK2 recombinante par le complexe IKK endogène immunoprécipité à partir de cellules HEK293T en utilisant un anticorps anti-NEMO. Ces expériences (réalisées par Changqing Li) confirment une activité de phosphorylation plus élevée pour le mutant IκBα C308L par rapport à IκBα de type sauvage et IκBα Y305S/Y307S .

Fait remarquable, nous avons également observé que l'addition d'un excès de peptide synthétique contenant le motif YDDΦXΦ d'IκBα inhibe la phosphorylation d'IκBα par l'IKK2 recombinante ou le complexe IKK endogène.

Puisque la phosphorylation d'IκBα conduit à sa dégradation, nous avons collaboré avec le Pr Alain Chariot (Université de Liège, Belgique) afin d'évaluer l'effet des mutations dans le motif de docking sur la dégradation d'IκBα. Dans ces expériences, IκBα de type sauvage et les mutants Y305S/Y307S et C308F ont été transduits dans des fibroblastes embryonnaires de souris (MEF) dépourvus des gènes de la famille IκB (IκBα, IκBβ, IκBe). Lors de l'activation de la voie canonique NF-κB par le TNFα , nous pouvons observer que la dégradation du mutant Y305S/Y307S est plus lente par rapport à celle de IκBα de type sauvage et de C308F IκBα.

Dans l'ensemble, ces résultats montrent que l'affinité de l'interaction de docking est corrélée à la fois à la phosphorylation et à la dégradation du substrat IκBα *in vitro* et *in cellulo*.

*Etudes structurales du complexe IKK2/IκBα peptide.* Afin d'obtenir de l'information structurale sur les mécanismes d'interaction du motif YDDΦXΦ, nous avons effectué plusieurs criblages de cristallisation des complexes IKK1 et IKK2 liés au peptide IκBα(297-317), soit de type sauvage ou contenant la mutation C308L. Ces expériences ont été réalisées par Chanqing Li et moi-même avec l'aide de Pierre Poussin Courmontagne de la plateforme de cristallisation de l'IGBMC (Strasbourg).

Les complexes d'IKK1 et d'IKK2 ont tous deux donné des cristaux, dont la plupart présentaient une mauvaise diffraction en utilisant le rayonnement synchrotron. J'ai contribué à la cristallisation d'IKK2 (1-669) EE lié au peptide IκBα (301-317) de type sauvage, pour lequel on a pu identifier un cristal diffractant à une résolution suffisante pour la détermination de la structure par remplacement moléculaire (diffraction anisotrope à 4.15-6,8 Å). La carte de densité électronique ainsi dérivée montre que le peptide interagit à l'interface de l'homodimère d'IKK2. La collecte des données et la

résolution de la structure ont été réalisées par Alistair McEwen de la plate-forme de cristallisation de l'IGBMC.

En raison de la faible résolution des données de diffraction, nous n'avions pas d'information sur les chaînes latérales du peptide IκBα. Par conséquent, nous avons opté pour l'utilisation d'agents de réticulation en conjonction avec la spectrométrie de masse (cross-linking mass spectrometry) afin de cartographier de manière indépendante le site de liaison du peptide sur IKK2. Ces expériences, qui ont été réalisées par Changqing Li en collaboration avec le groupe de Juri Rappsilber (Technisches Universität Berlin, Allemagne), ont utilisé l'agent de réticulation NHS/Sulfo-EDC et nous ont permis d'identifier plusieurs interactions entre les résidus acides du peptide et des résidus lysine conservées à l'interface du dimère d'IKK2. Sur la base des données cristallographiques et de spectrométrie de masse, nous avons pu calculer une structure 3D pour le complexe IKK2/IκBα à l'aide du logiciel Haddock [25]. En outre, cette structure a été corroborée par mutagenèse de la poche de liaison IKK2 (expériences réalisées par Changqing Li). Cette poche, qui est constituée des domaines SDD des deux sous-unités IKK2, est dominée par des résidus lysine conservées interagissant avec les résidus chargés négativement du motif YDDΦXΦ.

*Identification d'autres substrats interagissant avec IKK1 ou IKK2 via le motif YDDΦXΦ.* Des analyses bioinformatiques ont permis la détection du motif YDDΦXΦ dans d'autres substrats d'IKK2 (i.e. le paralogue IκBβ) et d'IKK1 (i.e. le substrat non canonique p100 et IRF7). J'ai donc conçu des peptides contenant le motif à partir de ces substrats et testé leurs interactions avec les protéines recombinantes IKK1 et IKK2 par des analyses pulldown. Les résultats montrent qu'IKK2 ne reconnaît que les motifs d'IκBα et d'IκBβ, tandis qu'IKK1 reconnaît les motifs de tous les substrats analysés (IκBα, IκBβ, p100 et IRF7). Afin d'approfondir l'interaction d'IKK1 avec p100 et IRF7, nous avons supprimé le motif YDDΦXΦ par mutagenèse dirigée dans le contexte des protéines de pleine longueur. Les résultats des analyses par GPCA montrent que cette suppression réduit fortement l'interaction entre IKK1 et p100. En revanche, la suppression du motif YDDΦXΦ d'IRF7 n'as pas d'impact significatif sur la liaison à IKK1, suggérant ainsi qu'il existe des mécanismes supplémentaires pour cette interaction.

Pour résumer, dans ce travail nous avons identifié le nouveau motif de docking YDDΦXΦ qui interagit à la fois avec IKK1 et IKK2. Ce motif est présent dans différents substrats de phosphorylation des deux sous-unités catalytiques. Nous avons évalué les contributions du motif à la phosphorylation par IKK2 *in vitro* et *in cellulo* et avons démontré une relation directe entre l'affinité de l'interaction de docking et phosphorylation du substrat IκBα. Nous avons également pu déterminer la structure 3D de l'homodimère IKK2 lié à un peptide contenant le motif YDDΦXΦ d'IκBα, ce qui

a montré que le motif interagit avec une poche chargée positivement contribuée par les domaines SDD des deux sous-unités d'IKK2. Ces résultats sont décrits dans le suivant manuscrit actuellement en préparation :

*Changqing Li[#], Stefano Moro[#], Kataryna Shostak, Francis O'Reilly, Andrea Graziadei, Alistair Mc-Ewen, Pierre Poussin-Courmontagne, Thomas Bachelart, Denise Martinez-Zapien, Mert Fiskin, Marie Laure Straub, Georges Orfanoudakis, Massimo Tommasino, Arnaud Poterszman, Vladimir Torbeev, Juri Rappsilber, Norman Davy, Alain Chariot and Katia Zanier.* A novel YDDΦXΦ linear motif mediates docking of the IKK kinase to its substrates: structural and functional insights.

[#] *co-first authors*

**Chapitre 6: Caractérisation de l'interaction entre NEMO et IκBα**

NEMO est la sous-unité régulatrice du complexe canonique d' IKK. Il permet l'interaction d' IKK avec le complexe de signalisation associé au récepteur *via* ses interactions avec les chaînes de poly-ubiquitine. Les travaux de Hoffman et collaborateurs montrent que NEMO interagit avec le substrat IκBα [28]. Dans ce chapitre, je présente les résultats préliminaires d'un ensemble d'analyses *in cellulo* et *in vitro* portant sur la cartographie des régions minimales d'IκBα interagissant avec NEMO et *vice versa*.

*L'interaction NEMO-IκBα implique le motif YDDΦXΦ d'IκBα.* Nous avons pu reproduire l'interaction NEMO-IκBα dans notre laboratoire en utilisant l'approche GPCA *in cellulo*. Par cette approche nous avons également pu observer que la mutation Y305S/D307 dans le motif YDDΦXΦ d'IκBα réduit fortement l'interaction avec NEMO. Afin de caractériser d'avantage cette interaction, j'ai exprimé et purifié la construction 6xHis-NEMO pleine longueur, et j'ai effectué des analyses de MBP-pulldown afin d'évaluer ses interactions avec IκBα pleine longueur et les peptides YDDΦXΦ dérivés d'IκBα et d'IκBβ. De manière surprenante, nous avons constaté que la protéine NEMO recombinante n'interagit ni avec la protéine IκBα entière ni avec les peptides YDDΦXΦ. Pour exclure un effet inhibiteur potentiel de l'étiquette MBP, nous avons analysé l'interaction entre les protéines recombinantes purifiées 6xHis-NEMO et 6xHis-IκBα par chromatographie d'exclusion stérique (SEC). Encore une fois, aucune interaction entre NEMO et IκBα n'a été observée, suggérant ainsi que des facteurs supplémentaires tels que le repliement de NEMO, des modifications post-traductionnelles ou d'autres protéines de pontage peuvent contribuer à l'interaction NEMO-IκBα *in cellulo*.

Ces résultats suggèrent que le motif YDDΦXΦ permet non seulement l'interaction avec les sous-unités IKK1 et IKK2 mais également avec NEMO. Cependant, il reste à établir si le motif est suffisant pour l'interaction ou si d'autres facteurs sont requis pour l'interaction.

*Chartographie de la région de NEMO impliquée dans l'interaction avec IκBα.* J'ai également contribué à cartographier la région de NEMO interagissant avec IκBα, en complétant une analyse par GPCA initiée précédemment, basée sur l'utilisation des mutants de délétion de NEMO. Les résultats suggèrent que la région d'interaction réside dans la partie centrale de NEMO comprenant les résidus 192 et 288. Conformément à cette observation, cette région est connue pour interagir avec d'autres protéines telles que la protéine virale vFLIP [28].

Enfin, nous avons commencé à explorer l'interaction entre la sous-unité NEMO et IκBα. Les résultats préliminaires suggèrent que l'interaction est médiée, d'un côté, par le motif YDDΦXΦ d'IκBα et, de l'autre côté, par la région centrale de NEMO. Cependant, nous n'avons pas pu reproduire cette interaction *in vitro* en utilisant des protéines recombinantes purifiées. Cela suggère que des facteurs supplémentaires contribuent à cette interaction dans les cellules. En accord avec cette observation, des travaux récents ont montré que la liaison aux chaînes de poly-ubiquitine est nécessaire au bon repliement de NEMO et à ses interactions avec les protéines partenaires [13 11]. Par conséquent, d'autres expériences seront nécessaires pour caractériser pleinement les mécanismes de l'interaction NEMO-IκBα.

**Chapitre 7 : Caractérisation biochimique de l'interaction de l'isoforme oncogènique de p73 ΔNp73α avec le complexe E2F4/p130 (projet secondaire)**

ΔNp73α est une isoforme oncogène du facteur de transcription p73, un homologue fonctionnel du suppresseur de tumeur p53. Cette isoforme est dépourvue du domaine de transactivation N-terminal (TAD) présent dans les autres protéines de la famille p53 et, par conséquent, fonctionne comme un antagoniste de la signalisation de p53 et p73. ΔNp73α est régulé positivement par des virus oncogènes et dans plusieurs tumeurs malignes humaines non virales résistantes à la chimiothérapie. Notre équipe a collaboré avec Massimo Tommasino (Centre International de Recherche sur le Cancer, Lyon) pour identifier des partenaires de liaison de ΔNp73α par des analyses protéomiques. Nous avons pu ainsi montrer que ΔNp73α interagit avec le complexe répressif transcriptionnel E2F4/p130 et que cela conduit à une coopération dans l'inhibition de gènes spécifiques parmi lesquels il y a des régulateurs négatifs de la prolifération cellulaire. Dans ce projet, j'ai contribué à la caractérisation biochimique de l'interaction entre ΔNp73α et la protéine E2F4. En particulier, nous avons pu montrer que la région désordonnée N-terminale de ΔNp73α interagit avec la moitié C-terminale de E2F4. De plus, j'ai pu

reproduire cette interaction par pulldown en utilisant une construction protéique minimale purifiée recombinante de ΔNp73α et de l'hétérodimère E2F4/DP1. Les résultats de ces expériences sont décrits dans les manuscrit suivant qui est actuellement en cours de révision :

*Valerio Taverniti [#], Hanna Krynska [#], Assunta Venuti, Marie-Laure Straub, Cécilia Sirand, Eugenie Lohmann, Maria Carmen Romero-Medina, Stefano Moro, Alexis Robitaille, Luc Negroni, Denise Martinez-Zapien, Murielle Masson, Massimo Tommasino and Katia Zanier.* The E2F4/p130 repressor complex cooperates with oncogenic ΔNp73α to promote cell survival in human papillomavirus 38-transformed keratinocytes and in cancer cells.

[#] *co-first authors*

# Part 1

# INTRODUCTION

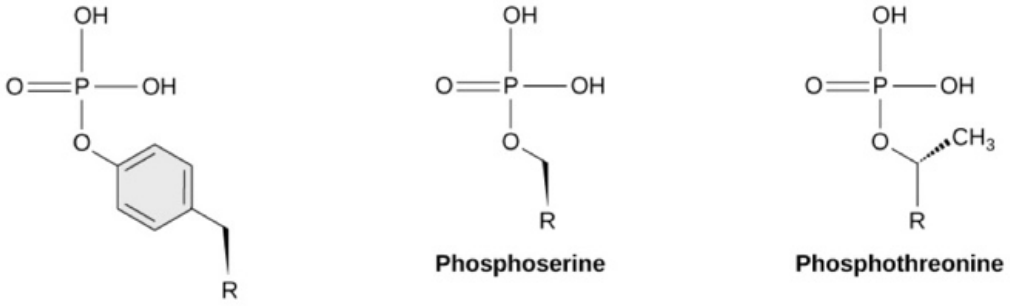



**Figure 1.** Structure of the phosphorylated side chains of tyrosine, serine and threonine. The phosphate is transferred on the hydroxyl group on the residue's side chains.

# Chapter 1
# Overview on the structural features and mechanisms of protein kinases

Protein kinases are enzymes that belong to the phosphotransferase group. These enzymes transfer a $PO_4^{3-}$ group from a high energy donor molecule, typically adenosine triphosphate (ATP) or analogues (such as GTP), to a substrate protein at a specific residue. The phosphate is transferred to the hydroxyl group present on the amino acid side chain according to the reaction:

$$C\text{-}ROH + PO_4^{3-} \rightarrow C\text{-}ROPO_3^{2-} + H^+.$$

In the case of serine/threonine kinases the target residue is either a serine or a threonine amino acid, whereas in the case of tyrosine kinases it is a tyrosine (Figure 1). For both types of kinases phospho-acceptor residues are usually surrounded by a typical consensus sequence. In addition, a small group of protein kinases called dual specificity kinases can phosphorylate tyrosine and serine/threonine residues [1].

## 1.1 Kinase families according to the human kinome classification

Kinases are low-efficiency enzymes. Their turnover rate, expressed as the number of substrate molecules converted into a product during a specific interval of time ($K_{cat}$ = n ° molecules/s), is extremely low (between 0.5/s and 1/s) compared to the majority of enzymes which have a $K_{cat}$ ranging between 10/s and 20/s. Indeed, in the phosphorylation-dependent signal transduction, the ability to selectively phosphorylate one substrate rather than another is more important than the amount of substrate that is modified [29].

Kinases are ubiquitous and they are essential for the correct modulation of various metabolic and cellular processes, especially in eukaryotic organisms. They are classified into three super families: serine/threonine kinases, tyrosine kinases and pseudokinases. In the human genome about 500 genes encoding kinases and pseudo-kinases are identified, constituting approximately 2% of all genes. The human kinome counts at least 518 members, organized in seven subfamilies [3].

The classification of kinases into different families originally proposed by Hanks and Hunter [30], is based primarily on sequence analyses of the kinase domain, on sequence and structural similarities in outer domains and on related biological functions. This classification describes 4 groups (cAMP-dependent protein kinase/protein kinase G/protein kinase C extended family or ACG kinases,

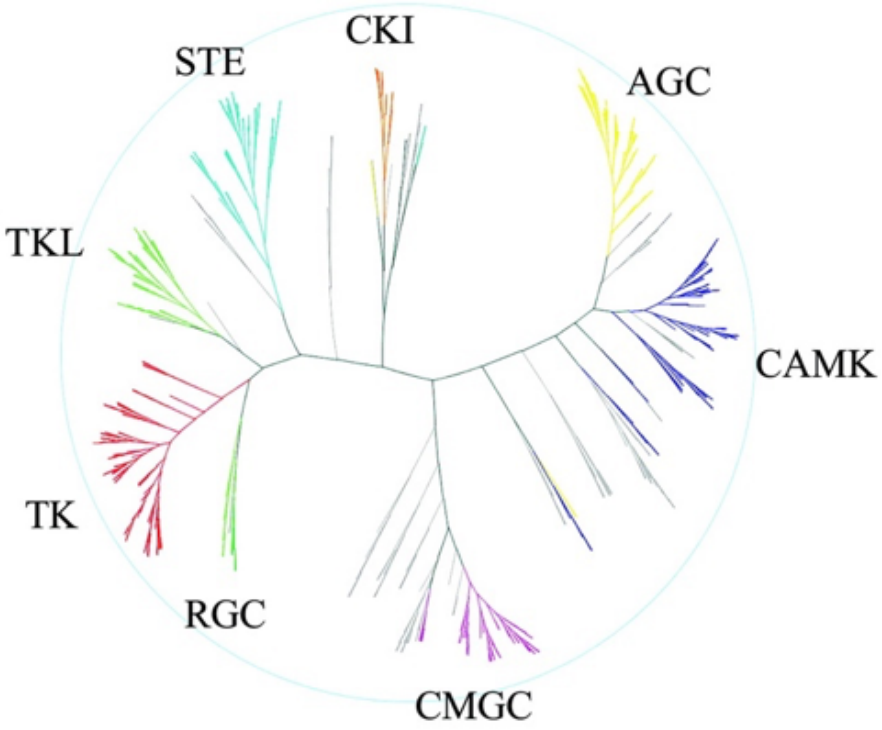



**Figure 2.** The human kinome evolutionary tree comprises eight major groups, 90 families and 145 subfamilies. The ninth group of heterogenous kinases is not shown. The classification takes count primarily on sequence similarities, similar primary structural features and biological functions. These groups count 518 kinases, 50 of them are considered to be inactive because of the lacking of key residues for the phosphorylation reaction (taken by Manning et al., 2002 [3]).

Calcium/calmodulin regulated kinases or MKs, CDK, MAP kinase, glycogen synthase kinase and CDK-like or CMGCs and protein tyrosine kinases or PTKs), 44 families and 51 subfamilies. A 5th group comprises heterogenous kinases which don't fit in any of these previous groups based on the criteria adopted for the classification. The increasing amount of structural and functional information on kinase proteins, led to the extension of the classification with the addition of other 4 groups (homologues of STE11 and STE20 or STEs, casein kinases or CK1s and tyrosine kinase-like or TKLs and receptor guanylate cyclase kinases or RGCs), 90 families and 145 subfamilies (Figure 2) [3].

*ACG family.* ACG group comprises the cyclic-nucleotide-dependent family, the protein kinase C family, the β-adrenergic receptor kinase (βARK), the ribosomal S6 family and other related kinases. ACGs are serine/threonine kinases that preferentially target residues flanked by basic amino acids such as arginine or lysine [31].

*CaMK family.* CaMKs activity is modulated by calcium/calmodulin thanks to a typical C-terminal binding domain. CaMKs are serine/threonine kinases, closely related to ACGKs and in a similar way they target residues flanked by basic amino acids [32].

*CMGC family.* CMGC kinases are serine/threonine proline directed enzymes that target residues included in a proline rich region and usually followed by a glutamate. CMGCKs are a large group of highly related kinases which exert numerous functions in signal transduction and that control cell metabolism and behavior. They are represented by several families of receptor and non-receptor kinases. These include the cyclin-dependent kinases (CDKs) and CDK-like kinases which control the cell cycle progression, the mitogen-activated kinases (MAPKs), which regulate cell behavior and metabolism and the glycogen synthase kinases (GSKs) discovered first as a key regulator of glucose metabolism [33].

*PTK family.* PTKs largely participate in the signal transduction, affecting cell behavior such as proliferation, survival, differentiation and other physiologic functions as organogenesis and tissue homeostasis. Impairment of PTKs activity is related to developmental issues and cancer. PTKs are exclusively tyrosine kinases [34].

*STE family.* STE kinases are classified into three families (STE7, STE11 and STE20), which trigger the MAPK signaling. The MAP4K STE20 targets and activates the MAP3K STE11, which in

turn phosphorylates the MAP2K STE7, leading to the activation of the MAPK cascade [35]. Several members of the MAPK signaling pathway, such as MAPK2K1 and ERK2, show dual specificity [1].

*CK1 family.* CK1 is a group of monomeric serine/threonine kinases, which regulate several cell functions such as cell proliferation, DNA repair and transcription, TFs shuttling, chromosomal segregation and other physiologic functions like the control of circadian rhythm [36].

*TLK family.* TLKs show high sequence similarities with TKs despite they are serine/threonine kinases. They are implicated in several signaling pathways such as angiogenesis, bone morphogenesis, neural development and survival [37]. RCGs are a small group of kinases related to the TKs, exclusive of metazoa. These receptors share a serine autophosphorylation domain, which is modulated by magnesium ions [38].

*Pseudokinases.* The kinase groups of the human kinome include at least 50 catalytically inactive kinases. These pseudokinases are enzymes that in their evolutionary path have lost the kinase function due to mutations in the kinase domain and which, however, are hypothesized to exert other functions in different signaling pathways as scaffolds or as allosteric effectors [39]. Kinases such as JAK and GCN2, possess two distinct kinase domains, one of which is catalytically inactive. The latter is thought to interact and modulate the activity of the active domain [40]. Some pseudokinases, including kinase suppressor of Ras (KSR2) and protein serine kinase H2 (PSKH2), exert exclusively scaffod functions. Receptor tyrosine kinases (RTKs) could act as dimerization partners for other kinases, inducing their activation [41].

*Atypical kinases.* A third group of early divergent kinases includes inactive kinases that show a low grade of similarity with the typical kinase domain, lacking of at least one key residue for the catalytic activity. Notably, this group contains also active kinases such as the human vaccinia-related kinase 1 (VRK1), an upstream regulator of p53, which catalyzes the phosphorylation reaction by a mechanisms that doesn't require the key residues of active kinases [42].

## 1.2 Structure of the kinase domain and activation mechanisms

Although kinases display high heterogeneity in their functional contributions, the architecture of the kinase domain and the structural changes required during the transition from the inactive to the active state are highly conserved [29].

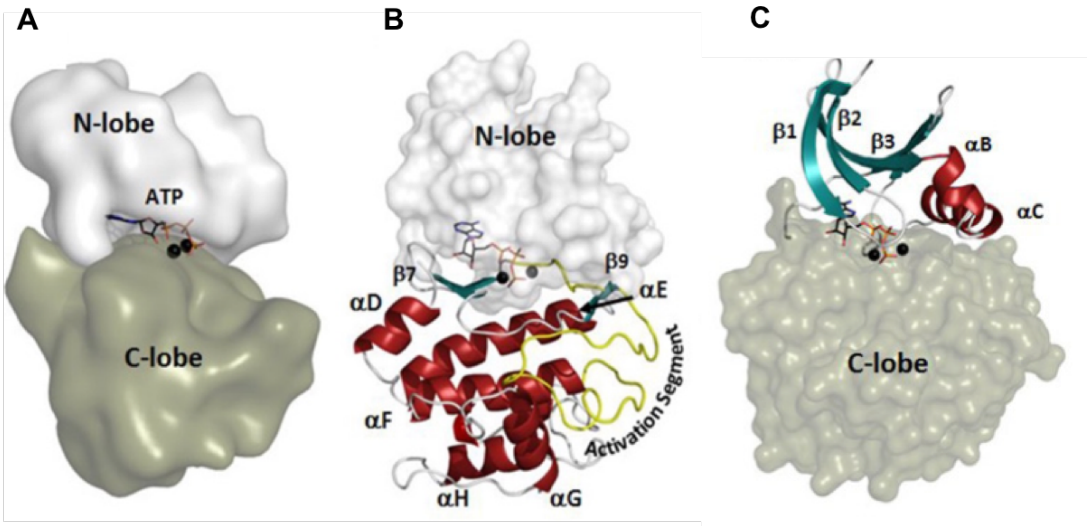



**Figure 3. A)** The kinase domain is organized into the N-lobe and C-lobe. **B)** The N-lobe is made by a five stranded β-sheet and a mobile αC-helix inserted between strands 3 and 4. The strands contain a group of hydrophobic amino acids that interact with the adenine ring of ATP. Two of them are part of the C-spine (taken from Taylor and Kornev, 2015 [56]). **C)** The C-lobe is made by 6 α-helices in eukaryotic protein kinases (EPKs). 2 short β-sheets participate in the stabilization of the extended region that contains the activation and the catalytic loops (taken from Taylor and Kornev, 2015 [56]).

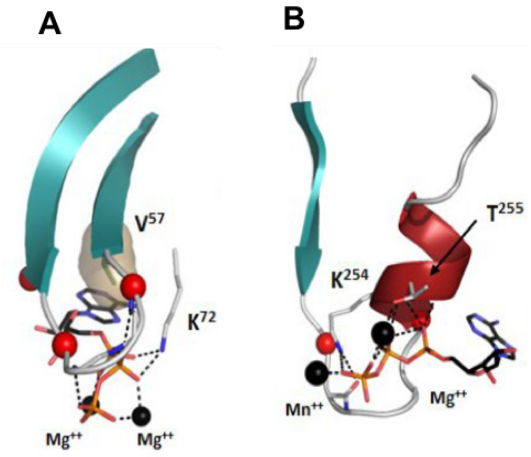



**Figure 4. A)** The G-loop is included between the β-strands 1 and 2 of the N-lobe. It harbors the glycine-rich motif that folds over the ATP, shielding the g-phosphate and orienting the adenine ring. **B)** The Walker loop (also called P-loop) is found instead of the G-loop (taken from Taylor and Kornev, 2010 [57]).

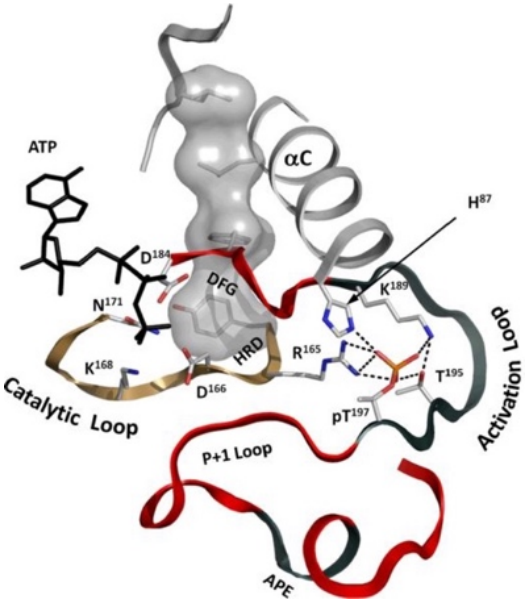



**Figure 5.** The catalytic region of the C-lobe is composed by a series of disordered loops harboring key residues and functional motifs for the correct activation and functioning of the kinase. The activation loop is targeted by upstream activating kinases and also shields the phosphate groups of the ATP. The catalytic loop folds over the phospho-acceptor residue of the substrate and induces the phsopho-transfer reaction. The P+1 loop stabilizes the substrate (taken from Taylor and Kornev, 2011 [4]).

### 1.2.1 The N- and C-lobes

The kinase domain is organized into two distinct lobes, conventionally called N-lobe and C-lobe (Figure 3A). The N-lobe consists of a 5 stranded β-sheet and two mobile α-helices (αC) that connect β3 and β4 (Figure 3B). The first 3 strands of the β-sheet (β1-β3) harbor two highly conserved motifs. The flexible loop between β1 and β2 (G-loop) has a glycine-rich motif (GXGXXG) and bends over the ATP γ-phosphate, while the strands accommodate and stabilize the adenine ring [43] (Figure 4A). The loop contains residues, which are a target for post translational modifications (PTMs) such as phosphorylation, ubiquitination, acetylation, O-GlcNAcylation and oxidation [44]. In some kinases, such as the adenylate kinase, the G-loop is substituted by the Walker-loop (P-loop) which exerts the same ATP coordination function despite it harbors a GXXXXGK(T/S) motif (known as Walker motif) and always connects a β-strand and an α-helix (Figure 4B) [45].

The β3 strand contains an AXK (Ala-X-Lys) motif, whose function is to further stabilize the ATP [46]. The lysine 72 (Lys72) of the AXK motif coordinates the ATP α and β phosphate groups and reaches out the αC-helix through a salt bridge with a conserved glutamate. All mutations of this lysine completely abolish the catalytic activity of the kinase domain [47].

The C-lobe is larger in size and exhibits higher variability among the kinases compared to the N-lobe. Usually, it consists of 6 α-helices (αD, αE, αF, αG, αH and αI) and one β-sheet made of four short β-strands (Figure 3C). Helices αD, αE, αF, αG and αH constitute a highly stable hydrophobic core, with the αG-helix exposed to the solvent. The main function of this core is to bind the substrate. The α-helices G, H and I, constitute a subdomain that exerts regulatory functions. In particular the αG helix, not only participates in the binding and stabilization of the substrate but also works as docking site for allosteric regulators and phosphatases that target the activation loop [48].

The C-lobe also comprises the catalytic loop and the activation segment (Figure 5). While the substrate is bound to the helical core, the catalytic loop located between β-strands 6 and 7 folds over the substrate and induces the transfer of the ATP γ-phosphate. This loop is quite rigid in active and inactive kinases as it tightly connects to the αF helix. At the N-terminus it harbors the conserved motif HRDLKXXN or HRDLKXXXN (in Ser/Thr kinases and Tyr kinases, respectively) [49]. The asparagine connects the phosphorylated activation loop (see below) to the catalytic loop by interacting with the phosphates and the aspartic acid acts as catalytic base that accepts the proton removed from the hydroxyl group of the phospho-acceptor residue during the phosphorylation reaction.

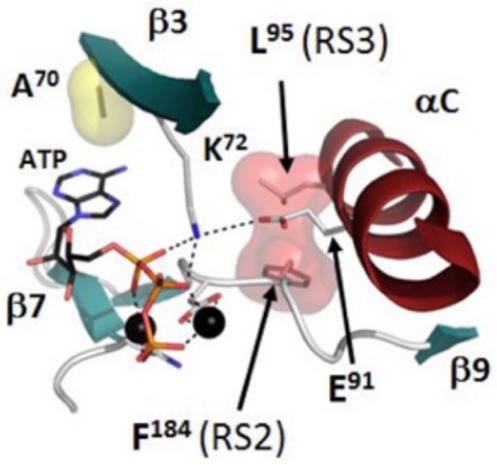



**Figure 6.** The K72 of β3 strand takes contact with the α and β phosphate of the ATP and stabilizes the ATP binding pocket by reaching the E91 on αC (taken from Taylor and Kornev, 2015 [56]).

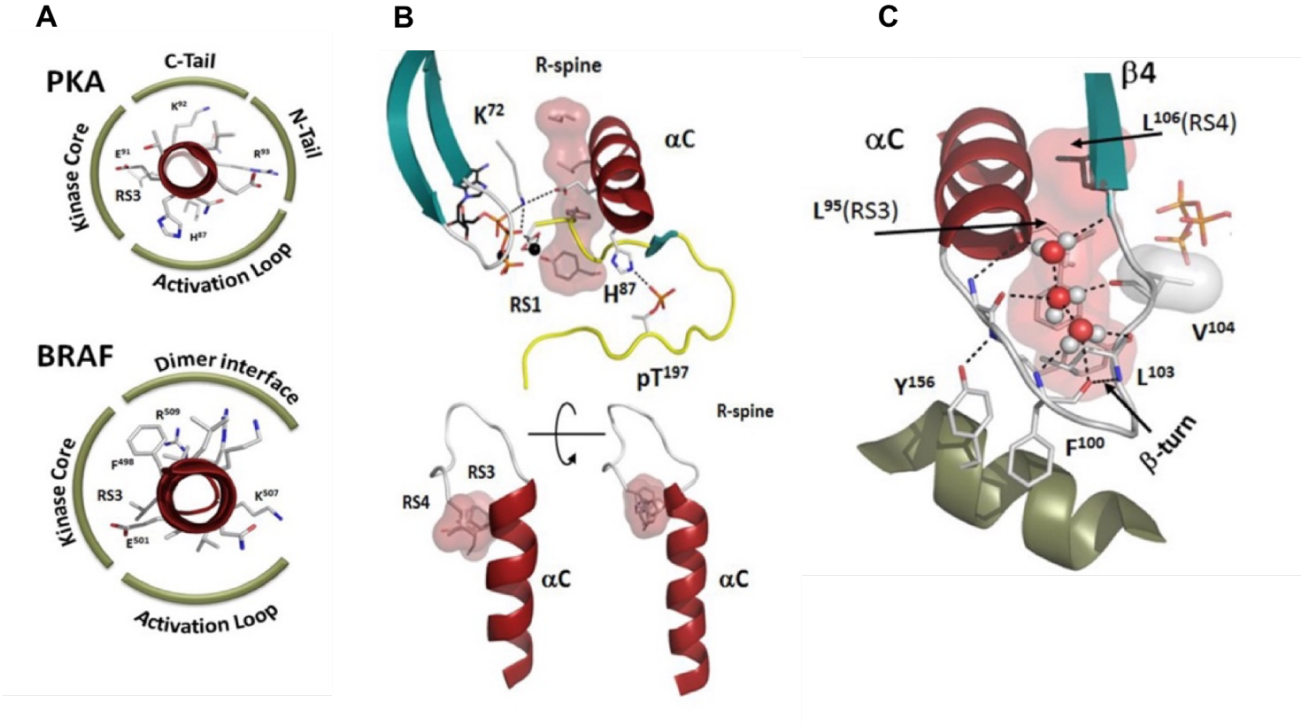



**Figure 7. A)** The αC-helix participates in the spatial arrangement and stabilization of different kinase domain elements plus other structural domains outside the core, such as the C- and N- tails in the protein kinase A (PKA). In other kinases such BRAF, the αC-helix is part of the dimerization interface. **B)** The αC-helix shifts away from the catalytic loop rotating around the αC- β4 loop defining the open configuration of the active kinase domain. **C)** the αC- β4 loop allows the movement of the αC-helix plus it reaches the αE-helix stabilizing the core (taken from Taylor and Kornev, 2015 [56]).

The activation segment shows the highest variability in terms of sequence and length among the kinase families. Generally, it measures 35-40 residues and it undergoes a disorder to order transition during kinase activation. The activation segment starts with the DFG motif located between β-strands 8 and 9. The conserved aspartic acid of the motif coordinates the $Mg^{2+}$ ion, which shields the ATP phosphate chain. In inactive kinases, the phenylalanine of the motif sits in the ATP pocket and shifts towards the αC-helix during the kinase activation [50]. After the DFG motif we can find the activation loop, which, in turn is followed by the P+1 loop. The main function of the P+1 loop is to stabilize the substrate and optimally orient the hydroxyl group of the phosphor-acceptor residue (P site) towards the ATP γ-phosphate by docking to the P site backbone and to the side chain of the following residue (P+1) [51]. The APE motif follows the P+1 loop. This motif takes contact with αH and with the connection loop between αH and αI helices and plays an important role in defining the active structure [52].

There are some key structural elements that mediate communication between N- and C-lobes. The β3-strand from the N-terminal lobe harbors a primary residue for the correct functioning of the kinase domain. Eukaryotic protein kinases (EPKs) have a universally conserved lysine (Lys72) that coordinate the α and β ATP phosphate groups and reaches out the αC-helix through a salt bridge with a conserved glutamate (Glu91) (Figure 6). Mutations of the lysine and the glutamate completely abolish the catalytic activity [47].

The αC-helix from the N-lobe is another important element of communication making several contacts residues of the C-lobe [46]. In inactive kinases, the N-terminus of the αC-helix is found in a disordered state and positioned near the activation loop, defining the typical closed configuration of the inactive kinase (Figure 7A). During the transition to the catalytically competent state of the PKA, the N-terminus of the αC-helix shifts away from the activation loop, drawing the open configuration of the active kinase domain where the histidine (His87) hydrogen bonds to the phosphorylated tyrosine (pTyr197) (Figure 7B) [53]. The C-terminus of the αC is strongly anchored to the C-lobe at the αC-β4 loop. The αC-β4 loop is a conserved loop that, despite being formally part of the N-lobe, dynamically belongs to the C-lobe. It is stabilized by the interaction with conserved residues of αC and β4 and protrudes towards the αE (Figure 7C). The loop works as a hinge around which the C-terminal domain of the αC rotates, allowing the helix to move away from the activation loop during the kinase activation (Figure 7B) [54].

The αF-helix is the architrave of the kinase domain. In fact, it is a large hydrophobic helix around which the R- spine and the C-spine (see section 1.2.2) are anchored at its N- and C-terminuses, respectively. In addition, the elements of the catalytic region (catalytic and activation loop, P+1 loop)

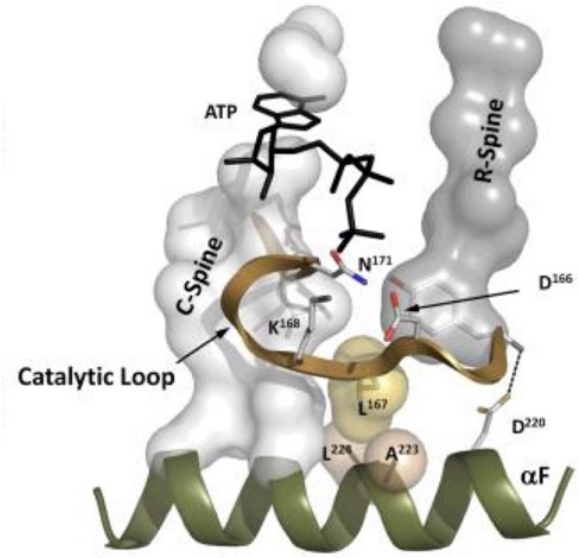



**Figure 8.** The R- and the C- spines are stably connected to the αF-helix which comes across the kinase core domain. Also, the catalytic region is stabilized by multiple contacts with the αF-helix (taken from Taylor and Kornev, 2011 [4]).

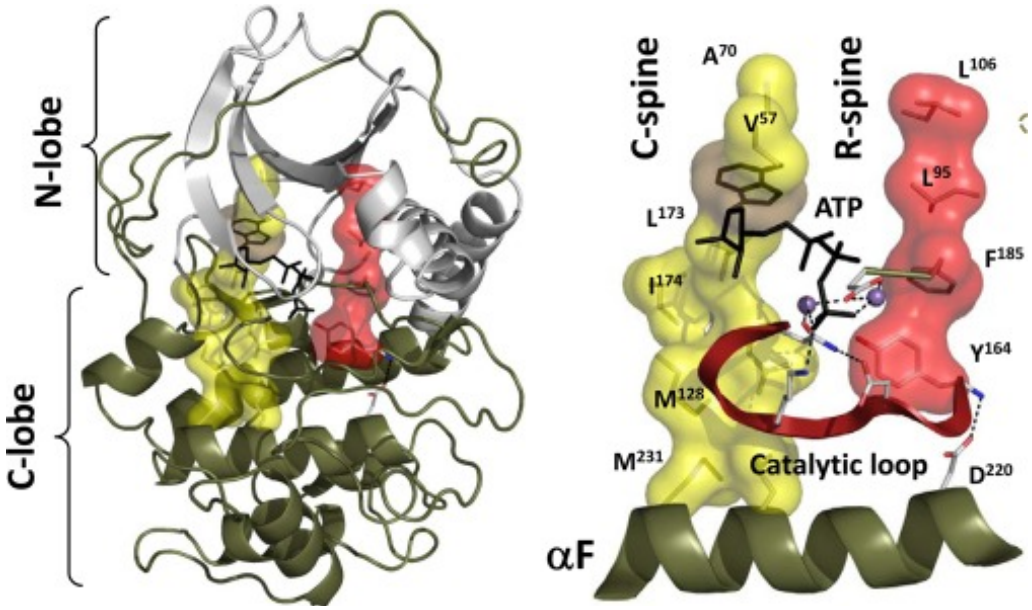



**Figure 9.** The hallmark of an active kinase domain is the assembly of the R and the C-spines. The R-spine is a transient structure made of 4 hydrophobic residues. R1 and R2 belong to the C-lobe, while R3 and R4 to the N-lobe. R1 comes from the HRD sequence that identifies the beginning of the catalytic loop while R2 from the DFG motif at the beginning of the activation loop. R3 and R4 are found in the αC and β4 strand respectively (taken from Shaw et al., 2014 [29]).

also assemble around the αF, representing a fundamental element for the general organization and stability of the kinase domain (Figure 8) [55].

### 1.2.2 Transition between inactive and active states of the kinase domain.

Kinases are generally maintained inactive by the activation loop, which bends over the catalytic loop, preventing any interaction with the substrates. The displacement of the activation loop is obtained by phosphorylation of conserved residues within the loop itself and, to date, two models have been described, namely the *trans* and *cis* phosphorylation models. The *trans* model predicts that upon dimerization, one kinase phosphorylates the activation loop of its partner. The *trans* phosphorylation is possible for both homo and heterodimers [29]. The *cis* autophosphorylation model is rare and not well characterized yet. The phosphorylation of the activation loop by the kinase itself, might be achieved during the protein synthesis or by the interaction with other effectors [58].

Despite their poor enzymatic efficiency, the transition from an inactive basal state to a catalytically competent configuration is rapidly achieved by the assembly of two pivotal spatial motifs of the kinase domain: the regulatory spine (R-spine) and the catalytic spine (C-spine) (Figure 9). With rare exceptions, these hydrophobic spines are highly conserved and ubiquitous in every active kinase, while they are never found in inactive kinases. Indeed, the phosphorylation of the activation loop induces the formation of hydrogen bonds within several elements of the kinase domain, leading to the formation C and R-spines [4]. The R-spine is assembled by four non-consecutive hydrophobic residues that belong to the C-lobe and the N-lobe, thus connecting the two lobes in the activate kinase core domain. The assembly of the R-spine is kinase-specific, determined by different regulatory mechanisms and precedes the assembly of the C-spine. Similarly to the R-spine, the C-spine is formed by residues that belong to both N- and C-lobes and it is completed and stabilized by the binding of the adenine ring of an ATP molecule, positioning the γ-phosphate towards the phosphor-acceptor residue of the substrate. Residues that align to create these spines belong to different elements of the kinase domain and they are rearranged in a discreet configuration only upon proper stimuli [4]. The R spine and the C spine are anchored to the N- and C-terminuses of the αF helix, respectively (see section 1.2.1) [55].

## 1.3 Kinase-substrate interaction

In contrast to the high degree of structural and functional conservation of the kinase domain, kinases are highly selective with regard to the substrates. Even in the case of closely related kinases, these enzymes are able to recognize and phosphorylate very different substrates. In general, selectivity can

be seen as a series of subsequent and progressively stringent levels that start from the primary discrimination of the type of phospho-acceptor residue and extend to include docking interactions outside of the catalytic region, allosteric regulatory mechanisms and cellular physiology such as increment of substrate concentration or cellular sub-localization.

### 1.3.1 Short linear motifs in the catalytic region determine a first selectivity criterium

Short linear motifs (SLiMs) are conserved short segments, from three to 11 amino acid residues, of protein sequences that have evolved and are optimized to mediate protein-protein interactions [60]. SLiMs are frequently found in intrinsically disordered regions and appear as conservation islands in non-conserved regions. SLiMs can be gained or lost by a protein by casual mutations in these regions, providing a mechanism of protein-protein interaction evolution and modification of existent protein networks [60]. Generally, only few residues are responsible for the interactions as they are the major contributors of the binding energy and also determine the specificity and the affinity. At least 80% of SLiMs are found in intrinsically disordered regions. The interaction mediated by SLiMs are weak, transient and reversible, making them an optimal tool to mediate different dynamic processes. SLiM-dependent interactions can be regulated by PTMs. SLiMs play also an important role in the regulation of enzymes such as kinases. The kinase domain contains motifs that participate in the core stabilization, in the determination of substrate affinity and selectivity and in the activation rearrangements [61].

*The DGF motif discriminates between Ser and Thr phospho-acceptors in serine/threonine kinases*. The kinase domain plays a role in determining the primary levels of selectivity of the kinases. This first level of selectivity is determined by the characteristics of the active site. In fact, within the active site there are minimal motifs, of one to three residues, which define the preference of the kinase towards the phospho-acceptor residue of the substrate. In the case of serine/threonine kinases, the +1 residue of the DFG motif at the beginning of the catalytic loop determines the selectivity of the enzyme. In particular, in serine kinases the DFG+1 position is occupied by a large hydrophobic amino acid (DFGϕ), while in threonine kinases the position is occupied by an amino acid with smaller side chains (DFGπ) [62].

*The P+1 loop participates in substrate recognition in serine/threonine kinases.* The second level of selectivity in serine/threonine kinases is determined at the P+1 loop, in the activation region of the kinase domain. P+1 loop harbors an APE motif and accommodates the phospho-acceptor residue of the substrate, orienting the side chain towards the γ-phosphate. The AGC, CaMK and STE

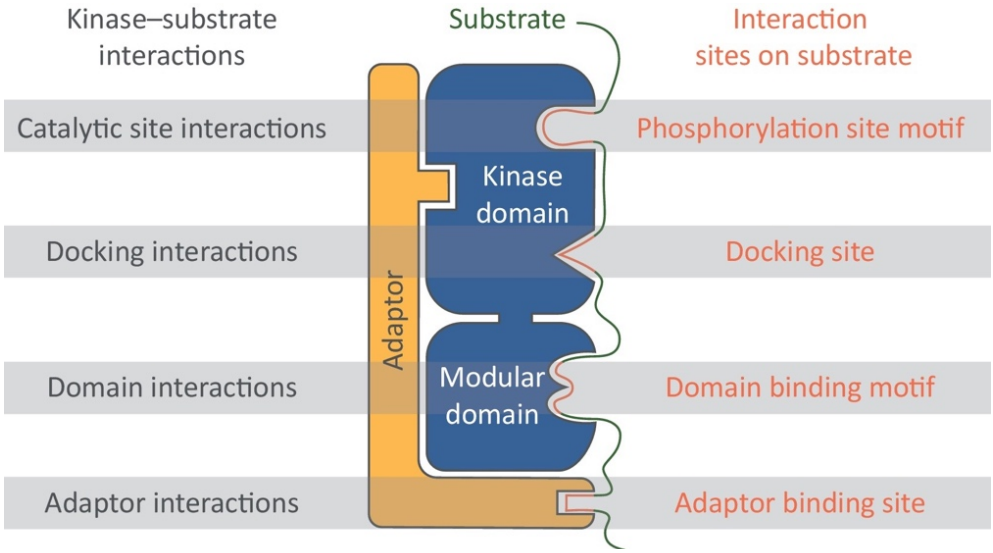



**Figure 10.** Kinase selectivity relies on different mechanisms that mediate protein-protein interactions. These mechanisms comprise interactions with proximal and distal docking groves in the kinase domain, interactions mediated by modular domains that do not belonging to the core, indirect interactions mediated by scaffold proteins (taken from Miller and Turk, 2018 [59]).

kinase families share common features in the P+1 loop. The AGC kinases PKA and PKB have a P+1 loop containing the conserved sequence (L/F)CGTPEYL. L/F, P and the second L confer a hydrophobic nature to the loop, thus, orienting the kinase towards a hydrophobic residue at the P+1 position on the substrate protein [64]. CaMK and STE family kinases have similar P+1 loop hydrophobic sequence features [65]. In CMGC kinases, a group which comprises MAPKs and CDKs, the P+1 loop has a secondary role in the substrate selection as it confers proline specificity to the P+1 position on the substrate. Despite the spatial organization of the P+1 loop is similar to the AGCKs, the P+1 pocket is occupied by the side chain of a flanking arginine in CDK2 or by the activation loop phospho-tyrosine in active ERK2. Also, the flat surface of the pocket prevents the formation of hydrogen bonds with the substrates [66].

The phosphorylation motif of CDKs in substrates is (S/T)*PX(R/K) (where * indicates the target residue) [67, 68], while MAPK phosphorylates substrates in the PX(S/T)*P motif [69]. The P+1 loop of GSKs is proline oriented but can accept also other substrates upon specific priming, obtained mainly by phosphorylation at P+4 position. In a similar fashion, SRPKs target serine/arginine rich proteins and are specific for P+1 proline and/or arginine. The pTyr residue found in ERK is substituted by a D in arginine-directed or by a serine in proline-directed SRPKs. In some other cases, such as for CK2s, the P+1 loop accepts all negatively charged residues [70]. The P+1 loop of CK1s exerts a poor selection on substrates since the P+1 pocket is entirely occupied by specific residues. These kinases recognize primed substrates, preventively phosphorylated in both C-terminal and N-terminal domains [71].

### 1.3.2 Docking interactions enhance substrate affinity

The mechanism described in section 1.3.1 are is not enough to determine the high level of specificity that characterizes each kinase. Other interaction mechanisms occurring out of the kinase domain participate to selectively bind substrates. Despite the kinase domain is highly conserved, kinases show wide differences in their overall architecture. Kinases use docking grooves within or outside their kinase domains to bind substrates. These grooves recognize specific linear motifs present on their substrates or in scaffold proteins that mediate the contact between kinases and substrates. Also, several structured domains, such as Src homology 2 and 3 (SH2/3) domains, provide contributions to kinase-substrate recognition (see section 1.3.3). These mechanisms are not mutually exclusive and can coexist to increase kinase selectivity [59] (Figure 10).

Kinases targeting a wide range of substrates are more likely to recognize SLiMs. This strategy is evolutionary convenient as single mutations in the motifs carried by the substrate can modify the

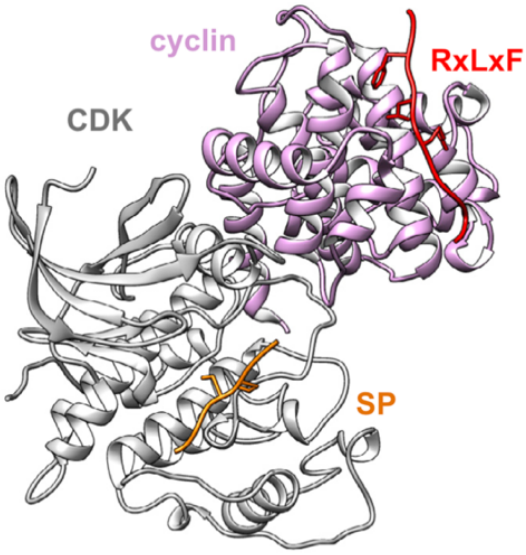



**Figure 11.** CDKs catalytic activity is regulated by the interaction with different cyclins that can shift the selectivity of the kinase towards one substrate or another. For instance, Cyclin A binds to the RXL (RXLXF) motif of the substrate whose phosphor-acceptor residue is properly oriented towards the catalytic domain (SP) of the kinase (taken from Faustova et al., 2021 [63]).

affinity towards one or another kinase. On the contrary, highly specific kinases interact with their substrates by specific domains and large interaction surfaces [33].

*Docking interactions mediated by the kinase domain*. MAKPs share a common phosphorylation consensus (S/T-P) but the various members show also a typical subset of targets. MAPKs have at least two common docking domains (CDDs), which mediate the interactions with substrates and other interactors such as phosphatases, scaffolding proteins and allosteric regulators. A first CDD, located in the C-lobe in opposition to the catalytic cleft, recognize the D-motif present on substrates, a region of 13-16 amino acids with sequence $X\Phi XX(R/K)_2X_{2-6}\Phi X\Phi$ for the or the kinase-interacting motif (KIM) with sequence $(L/V)X_2(R/K)_2X_5L$) [72]. This motif folds in a N-terminal helix and a stretched C-terminus and takes contact with the CDD by several hydrogen bonds. The same CDD recognize a second class of D-motifs, called reverse D-motifs due to the opposite orientation they assume, whose sequence features the sequence $\Phi X\Phi XX\Phi XLX_{4-6}\Phi$ $XX\Psi\Psi$ [73]. A second small CDD is found in proximity of the catalytic region and recognizes the F-motif $\Omega X\Omega$ ($\Omega$ is an aromatic residue) [74].

AGC kinases have a unique and conserved HM pocket (hydrophobic motif binding pocket) found in the N-lobe at the C-terminal region of the activation loop. The HM pocket recognizes and docks to disordered hydrophobic patches containing HM motifs (FXXF) on substrates. Interestingly, AGC kinases have a C-tail containing HM motifs, which fall into the HM pocket and act as self-regulatory elements or PDK1-interacting motif, upon phosphorylation [75].

SRPKs target splicing factors at SR (serine/arginine rich) domains. Phosphorylation occurs at different serine residues, generating sequence patterns that modulate the SR family-dependent alternative splicing. The docking grove in SRPKs is formed by helices F and G of the kinase core domain and binds to different SR domains (of the substrate in a non-hydrophobic based interaction [76].

*Docking interactions mediated by regulatory proteins.* In the case of CDKs, the interaction with substrates and inhibitors are regulated by cyclin proteins, which determine the specificity and increase the selectivity of the kinase. Cyclins harbor hydrophobic patches (hp) that act as docking domains by recognition of different SLiMs. For instance, mitotic cyclin B recognizes a short sequence containing the linear motif PXF [63], G2-phase cyclins bind to the G2 check-point substrates by docking to RXL$\phi$ motif while the S-phase cyclins recognize and bind to the NLXXXL docking motif. These two motifs are not mutually exclusive (Figure 11) [77].

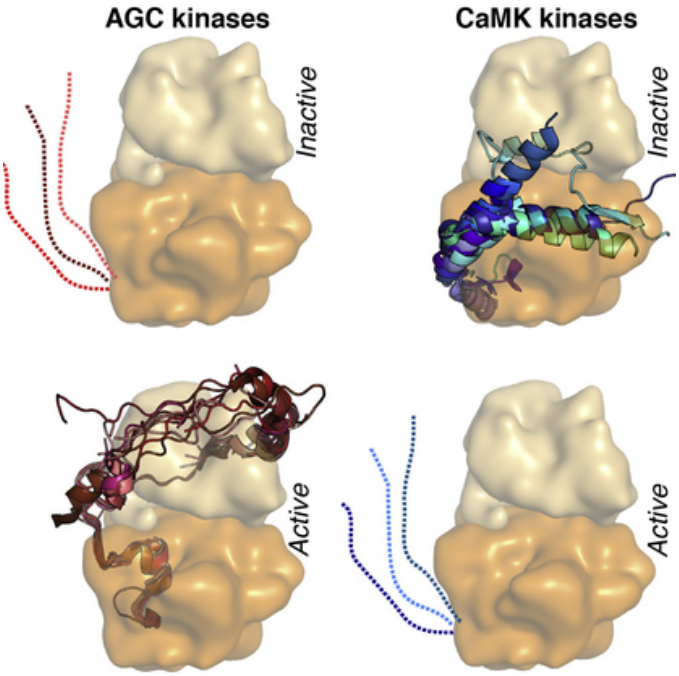



**Figure 12.** ACG and CaMK kinases show two typical mechanisms of regulation based on exclusive elements that interfere with the kinase core domain. ACG kinases have a partially disordered C-tail that hosts HP motifs, which in inactive kinase fall into the HP pockets of the kinase core domain. Phosphorylation of HP motifs induce the rearrangement and the subsequent shift of the tail that participate to assemble the R-spine. Similarly, CaMK kinases have a C-tail which docks into the ATP binding domain via substrate-like motifs. The tail shifts away, freeing the ATP binding pocket upon interaction with the calcium/calmodulin complex (taken from Gógl et al., 2019 [61]).

*Docking interactions that participate in the regulation of kinase activity.* Some kinase groups have evolved mechanisms of activation based on short motifs within and out of the kinase domain. The self-inhibitory/self-regulatory C-tail of AGC kinases contains HM motifs, which occupy the HM pocket, in proximity of the catalytic loop. Phosphorylation of the activation loop and of the phospho-acceptor residues of the HM motif induces the rearrangement of the C-tail and the assembly of the R-spine, rendering the kinase prone for catalysis [78]. CAMKs kinases also show a conserved helical C-terminal tail that contains SLiMs that mimic the substrate docking motifs. These kinases can reach the active conformation but they remain catalytically inactive as the C-tail occupies the ATP binding pocket or the substrate binding domain. The complex calcium/calmodulin binds to a docking site on the C-tail, inducing its displacement and opening the ATP binding pocket or allowing the substrate binding (Figure 12) [79].

*Wide interaction surfaces that increase specificity.* Some kinases target a limited number of substrates, showing a high specificity. In this case, interactions are mediated by wide surfaces and multiple contacts rather than linear motifs. For instance, LIM kinase binds to its substrates by a wide region of the C-lobe, orienting the substrate in a way that the phospho-acceptor residue enters the catalytic crevice bypassing any interaction with the activation loop [80]. Similarly, GRK phosphorylates the C-terminal tail of the GPCR at multiple sites upon binding. The interaction occurs between the kinase domain, the RH domain and the receptor C-terminal loops by direct contacts, defining a wide interaction surface [81].

*Substrate priming by phosphorylation.* Several kinases are unable to recognize and bind their own substrates in their native forms. In particular, the phosphorylation of specific residues on the substrate is optimized to generate phospho-motifs functional to the interaction with other proteins. For instance, the glycogen synthase kinase (GSK) phosphorylates its substrates upon priming by phosphorylation of a specific serine residue [82]. Frequently, the substrate priming by phosphorylation induces a phosphorylation cascade that regulates the activity and the functions of the substrate. In some case multiple phosphorylation leads to a signal amplification while differential phosphorylation patterns can change the functional parameters of a protein, modifying the affinity for other partners or substrates [83]. For instance, CDK1 counts several substrates, all of them containing multiple phosphorylation sites. The different phosphorylation patterns created by upstream kinases at different stages of the cell cycle, determine the affinity towards the kinase [84]. These clusters of phosphorylation sites usually are found in intrinsically disordered regions [85].

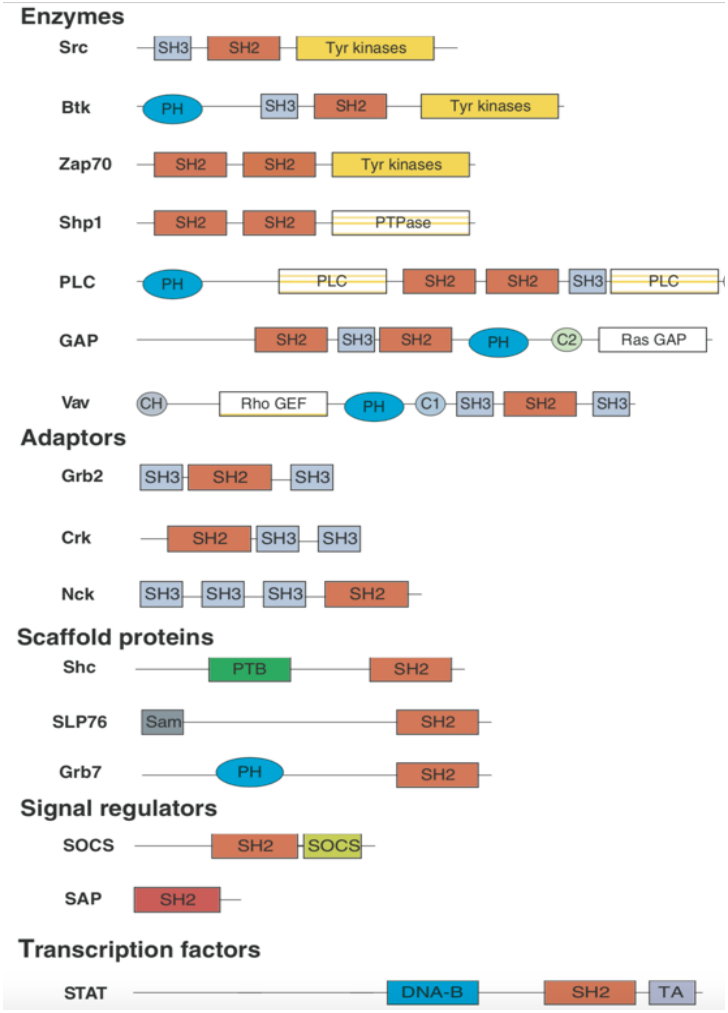



**Figure 13.** Interaction domains are self-stabilizing sequences of proteins with a defined and stable folding. These domains have the capacity to recognize and bind to different linear or non-linear motifs, mediating protein-protein interaction mechanisms. Interaction domains are classified in families based on the structural and sequence features. Proteins can host more than one interaction domain contemporarily. SH2, SH3 and PTB interaction domain families have notable importance in the kinases signaling mechanism as they generally recognize phospho-motifs (taken from Schlessinger and Lemmon, 2002 [86]).

### 1.3.3 Docking interaction mediated by ligand binding domains

Targeting and docking occur also by interactions mediated by ligand binding domains, which are structurally conserved interaction modules that can be classified into families based on sequence and structural analyses (Figure 13). These domains recognize SLiMs found in substrates and in other binding partners such as scaffold proteins or phosphatases [87]. Ligand binding domains are found both in kinases and in their substrates.

*SH2 domain.* The SH2 (Src Homology 2) domain is a structurally conserved protein domain of around 100 residues, organized in 2 α-helices and 7 β-strands that recognizes and binds to phospho-tyrosine motifs [88] with a length of 3 to 6 amino acid residues. SH2 domains are found in 111 human proteins [89]. SH2 domains are commonly found in substrates of RTKs and in accessory components of multi-subunit receptors for cytokines, antigens and other pro-inflammatory molecules. SH2 domain targets the RTKs phospho-tyrosine residues located in the C-terminal intracellular domain or in the kinase core domain, out of the activation loop. SH2 domain-containing proteins dock to the C-terminal domain of the RTK, correctly positioning for phosphorylation. Despite the majority of binding energy is given by the interaction with the pTyr [90], the specificity and the affinity are determined by the sequence immediately following the targeted pTyr [91].
Clinically relevant kinases, such as c-Abl which is responsible for the lymphoblastic leukemia, carry one or more SH2 domains. Thus, targeting the SH2 interaction interface could be used as a treatment for Abl-dependent leukemyas [92].

*SH3 domain.* The SH3 domain is a small docking domain composed of five β-strands organized in a β-sandwich and a short α-helix, which docks to the helical PXXP motif of interacting partners. SH3 domains have two di-peptide binding pockets, plus a third negatively charged pocket (specificity pocket). Thus, the orientation of the substrate is determined by the position of a positive charge immediately before or after the PXXP motif (+XPXXP or PXXPX-) [93]. Also, SH3 domains bind non-consensus binding motifs. These motifs do not adopt the classic helical conformation and may not occupy the XP binding pockets at the same time. Instead, substrates containing non-consensus motifs establish contacts with the SH3 surfaces (specificity zone). The SH3 domain also binds to other substrates by a mechanism that is independent of the PXXP binding pocket. These atypical interactions are based on contacts within extensive surfaces and other tertiary elements [94].

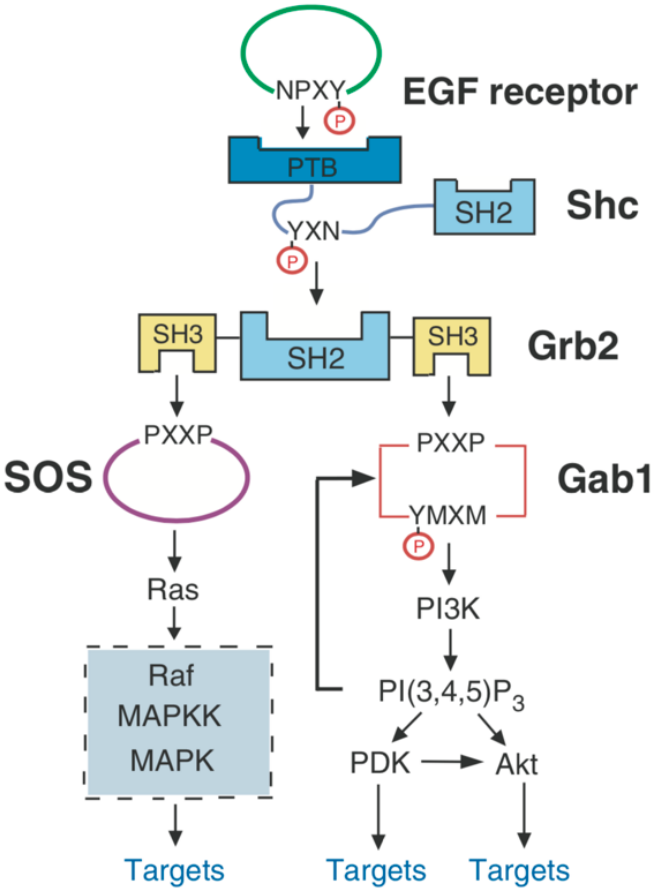



**Figure 14.** The signaling cascade is based on a series of subsequent phosphorylation reactions in synergy with docking domains present on adaptor and scaffold proteins. In the case of the EGFR, upon activation and auto-phosphorylation of the C-tail, Shc is recruited to the membrane by the interaction with its PTB domain and the newly generated phospho-motif. Then, Shc is phosphorylated and the new phospho-motif by Grb2 SH2 motif. Eventually, Grb2, thanks to 2 distinct SH3 domains, leads to the activation of the MAPK and the PI3K signaling (taken from Schlessinger and Lemmon, 2002 [86]).

*PTB domain.* Phospho-tyrosine binding (PTB) domains are wide-spread in the human interactome thanks to different mechanisms of interaction with their binding partners. PTB are usually found in cytosolic proteins that bind to cell surface receptors by docking to the NXXY motif. Specificity is improved by the residues flanking the core motif. Interestingly, the ability of PTB domains to bind the NXXP motif can be independent (or partially dependent) of tyrosine phosphorylation. For instance, PTBs in RTKs bind to their substrates in a phosphorylation-dependent manner, while other PTBs, such the ones found in Numb and Dab proteins, bind regardless or even preferentially to unphosphorylated NXXP motifs [95].

*Phospho-serine/threonine-containing motif binding domains.* pSer/Thr binding domain comprises several docking domain groups that preferentially dock to phospho-serine or phospho-threonine containing motifs. For instance, in 14-3-3 proteins, scaffold proteins implicated in the class switch recombination of B cells, contain docking domains that recognize the sequence RSXpSXP, while Fork Head associated (FHA) domains target phospho-threonine motifs in which the +3 residue determines further selectivity [96]. MH2 docking domains, typical of Smad proteins, bind phospho-serine motifs on the C-terminal domain of TGF β-receptors, positioning the substrate for optimal phosphorylation and subsequent dimerization. MH2 and FHA domains have similar conserved structures featuring a β-sandwich made of 11 β-strands connected by short loops and/or small α-helices [97].

*Cooperation between ligand binding domains.* Multiple docking domains cooperate in cell signaling. In the epidermal growth factor receptor (EGFR) signaling, upon receptor activation, the scaffold protein Shc binds to the phospho-motif NPXpY at the C-terminal region of the receptor, *via* the Shc PTB domain [98]. This results in the receptor-dependent phosphorylation of two phospho-acceptor tyrosine residues of Shc, generating two distinct pYXN phospho-motifs. Subsequently, Shc recruits the adaptor protein Gbr2, which docks to the Shc pYXN motif by a SH2 domain. Gbr2 has two distinct SH3 domains, which bind proline-rich motifs (PXXP) on substrates such as SOS and Gab1. SOS is a guanine nucleotide releasing factor which, once recruited at the membrane level, eventually leads to the activation of Ras and the triggering of the MAPK signaling pathway [99]. The adaptor protein Gab1 is phosphorylated by the EGFR in multiple sites generating the phospho-motif pYMXM. The p85 regulatory subunit of the phosphoinositide 3-kinase (PI3K) binds to phosphorylated Gab1 by docking to the pYMXM motif thanks to a SH2 domain, ultimately resulting in the activation of the Akt-dependent signaling pathway (Figure 14) [100].

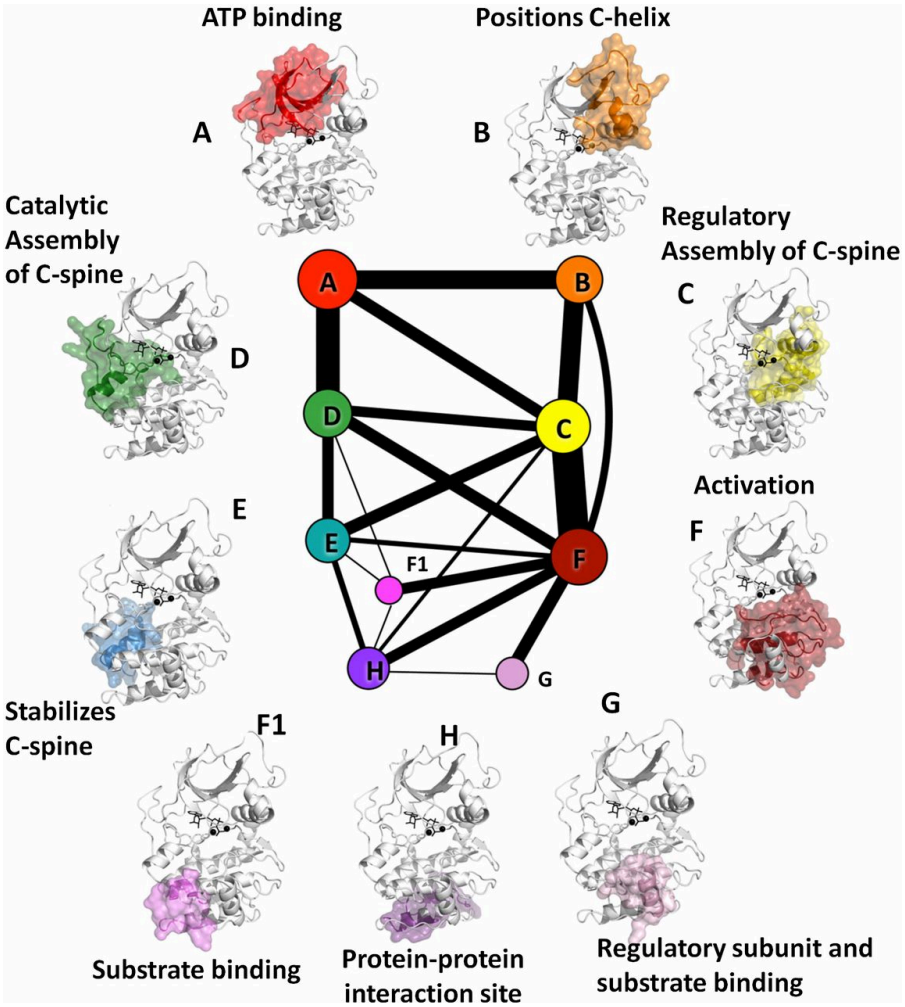



**Figure 15.** Schematic representation of dynamic communities of a MAPK showing the deep interconnection between different structural elements that constitute distinct functional units (taken from McLeoden et al., 2014 [6]).

## 1.4 Regulation of kinase activity

In cell signaling, hormones, neurotransmitters, pro-inflammatory molecules, growth and differentiation factors, physical and mechanical stress produce a biological answer by activating cellular kinases. A series of interlocked phosphorylation reactions generate a biochemical flow that evokes a specific cell response. Upon receptor activation, the upstream kinase initiates a phosphorylation cascade where substrates (downstream kinases or kinase activating factors) are subsequently activated [101]. Eventually, this cascade leads to the activation of transcription factors which start the transcription of specific gene sets, in a stimulus-dependent manner.

The fine tuning of the kinase activity occurs through an ensemble of different regulation mechanisms (Figure 15). PTMs, binding with small molecules or effectors and distal amino acid variations participate to regulate the activity of an enzyme inducing structural rearrangements and impairing the dynamic properties of the kinase domain. The assembly of these transient functional structures is a critical step that brings the kinase into a dynamical committed state [4], [5]. In this configuration, another subtle mechanism of regulation is observed, characterized by small sub-structural movements, intramolecular vibrations that define virtual dynamic communities modifying the features of the kinase activity, without altering the overall structure of the kinase domain [6].

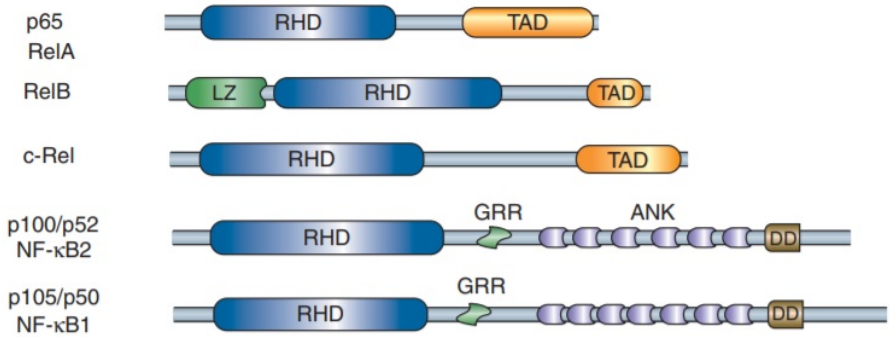



**Figure 1.** Domain architecture of NF-κB/Rel family proteins. RHD: REL homology domain; LZ: leucine zipper motif; TAD: transactivation domain; GRR: glycine rich sequence; DD: dimerization domain, ANK or ARD: ankyrin repeat domain (taken from Oeckinghaus & Ghosh, 2009 [106]).

# Chapter 2

# The IKK protein family in NF-κB signaling and beyond

The inhibitor of κB kinase (IKK) family consists of five proteins, namely the four active kinases IKK1 (IKKα), IKK2 (IKKβ), IKKE and Tank binding kinase 1 (TBK1), and the scaffold/regulatory NF-κB essential modulator (NEMO or IKKγ) protein. IKK1, IKK2 and NEMO assemble into different IKK complexes with variable stoichiometry and subunit composition. The functionally relevant IKK complexes include the canonical IKK (IKK1/IKK2/NEMO) complex and the IKK1 homodimer. These complexes are best known as master regulators of NF-κB signaling although they also play key roles in other important signaling pathways and transcriptional regulation. IKKE and TBK1 are known as IKK-related kinases and play a fundamental role in the interferon pathway. In this chapter I mostly focus on IKK1, IKK2 and NEMO and their regulatory mechanisms in NF-κB signaling.

## 2.1 Overview on the molecular aspects of the NF-κB signaling

The nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) is a family of eukaryotic inducible nuclear transcription factors (TFs), which regulate the expression of specific sets of genes implicated in several physiological functions, including immune and inflammatory responses. In addition, NF-κB is involved in other functions concerning cell behavior such as adhesion, differentiation, proliferation, autophagy, senescence and apoptosis.

### 2.1.1 NF-κB protein family

The NF-κB family consists of five members, grouped into two subfamilies, namely the Rel and NF-κB subfamilies. The Rel subfamily comprises RelA (p65), RelB and c-Rel, whereas the NF-κB subfamily consists of NF-κB1 (p105/p50) and NF-κB2 (p100/p52) (Figure 1). All members share a highly conserved DNA-binding/dimerization domain called REL homology domain (RHD) and, through this domain, they can form both homo and hetero-dimers.

Rel subfamily proteins are active TFs, which contain a transactivation domain (TAD) within their C-terminal regions that allows for the interaction with other basal transcription factors such as TATA binding protein (TBP), TFIIB, E1A binding protein 300KD (p300) and CREB binding protein (CBP)

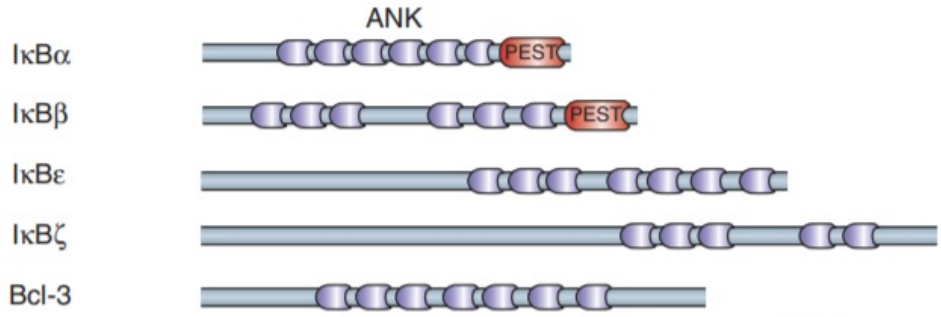



**Figure 2.** IκB family. ANK: ankyrin repeats; PEST: proline (P), glutamic acid (E), serine (S), and threonine (T) rich sequence.(taken from Oeckinghaus & Ghosh, 2009 [106]).

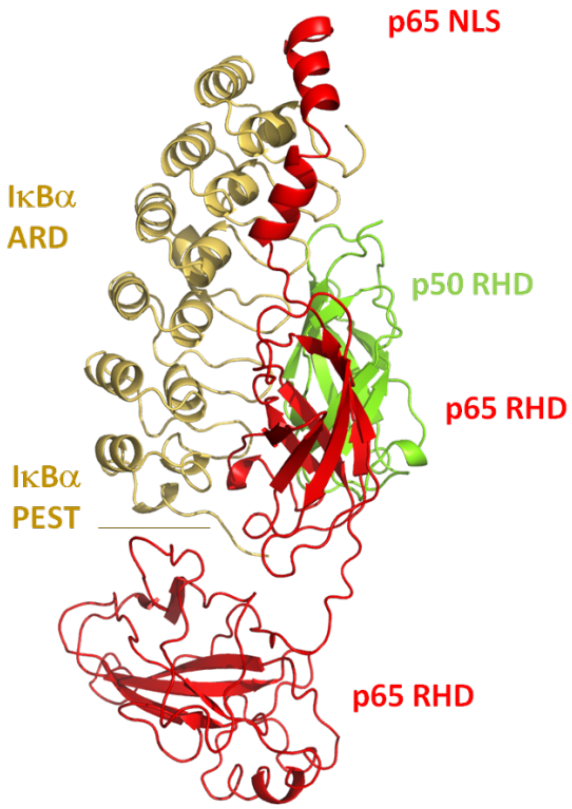



**Figure 3.** Structure of IκBα/p65/p50 complex obtained by X-ray diffraction at a resolution of 2,30 Å (PDB: 1NFI). The NF-κB proteins dimerize *via* their RHDs while IκBα masks the NLS (KRKR, 301-304) of p65 by direct interaction with the ARD (adapted from Ghosh et al., 2012 [111]).

[102]. The NF-κB subfamily proteins do not possess a TAD domain but feature multiple copies of ankyrin repeats within their C-terminal regions. The entire NF-κB1 (p105) and NF-κB2 (p100) proteins are present in the cytoplasm as inactive precursors. Upon activation of NF-κB signaling, these inactive precursors undergo partial degradation of the ankyrin repeat region, thus giving rise to the mature p50 and p52 subunits, respectively [103].

In general, NF-κB dimers recognize and bind to a 9-10 base pair DNA sequence called κB site, which shows a wide variability (5′-GGGRNWYYCC-3′; R: A/G, N: any nucleotide, W: A/T, Y: C/T). However, not all dimer combinations are allowed. The p50/p65 heterodimer represents the most common Rel dimer as it is ubiquitously expressed. On the contrary, dimeric complexes of RelA/RelA, RelA/c-Rel, RelA/p52, c-Rel/c-Rel, p52/c-Rel, p50/c-Rel, p50/p50, RelB/ p50, and RelB/p52 are expressed in specific cell types. RelB can assemble only heterodimers with p50 or p52, while p50/c-Rel is the main heterodimer constitutively expressed in mature B cells [104]. Specific stimuli and cell conditions participate to determine which dimers are activated, thus triggering the transcription of a related set of genes. These networks have both distinct and overlapping functions, which allow regulation of a given response, depending on the overall physiologic or pathologic conditions. The specific selection of a given NF-κB dimer species for activation of gene expression depends on a variety of factors including sequence variability of κB sites, the protein-protein interaction features of each dimer, the cell type and the cell cycle phase [105].

### 2.1.2 IκB protein family

The inhibitor of nuclear factor κB (IκB) family is a group of five proteins, which bind to NF-κB Rel subunits, causing their retention in the cytoplasm, thus preventing the activation of the NF-κB pathway in absence of proper stimuli. IκB proteins are IκBα, IκBβ, IκBε, IκBζ and Bcl-3 (Figure 2). All IκB proteins are characterized by the presence of five to seven ankyrin repeats, which constitute the ankyrin repeat domain (ARD), responsible for the interaction with the RHD domain of NF-κB proteins. This latter interaction masks the nuclear localization sequence (NLS) of NF-κB TFs (Figure 3), leading to the sequestration of IκB/NF-κB complexes in the cytoplasm. IκBα and IκBβ contain a PEST sequence (a peptide sequence rich in these amino acids) within the C-terminal intrinsically disorder region, which increases susceptibility to proteolysis [107].

IκBs bind to NF-κB complexes with different selectivity. IκBα and IκBβ inhibit preferentially c-Rel and RelA containing dimers, with IκBα displaying a higher affinity for the complex assembled by RelA and p50 (Figure 3). Bcl-3 and IκBζ mainly associate with p50 and p52 homodimers [108]. Other differences between IκB members regard timing and mechanism of expression. Whereas IκBα is constitutively expressed, IκBβ is expressed in a stimulus-dependent manner and with a slower

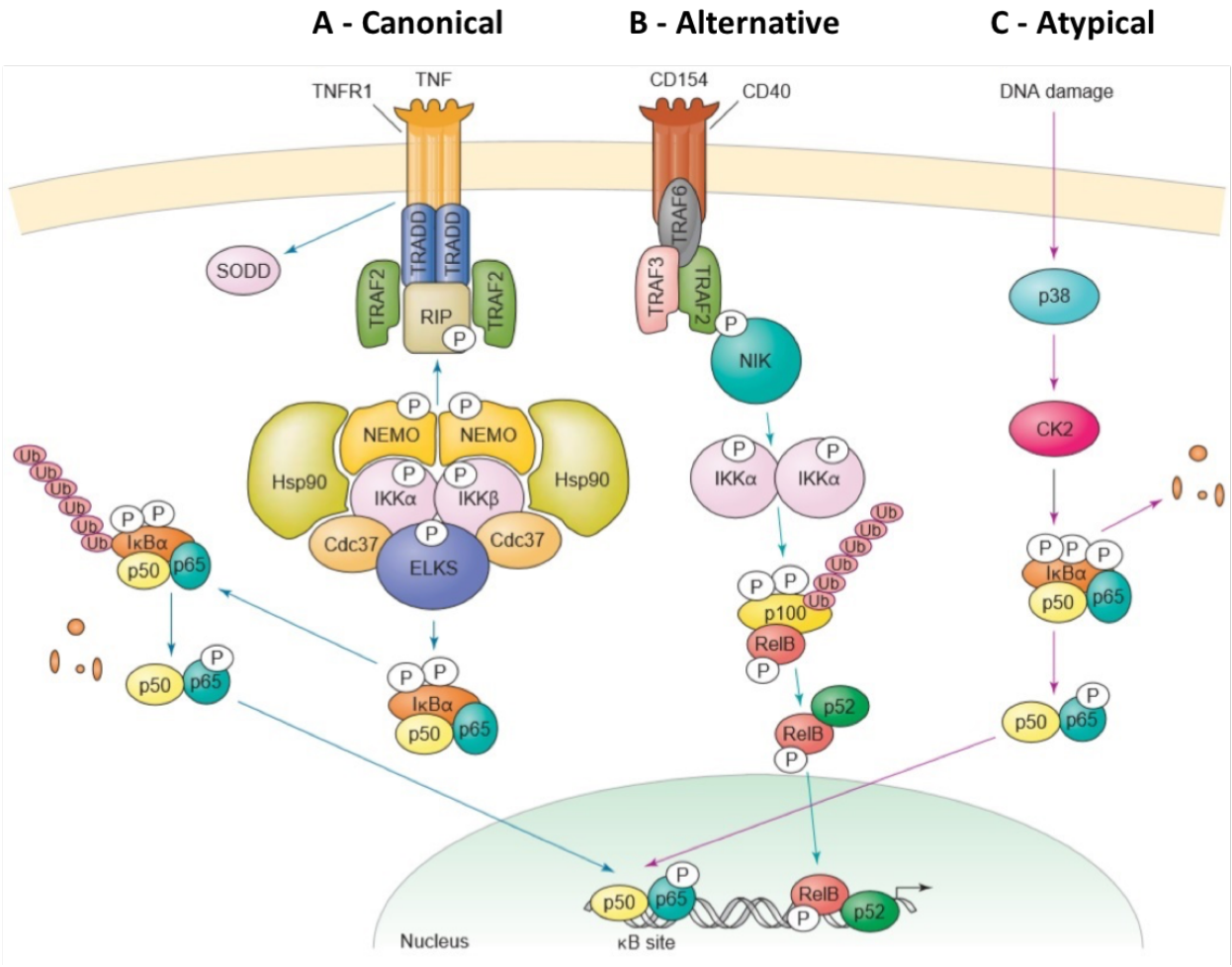



**Figure 4.** Schematic representation of the NF-κB activation pathways: **A)** canonical, **B)** alternative and **C)** atypical pathways (taken from Viatour et al., 2005 [118]).

kinetics [109]. In spite of this, IκBα and IκBβ seem to have overlapping functions, since the knockout of one of the two does not affect significantly the NF-κB response [108].

Interestingly, Bcl-3 and IkBζ are nuclear proteins that play a dual role in the control of gene expression. Bcl-3 contains a transactivation domain, which binds to p50 and p52 homo and heterodimers. This leads to removal of homodimers lacking of the TAD domain from κB sites and promotes the activity of heterodimers [103]. On the contrary, over-expressed Bcl-3 inhibits NF-κB activation by competition with other NF-κB containing TAD members to target DNA responsive elements. This double behavior of Bcl-3 is under the control of PTMs [110].

### 2.1.3 Canonical, alternative and atypical NF-κB pathways

Two main NF-κB signaling pathways have been described, namely the canonical or classical pathway and the alternative pathway. In addition, a number of other pathways of NF-κB activation, collectively named atypical pathways, also exist. The three pathways are selectively activated based on the nature and the combination of stimuli, the target receptor, the cell type considered and the NF-κB dimer composition [112]. The canonical and alternative pathways depend on IKK kinase complexes, whereas the atypical pathways of NF-κB activation are IKK independent (Figure 4).

*Canonical pathway.* The canonical pathway has a fast activation kinetics. A key role in the regulation of this pathway is provided by the canonical IKK complex, whose core consists of one copy of each of the two catalytic subunits (IKK1 and IKK2) and two copies of the NEMO regulatory subunit. Several proteins have been shown to associate to the canonical IKK core complex, including chaperon Hsp90 and cdc37, thus giving rise to a heterocomplex that ranges between 700 and 900 kDa [113].

The canonical pathway is activated by a wide range of pro-inflammatory factors and cytokines such as tumor necrosis factor α (TNFα) interleukin-1 (IL-1), Ectodysplasin A1 and 2 (EDA1/2), lipopolysaccharide binding protein (LBP), lymphocyte antigen 96 (MD2), major histocompatibility complex (MHC) and viral antigens. In the presence of these molecules, cellular immunity receptors, in particular the tumor necrosis factor receptor (TNFR), the Toll-like receptor (TLR) and the T cell and B cell receptors (TCR and BCR) induce the recruitment at the cytoplasmic domain of different scaffold proteins and enzymes including the TNF-receptor-associated factor 2 (TRAF2), the TNF-receptor associated death domain protein (TRADD) and the receptor interacting protein 1 (RIP1).

TRAF2 is a scaffold protein, which mediates the establishment of a signaling platform assembled by the E3 ubiquitin ligase TRADD and the kinase RIPK1. The TRAF/TRADD complex polyubiquitinates RIPK1, which in turn recruits the TGFβ-activated kinase (TAK1) through the

interaction with TAK binding proteins (TAB) 2/3, which recognize and bind to K63-linked polyubiquitination chains [114]. At the same time, RIPK1 also recruits the IKK complex to the signaling platform through the interaction of the UBAN domain of the NEMO subunit with the polyubiquitin chains of RIP1 [115]. Hence, TAK1 activates the IKK1 catalytic subunit by phosphorylation of serine residues 176 and 180 within the activation loop of IKK1. In turn, IKK1 activates the IKK2 subunit by phosphorylating serine 177 and 181 of the activation loop of IKK2 [114, 116]. In addition to this widely accepted mechanism, other routes to IKK2 activation seem to exist, which require kinases such as PDK1, SGK and MEKK3. These kinases seem to act through recruitment mechanism similar to that of TAK1, which are based on the polyubiquitination of specific protein scaffolds [117]. Subsequently, activated IKK2 phosphorylates IκB proteins at specific serine residues of the signal responsive region (SRR) at the N-terminus (IκBα on Ser36 and Ser36, IκBβ on Ser19 and Ser23, IκBε on Ser18 and Ser22) [118]. Such phosphorylation generates a degron motif with the consensus DpSGXXpS that is recognized by the Skp1-Culin-Roc1/Rbx1/Hrt-1-F-box E3 ligase (SCF-bTrCP) through its β-transducin repeat containing subunit (β-TRCP) [119]. This induces the polyubiquitination of IκB proteins (K21 and K22 of IκBα), leading to their degradation *via* the 26S proteasome. Hence the NF-κB dimers can now translocate in the nucleus to induce gene expression [120].

Interestingly, while IκBα is quickly degraded with consequent fast rise of NF-κB activation, phosphorylation of IκBβ induces a slower degradation mainly due to the intrinsic features of the sequence containing the phospho-acceptor residues, which determines a lower phosphorylation efficiency by IKK [109].

The dissociation of the IκB/NF-κB complex is a crucial step for the pathway activation but other biochemical modifications are required for an efficient gene transcription. Indeed, in response to LPS stimulation, RelA, c-Rel and p50 are phosphorylated at a highly conserved serine residue located in the RHD (Ser276, Ser267 and Ser337, respectively), by the cytosolic protein kinase A (PKAc), which increases the nuclear translocation rate [121] or modifies the dimerization preferences [122]. These residues are also targeted by nuclear MSK1 kinase. The phosphorylation may impede the interaction between the N- and C- terminal regions of the NF-κB subunits, increasing the DNA binding affinity and inducing the recruitment of co-activator p300/CBP [123]. Ser536 in the TAD domain of RelA is another phospho-acceptor residue targeted by different kinases including IKK2. IKK2-mediated phosphorylation of this residue occurs as second cytosolic event after IκB dissociation and may optimize the nuclear translocation or increase the transactivation potential of RelA [124].

Stimulation of the canonical NF-κB pathway also induces IKK2-mediated activation of NF-κB1 (p105), which is processed to its active form p50 in a stimulus-dependent manner. In this process, IKK2 phosphorylates serines 927 and 932, creating a SCF β-TrCP binding site, which induces the

polyubiquitination of p105 that leads to the full degradation of most p105 [125]. Only a small amount is converted into active p50 and this depends on a glycine rich region (GRR) located at the C-terminus of the RHD domain able to break off the proteolytic process [126].

*Alternative pathway.* The non-canonical or alternative pathway is activated by a smaller number of molecules including lymphotoxin β, B cell activating factor (BAFF), the CD40 ligand and several human viruses [127]. The IKK species regulating this pathway are homodimers of IKK1. Activation of receptors such as LTβR, BAFFR, CD27, CD30, CD40, RANK, Fn14, TNFR2 and TLR4 [128] causes the recruitment of TRAFF2, 3 and 6, which in turn activate the NF-κB-inducing kinase (NIK). NIK activates IKK1 homodimers by phosphorylation in the activation loop. IKK1 dimer phosphorylates p100 at serines 866 and 870 located in the C-terminal region of the protein [129]. These two phospho-groups generate an SCF β-TrCP binding site, thus inducing polyubiquitination and subsequent partial cleavage and degradation of C-terminal inhibitory sequence [130]. The active p52 moves into the nucleus in form of heterodimer with RelB [131].
IKK1 also phosphorylates other residues of p100 RHD, in particular serines 99, 108, 115 and 123. These modifications may play a role in DNA binding or other protein-protein interactions [132]. Interestingly, p100 is mainly associated with RelB subunits but it also regulates RelA containing dimers in mature T cells through direct protein-protein interaction, inducing a negative feedback that participates in control of the canonical pathway [133].

*Atypical pathways of NF-κB activation.* The atypical pathways are IKK independent and induced by DNA damage due to ultraviolet irradiation or exposure to toxic molecules such as doxorubicin or by oxidative stress. DNA damage leads to activation of CK2 through the MAPK pathway. CK2 phosphorylates IκBα at residues within the PEST region (Ser283, Thr291 and Thr299), leading to calpain-mediated degradation of IκBα [134]. Reactive oxygen species induce instead the phosphorylation of IκBα at Tyr42 by PTK. This leads to the generation of a SH2 binding motif that is recognized by SH2 domain containing proteins, thereby inducing the dissociation of the IκBα-NF-κB complex[135].
Atypical activation of NF-κB pathway occurs also through modifications of the RelA subunit. Upon TNF stimulation, RelA is phosphorylated at Thr254 creating a pT-P motif. This motif is recognized by the peptidyl-prolyl isomerase Pin-1, which breaks off the RelA-IκBα interaction and prevents RelA SOCS-1-dependet ubiquitination and consequent degradation [136]. RSK1 is another kinase which phosphorylates RelA at Ser536 upon activation of tumor suppressor p53. This leads to the activation of the NF-κB pathway in an IKK-independent manner. Specifically, RSK1 phosphorylates nuclear

RelA, interrupting the nuclear/cytoplasmic shuttle and consequently increasing the nuclear fraction of RelA [137].

### 2.1.4 Other functions of IKK1 and IKK2 in NF-κB signaling

In canonical and alternative pathways, the IKK kinases activate NF-κB by phosphorylation of IκB proteins or p100/p105. In addition, IKK1 and IKK2 can also regulate NF-κB *via* other targets. IKK1 is found in the cytosol and also in the nucleus where it promotes chromatin remodeling and transcriptional regulation. Nuclear IKK1 targets several proteins such as the transcriptional co-repressor SMRT. This protein is a transcriptional repressor since it recruits the histone deacetylase 3 (HDAC3) to the promoter of NF-κB target genes such as IL-8 and other cell survival related genes [138]. Nuclear IKK1 simultaneously phosphorylates chromatin-bound SMRT-RelA protein complex, thus preventing the recruitment of HDAC3 to chromatin and favoring the recruitment of p300. In turn, p300 catalyzes the acetylation of RelA, which acquires full transcriptional activity [139]. IKK1 also interacts with CBP and RelA, which results in direct phosphorylation and subsequent acetylation of histones H3 at the promoters of NF-kB genes, thus inducing their transcription [140]. Particularly, IKK1-dependent phosphorylation of CBP shifts the affinity of CBP towards NF-κB promoters rather than p53, favoring cell growth [141]. On the other hand, in macrophages phosphorylation of RelA by IKK1 could also lead to a decrease of its transcription potential as it favors RelA degradation and turnover [142].

In T cells, IKK2 exerts an inhibitory activity against NF-κB gene expression by targeting the CARD11-BCL10-MALT1 (CBM) signalosome complex, a multi-protein complex, which is assembled upon TCR activation and includes also Bcl10. The task of the CBM complex is to trigger the canonical NF-κB pathway by recruitment and activation of IKK2. Subsequently, IKK2 targets Bcl10, causing the dismantling of the CBM complex and silencing of NF-κB signaling [143].

## 2.2 IKK functions beyond NF-κB signaling

IKK kinases are involved in a considerable number of pathways in a NF-κB independent manner.

*IKK1*. IKK1 favors cell cycle progression by phosphorylation of steroid receptor co-activator 3 (SRC-3) upon TNFα exposure, leading to the expression of cyclin D1 and other proliferation genes [144]. IKK1 is also implicated in the regulation of the immune response *via* the interferon (INF) pathway. IKK1 targets the IRF7 transcription factor in dendritic cells. Phosphorylated IRF7 triggers the expression of virus-inducible genes, such as the type I interferon gene and other INFα genes [145].

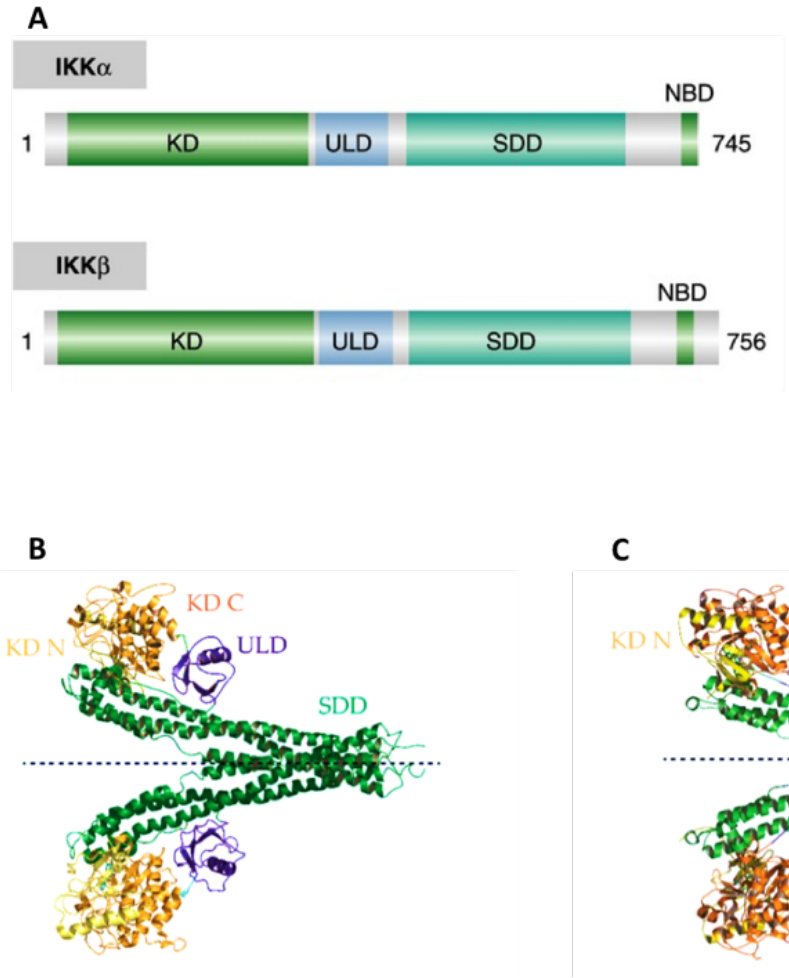



**Figure 5. A)** Schematic representation of functional domains of IKK1 and IKK2 homodimers. The catalytic subunits show a high degree of similarity in terms of domains and spatial organization. Comparison between crystal structure of **B)** IKK1 and **C)** IKK2 homodimers. The N-lobe and the C-lobe of the kinase domain are colored in yellow and orange respectively. The Scaffold/dimerization and the ubiquitin-like domain are colored in green and blue. The IKK1 and IKK2 homodimers show a high structural similarity with a 50% identity and approximately 70% of sequence homology. (taken from Prescott & Cook, 2018 [153]).

Notably, IKK1 can prevent cell proliferation by a kinase-independent mechanism. Indeed, nuclear IKK1 takes part, together with Smad3, in a transcriptional complex that promotes the expression of genes such as *Mad1*, which downregulate cyclin expression and increase sensitivity to apoptosis [146]. The MAPK signaling pathway is activated by IKK2 upon LPS, TNFα and B cell stimulator CD40 ligand, by phosphorylation of p105, which can form dimers with proteins other than NF-κB related TFs, such as Tpl2 (MAP3K). The processing of p105 to p50, on one hand, leads to the activation of the NF-kB canonical pathway and on the other induces the dissociation of Tpl2 that in turn activates MEK1 (MAP2K) and subsequently the classical MAPK ERK1/2 pathway [147]. By contrast, IKK2 also phosphorylates Docking protein 1 (Dok1), a scaffold protein that acquires the capacity of inhibiting ERK1/2, thus silencing the MAP signaling [148].

IKK2 represents a link between inflammation and tumor progression thanks its ability to target tumor suppressors such as FOXO3, TSC1 and p53 to degradation. FOXO3 is a pro-apoptotic factor that induces the expression of pro-apoptotic genes such as *PUMA* and *Bim* and also shuts down the expression of anti-apoptotic factors like FLIP. Phosphorylation of FOXO3 by IKK2 induces favors cell survival [149]. IKK2 phosphorylation of TSC1 allows for activation of mTOR signaling and enhancement of angiogenesis by expression of growth factors [150]. Finally the loss of p53 *via* IKK2 could be highly related to tumor progression and acquisition of malignancy in several solid cancers [151].

*IKK1 and IKK2.* In some cases, IKK1 and IKK2 phosphorylate the same substrate leading to different outcomes. For instance, β-catenin, a subunit of the cadherin protein complex implicated in the Wnt signaling, is targeted by both IKK subunits that generate different phosphorylation patterns. Phosphorylation of β-catenin by IKK1 leads to an increase of its transcriptional potential, while the phosphorylation by IKK2 results in repression of β-catenin activity[152].

## 2.3 Structure of IKK1, IKK2 and NEMO proteins

### 2.3.1 The catalytic IKK1 and IKK2 subunits

IKK1 and IKK2 have a high degree of similarity, with 51% of sequence identity and practically identical overall 3D structures. Both proteins comprise an N-terminal kinase domain (KD), a ubiquitin-like domain (ULD), a long scaffold/dimerization domain (SDD) consisting of 3 antiparallel α-helices and a C-terminal NEMO binding domain (NBD) (Figure 5A). Homo and hetero-dimerization of IKK1 and IKK2 are mediated by the SDD domains, which form a wide dimerization interface. In addition, several residue-specific contacts exist between the KD, ULD and SDD domains, which contribute to hold together the overall IKK1 and IKK2 structure. The resulting IKK1

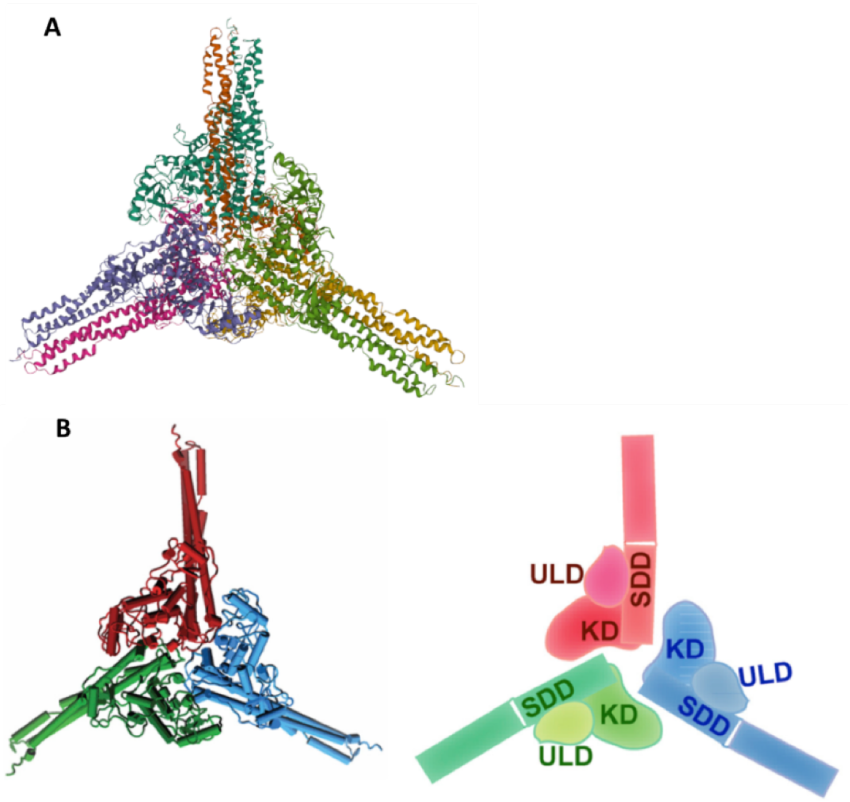



**Figure 6**. The hexameric organization of the IKK1 subunit is obtained by the interaction of three IKK1 dimers, which dispose themselves in a planar fashion, resulting in a 3-spike disc shape. **A)** Ribbon, **B)** cartoon and schematic representations of the hexamer (taken from Polley et al., 2016 [9]).

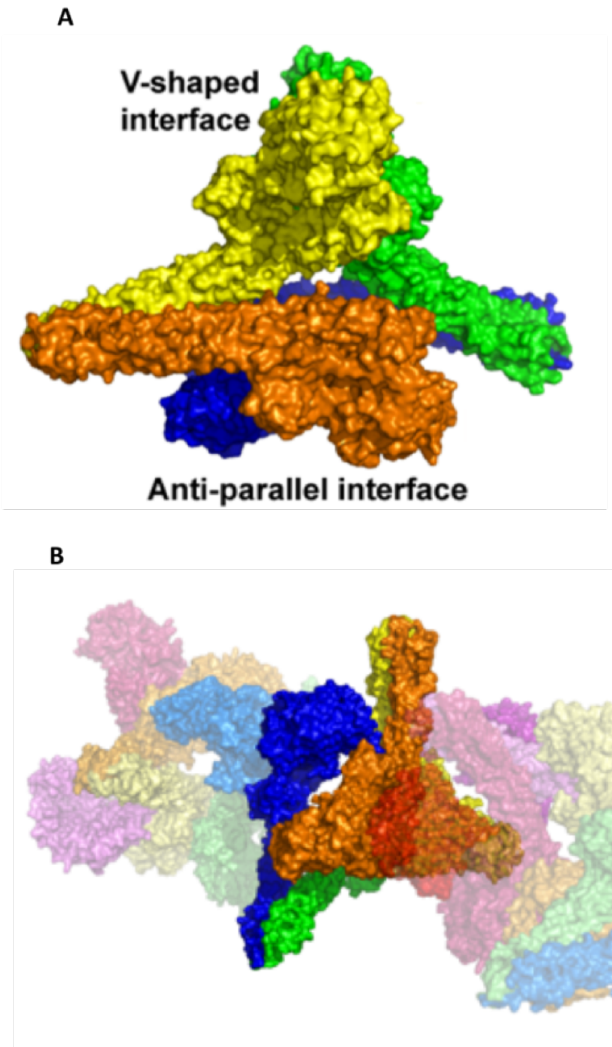



**Figure 7. A)** IKK2 dimers form tetramers with an anti-parallel disposition where the KD-ULD domains of a dimer interact with the SDD domains of a neighboring dimer. **B)** Tetramers pack on each other to form a continuous chain-like oligomer (taken from Polley et al., 2016 [7]).

or IKK2 homodimers have a characteristic scissors-like shape with a certain degree of flexibility. In fact, as a pair of scissors, the dimer swings from a close to an open conformation (Figure 5B-C).

*IKK1.* Human IKK1 is a 745 amino acid protein encoded by the CHUCK gene (ENSG00000213341), located on the chromosome 10. IKK1 is activated upon phosphorylation of Ser176 and Ser180 of the SXXXS motif within the activation loop. In 2016 Ghosh and collaborators determined two the structure of a large fragment of human IKK1 including residues 10-667 and the phosphomimic mutation S176E/180E in the activation loop that renders the kinase constitutively active (pdb:5TQW) [9]. In this work they first solved a cryo-EM structure, which was then used to solve the X-ray structure by molecular replacement at a resolution of 4.5 Å.
In these structures, three IKK1 dimers self-associate to form a hexamer. The hexamer shows a planar symmetrical shape, with a spherical core constituted by the KD domains and the SDDs domains protruding outwards forming a 3-spiked structure (Figure 6). The trimerization takes place by interactions between KD domains and between KD and SDD domains. More precisely, the KD domain of a dimer interacts with the SDD domain of a second dimer through specific key residues and secondary structures. This feature is exclusive of IKK1 and it may have biological relevance. Indeed even though the hexamer structure is not observed in cells, mutation of critical residues at the trimerization interface leads to impairment of the alternative NF-κB signaling pathway [9].

*IKK2.* IKK2 is a 756 amino acid protein, encoded by the gene IKBKB (ENSG00000104365) on chromosome 8. The IKK2 subunit is activated upon phosphorylation of Ser177 and Ser 181 in the activation loop by IKK1 or RIP1 [114]. The x-ray structure of the IKK2 protein from *Xenopus laevis* was published in 2011 (PDB: 3QA8, 3.6 Å resolution) [8], whereas two x-ray structures of the human IKK2 protein were published in 2013 (PBD: 4E3C, 4 Å resolution and PDB: 4KIK, 2.8 Å resolution) [154], [7]. For all structures the IKK2 construct is very similar, which corresponds to residues 1-669 of *huma*n IKK2.
In the *Xenopus* IKK2 structure the dimer is in a close conformation, whereas in the *human* structure it is in the open conformation. This open conformation is thought to allow for oligomerization of the dimers. The x-ray structure of human IKK2 revealed that homodimers assemble into asymmetric tetramers (dimer of dimers) (Figure 7A). The KD-ULD N-terminal portion of a dimer and the SDDs of the neighboring dimer make contacts to form a V-shaped interface. These tetramers interact with each other in the crystal to form an elongated continuous chain. Interestingly, mutations at residues critical for the V-shaped interface of the tetramer correlate with lower IKK2 activity in cells [7]. Hence, it is proposed that these tetramers might transiently appear during the activation of IKK (Figure 7B).

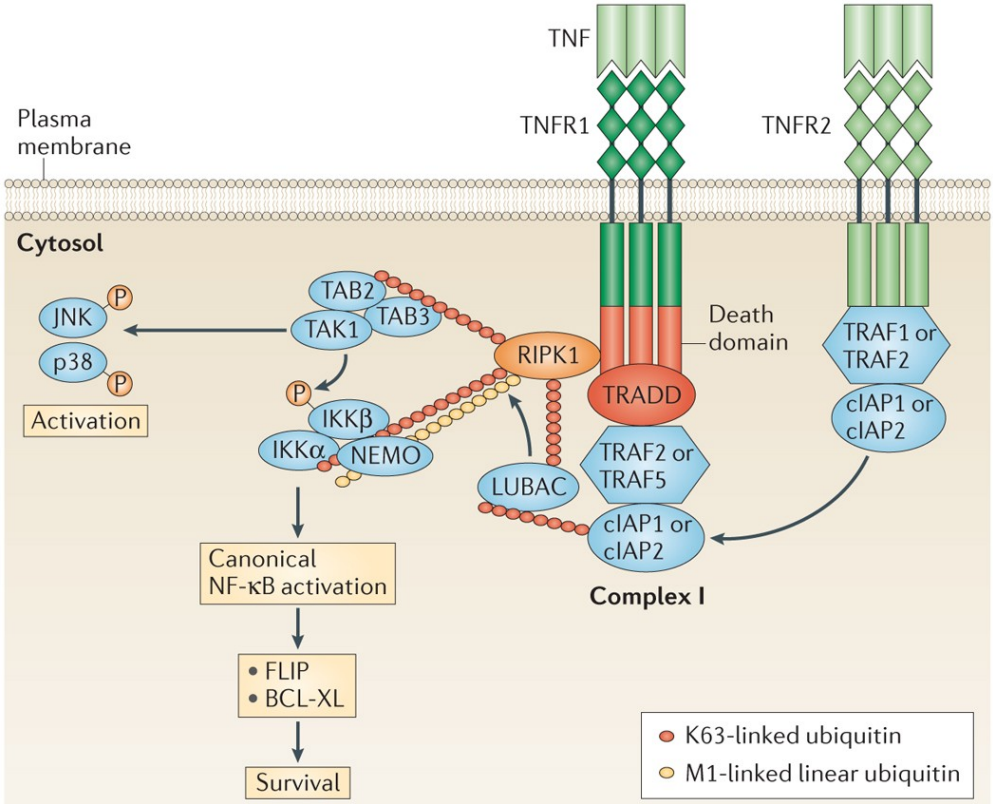



**Figure 8.** TNF activation induces the recruitment of the signaling complex made by TRADD and TRAF2/5. RIPK1 is polyubiquitinated by TRAF and functions as docking site for TAB2/TAB3/TAK1 complex. Modification of the RIPK1 polyUb chains with Met1 side chains by LUBAC recruit the IKK canonical complex by the interaction with NEMO. These two independent events lead to the activation of the IKK2 by TAK1, triggering the NF-κB canonical signaling and related pathways such as the MAPK signaling (taken from Brenner et al., 2015 [162]).

### 2.3.2 Structural and functional properties of NEMO

NF-κB essential modulator (NEMO or IKKγ) is a 419 amino acid protein, encoded by the IKBKG gene (ENSG00000269335) on chromosome X. NEMO is the scaffold and regulatory protein of the IKK canonical complex, which was discovered in 1988 in the laboratory of Alain Israel at the Institut Pasteur [155]. NEMO is one of the few, if not the only, non-redundant element of the NF-κB signaling pathway and, thus, it is considered the gatekeeper of the canonical pathway. Indeed, NEMO has the unique role of translating the wide range of the NF-κB activation stimuli into the recruitment and activation of the IKK canonical complex. NEMO is also needed for a correct IKK activity since it is necessary to target IκB proteins by IKK2 [10]. In the canonical complex, each IKK1/2 heterodimer binds to two copies of NEMO, which interact with each other through coiled coil structures along the entire length of the molecule.

*NEMO signaling.* NEMO signaling is strongly regulated by polyubiquitination (polyUb) [156]. NEMO recognizes and binds to a hybrid system of polyUb chains assembled by a main set of Lys63-linked polyUb enriched with Met1-linked polyUb branches. The affinity of NEMO for Met1-polyUb chains is one hundred times higher than for Lys63-polyUb [157]. NEMO is primarily implicated in the activation of the NF-κB canonical complex by addressing the IKK activity towards the IκB proteins, upon stimulation of immune receptors. Upon TNFα stimulation, TRADD is recruited by the C-terminal region of the trimerized TNF receptor together with the ubiquitin ligase TRAF2/5 and the scaffold protein RIPK1 [158], which is modified by the addition of Lys63-linked polyUb chains. The complex TAB2/3 and the IKK activating kinase TAK1 is recruited to this signaling platform by docking to the polyUb chains of the C-terminal ubiquitin binding domain (UBD) of TAB2/3 [159]. Independently, the IKK canonical complex is recruited to this sub-membrane platform by NEMO, which docks to the polyUb chains of RIPK1 through its UBDs. Consequently, IKK2 is activated by the phosphorylation of the activation loop by TAK1/IKK1. Interestingly, proper activation needs Lys63 polyUb of NEMO. It is possible that the proper positioning of IKK2 is achieved by a second interaction between poly-ubiquitinated NEMO and TAB2/3 *via* the TAB zinc finger domain [160] (Figure 8).

The NF-kB activation triggered by IL-1 exposure is thought to rely on a similar polyUb-based mechanism. IL-1 receptor activation induces the recruitment of a signaling complex called Myddosome (Myd88/IRAK4/1/2). IRAK1/2 is polyubiquitinated by TRAF6 and recruits TAB/TAK1 complex. At the same time, Met1 polyubiquitination of RIP1K4 recruits the IKK canonical complex

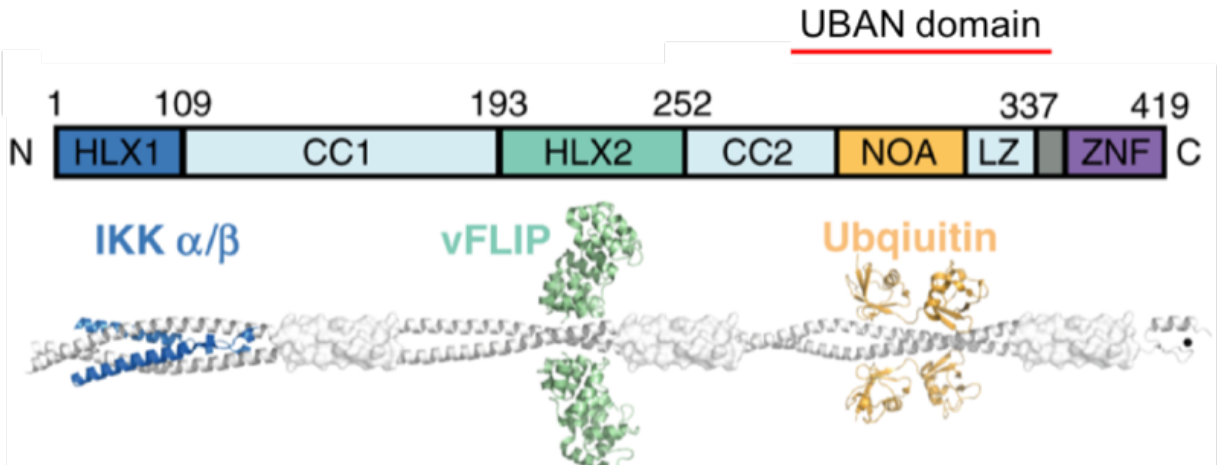



**Figure 9.** NEMO domain architecture and structural model of the NEMO dimer (gray) displaying the interactions with IKK1 and IKK2 (blue), vFLIP (green) and di-ub iquitin (yellow) (taken from Sadek et al., 2020 [165]).

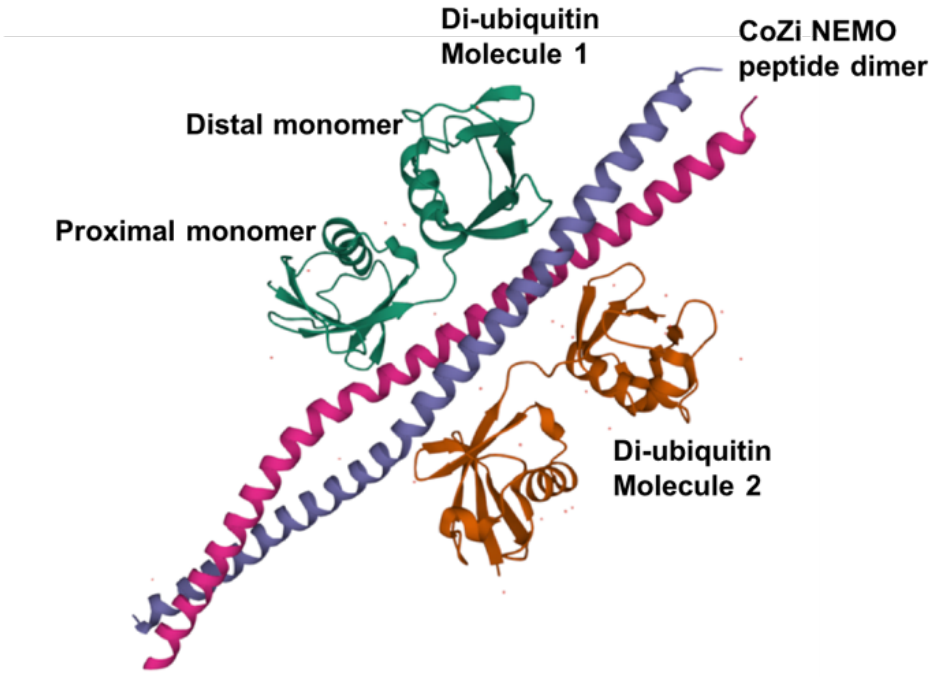



**Figure 10.** The structure of NEMO CoZi region homodimers bound to two molecules of di-Ub, resolved at 2.7 Å, reveals how linear polyUb chains bind to the UBAN motif. Observing the green di-Ub molecule 1, the monomer atop is the distal, while the lower is the proximal molecule. (PDB: 2ZVO) [14].

*via* the interaction with NEMO. These two independent events lead to the activation of the IKK2 by TAK1, triggering the NF-kB canonical signaling and related pathways such as the MAPK [161].
In addition to NF-kB signaling, NEMO functions extend to other pathways, such as the ATM pathway in response to DNA damage [163]. NEMO also interacts with calbindin D28K, a brain calcium binding protein which is expressed upon neurotrophic grow factor stimulation by NF-κB activation, with the transmembrane synaptotagmin 1 implicated in the release of neurotransmitters in the synaptic cleft, and with CDK2 [164].

*Structural properties of NEMO.* NEMO has an extended helical structure with a significant amount of disordered regions. In the canonical IKK complex, two copies of NEMO dimerize *via* extended coil-coil regions (Figure 9). The N-terminus of NEMO contains the kinase binding domain (KBD) that interacts with the NBDs of IKK1 and IKK2. This region is followed by a first coiled-coil domain (CC1). The central portion of NEMO contains the second helical domain (HLX2), which corresponds to the TANK binding domain (TBD) and that presumably mediates the interaction with other binding partners, regulatory proteins and also viral products such as the protein v-FLIP. The HLX2 is followed by a second coiled-coil domain (CC2), the NEMO-optineurin-ABIN (NOA) and a leucine zipper (LZ) domains. These latter three domains harbor the ubiquitin-binding in ABIN and NEMO domain (UBAN) [11]. Finally the C-terminal portion includes a Zinc Finger domain (ZNF) [161].
In the native state, NEMO appears to be a molten globule composed by a core of α-helices interspersed by disordered regions. Interestingly, NEMO assembles an extended lattice network by oligomerization in cooperation with IKK2 near the surface membrane. These networks are constitutively present and they feature clusters of NEMO connected by stretched NEMO dimers in presence of Ub chains. Upon IL-1 exposure, the lattice structure rearranges and thus it might be implicated in IKK activation [13].
Due to the high content of intrinsically disordered regions, all crystallization attempts of full-length NEMO failed. However, several structures of fragments of the protein have been determined by x-ray crystallography [11, 12, 13, 14, 15].
The UBAN motif is particularly important for the correct functioning of NEMO. The structure of two molecules of di-ubiquitin bound to a dimer of the UBAN motif shows two distal and two proximal ubiquitin molecules (PDB: 2ZVN). The distal Ub unit docks to NEMO by a hydrophobic patch, while the docking with the proximal Ub unit is mainly based on polar interactions (Figure 10) [166]. It is thought that the structural features of NEMO result in a high degree of flexibility, which allows for conformational changes in presence of polyUb chains. In particular, the presence of long polyUb chains induces the expansion of the folded core, with several hydrophobic residues exposed to the

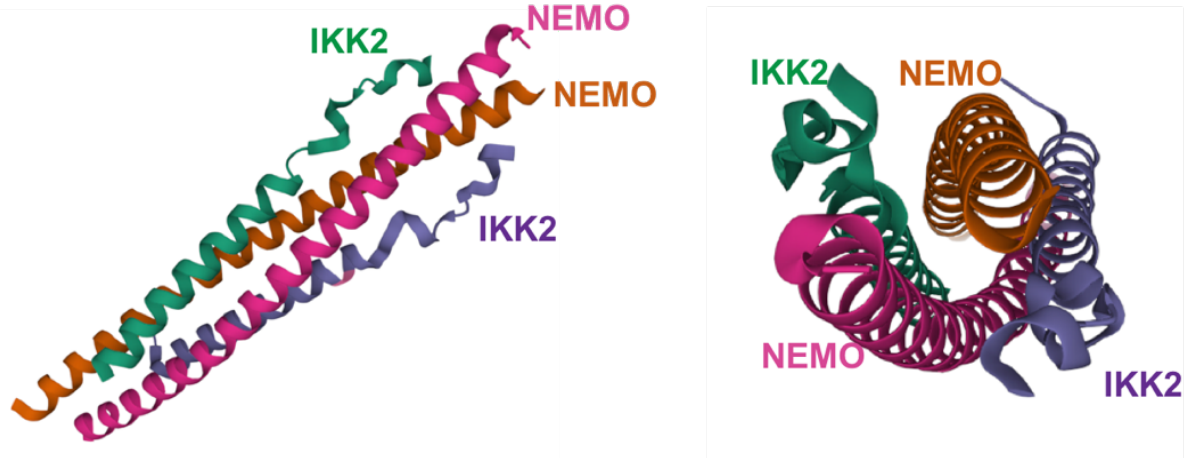



**Figure 11.** Longitudinal and depth view of the NEMO KBD bound to the IKK2 NBD complex. The structure shows that the interaction occurs between dimeric peptides at a stoichiometric ratio of 1:1 (NEMO peptides fuchsia and orange, IKK2 peptides purple and green) (PDB: 3BRT) [167].

solvent, which otherwise would be masked. These residues function as binding sites for different partners such as IκBα and other regulatory molecules [11].

IKK1 and IKK2 are both capable to bind to NEMO *via* their C-terminal NBDs. However, IKK2 has a higher affinity compared to IKK1 [167]. Based on the crystal structure of the IKK/KBD complex (PDB: : 3BRT), the dimeric peptides of the NEMO KBD interact with IKK NBD peptides and form a symmetrical tetrameric α-helix bundle [167] (Figure 11).

In the context of the NF-κB canonical signaling, NEMO plays a fundamental role in addressing the IKK2 kinase activity towards the IκB proteins, in particular the IκBα protein. The work of Hoffmann and collaborators suggests that NEMO directly binds to IκBα, participating in the spatial disposition of IκBα for proper phosphorylation. This work also suggests that the ZF is implicated in the interaction [27].

Given its central role in the regulation of NF-κB signaling *via* protein-protein interactions, NEMO is targeted by several proteins from pathogens. To date there is one x-ray structure of NEMO bound to a viral protein, which is the vFLIP protein from Kaposi's sarcoma-associated herpesvirus (KSHV). The crystal structure of the NEMO/vFLIP complex reveals that the interaction occurs within the HLX2 region of NEMO (Figure 9) [28].

## 2.4 IKK-related kinases TBK1 and IKKE

TANK-binding kinase 1 (TBK1) and IKKE are two IKK-related kinases, which exert the important role of coordinating the NF-κB and the interferon (INF) pathways. Interferons are a group of cytokines implicated in the immune response against viral infections and cancer cells. There are at least three main interferon groups: type 1 (INFα, β, ε and κ), type 2 (INF γ) and type 3 (INF λ).

### 2.4.1 TBK1 and IKKE in the NF-κB and INF pathways

The expression of interferon molecules is the result of the activation of specific PRR receptors, which are sensitive to a wide range of stimuli, such as bacterial lipids, glycans, ssDNA and cytosolic DNA. The interferon regulatory factor family (IRF) comprises 9 nuclear factors (named IRF1 to IRF9) that regulate the expression of type 1 interferon coding genes. The IRF factors can form homo and heterodimers and trigger the antiviral response also by inducing immune cell maturation and differentiation [168]. Normally, dimerization is required for activation and it is achieved by phosphorylation catalyzed by different upstream kinases. The IKK-related kinases TBK1 and IKKE are responsible for the phosphorylation and activation of IRF3 and 7 [169].

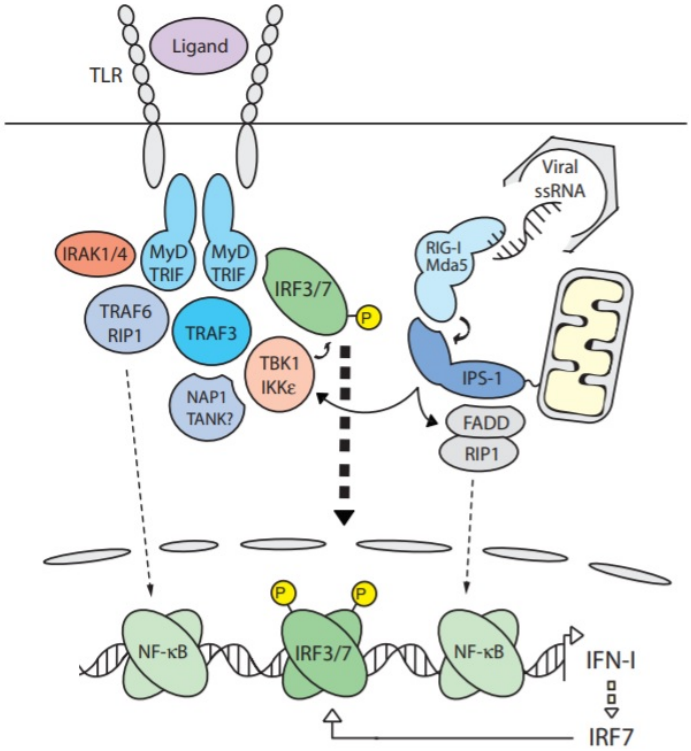



**Figure 12.** The INF pathway requires the cooperation of NF-κB signaling to obtain full activation. Indeed, the majority of stimuli captured by immune system receptors induce the activation of both the signaling pathways, through complex and interconnected signaling mechanisms. TLR activation induces the formation of a signaling platform whose core is represented by MyD88 and TRIF. Several scaffold proteins complete the complex and recruit IKK activating kinase RIP1 starting the NF-kB signaling and the IRF 3/7 factors together with TBK1 or IKKE. Also RIG cytosolic receptorsinduce the activation of both RIP1 and IKK-related kinases in a mitochondrial-LPS-1 dependent manner. (taken from Hacker & Karin, 2006 [169]).

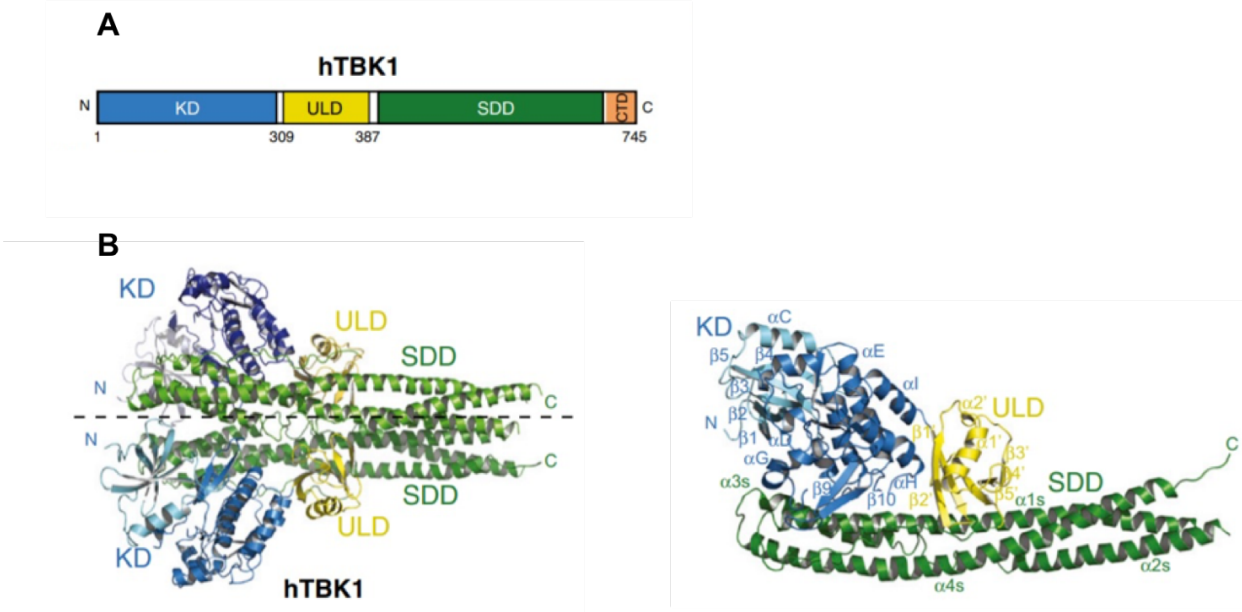



**Figure13. A)** Domain architecture of TBK1. **B)** 3D structure of TBK1 showing the dimer and the isolated monomer subunit with the KD domain displayed in blue, the ULD domain in yellow and the SDD domain in green (taken from Larabi et al., 2013 [16]).

The INF and the NF-κB pathways are deeply connected. For instance, upon TLR activation and dimerization, a signaling platform is assembled at the C-terminal domain by the scaffold proteins MyD88 and TRIF. MyD88 recruits TRAF3/IRF7 [170] or TRAF3/IRF3 [171] complexes . Then, TBK1/IKKE heterodimers bind TRAF3 through scaffold proteins and induce the phosphorylation of IRF factors. The phosphorylated factors dimerize and translocate in the nucleus where they start the expression of INF 1 genes. In parallel, the Myddosome leads to NF-κB signaling activation *via* the IKK kinase complex.

TBK1/IKKE are activated also by cytosolic ssDNA. The viral DNA is detected by intracellular PRRs (pathogen recognition receptors) such as RIG receptors RIG-I or Mda5. These receptors induce the activation of IKK-related kinases in cooperation with mitochondrial LPS-1 in a TRAF3-dependent manner [172] (Figure 12).

### 2.4.2 Structural features of IKKE and TBK1

TBK1 shows 64% homology to IKKE. In contrast to the broad expression of IKK1 and 2, TBK1 and IKKE are exclusive of a restricted number of cell types. In particular, IKKE is preferentially expressed in T cells while TBK1 is constitutively expressed in embryonic fibroblast. However, transient expression is observed upon immune receptor activation [169]. The TBK1 activation is obtained by trans-autophosphorylation at Ser172 in the activation loop [173] or by IKK2 phosphorylation [174]. TBK1 its antiviral activity can be enhanced by K63-linked poly-ubiquitination at several residues.

To date, the structure of IKKE has not been resolved yet. The crystal structures of both active and inactive human TBK1 bound to an inhibitor wer solved by x-ray crystallography at resolutions of 2.8 and 4 Å, respectively (PDB: 4IWQ) [16]. Like IKK1 and IKK2, TBK1 forms homodimers and displays the trimodular domain architecture consisting of the KD, the ULD and the SDD. The three domains are strongly associated and the dimerization occurs through several interaction surfaces. The dimer shows the same IKK scissors-like shape although with a more compact and closer structure (Figure 13).

# Chapter 3

# Physiopathology of NF-κB signaling

Since the IKK kinase is considered the master regulator of NF-κB signaling, here I describe some of the most prominent functional contributions of NF-κB signaling, focusing mainly on its roles in the immune response and apoptosis. In particular, NF-κB signaling is considered a pivotal element of the innate immune system. Its numerous physiological functions concern not only the activation of the immune response but also its modulation through the healing processes and eventually its shutdown. In addition, NF-κB is also required for proper selection and differentiation of B and T cells.

Besides its functions in immunity, NF-κB plays important roles in other areas such as embryonic development and neurometabolism. However, these latter contributions of NF-κB will not be covered in this thesis.

Given the important contributions in several physiological processes, deregulation of NF-κB signaling gives rise to different pathological conditions, such as self-immune diseases, immunodeficiency and cancer.

## 3.1 NF-κB in inflammation and immune response

Inflammation is an innate non-specific defense in response to physical, chemical and biological harmful agents. Its ultimate goal is the elimination of the initial cause of the damage through the destruction, the dilution or the confinement of the damaging element and restoration of tissue functions through the healing process. Inflammation occurs through a dynamic sequence of elementary processes, including initial vasodilatation, increased vascular permeability, which allows the passage of fluids (edema) and leukocyte extravasation in the damaged area. The enormous variety of damaging agents lead to the activation of similar inflammation processes that are determined by the release of a common set of endogenous pro-inflammatory molecules. This means that innate immune cells react to different stimuli with few common responses. The NF-κB pathway plays a crucial role in the modulation of the innate and adaptive immune response as it controls the expression of pro-inflammatory genes and also signal termination genes. Thus, deregulations of the NF-κB pathway leads to the development of inflammatory diseases [175].

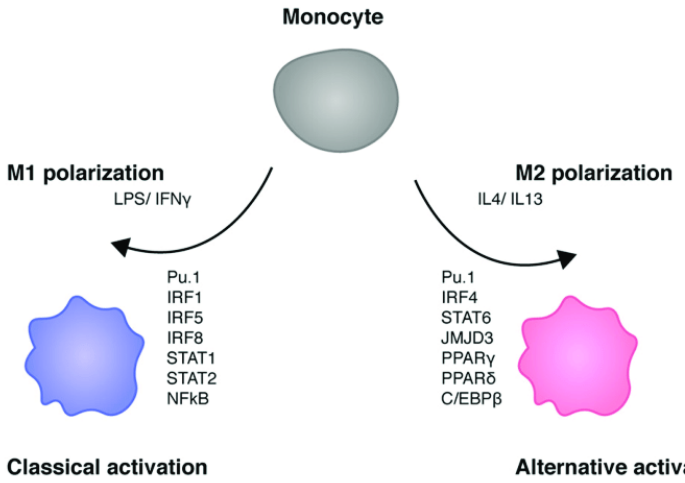



**Figure 1.** Monocyte polarization towards M1 and M2 macrophages. The pro-inflammatory and pro-cancer M1 phenotype is induced by PAMPs and DUMPs such as LPS (classical activation) *via* the activation of several transcription factors such as NF-κB. M2 anti-inflammatory phenotype is reached upon IL4/IL13 (taken from Platanitis and Decker, 2018 [176]).

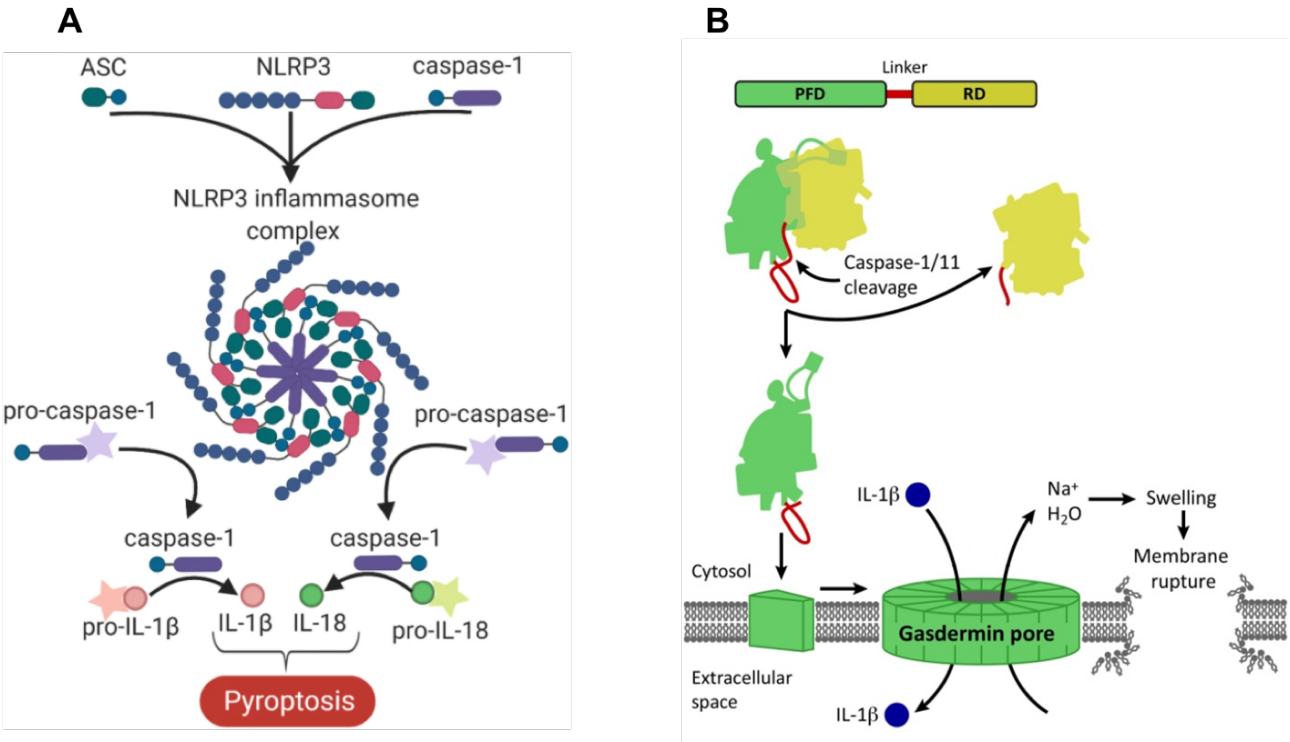



**Figure 2.** A) Inflammasome assembled by NLRP3 and ASC. Pro-caspase 1 is recruited to the inflammasome which results in its activation and subsequent cleavage of pro-IL1b/18, triggering the pyroptotic process in macrophages. (taken from McDonald et al., 2020 [177]). B) Caspase 1-activated gasdermin PFD monomers assemble a pore in the cell membrane that induces cell collapse (adapted from Kovacs and Miao, 2017 [178]).

### 3.1.1 Functions of NF-κB in innate immune cells

NF-κB signaling controls the activation of innate immune cells such as macrophages, dendritic cells and neutrophils. These cells detect several exogenous elements, called pathogen-associated molecular patterns (PAMPs) and damage associated molecular patterns (DAMPs), *via* their primitive pattern recognition receptors (PRRs). Mammals express five subgroups of PRRs: the toll-like receptors (TLRs), the RIG-I-like receptors (RLRs), the NOD-like receptors (NLRs), the C-type lectin-like receptors (CLRs) and the cytosolic DNA sensor.

In macrophages, the interaction of DAMPs and PAMPs with PRR receptors activates the canonical NF-κB pathway, inducing the expression of pro-inflammatory molecules such as cytokines, chemokines and other inflammatory mediators [179]. In turn, pro-inflammatory molecules activate MyD88-dependent TLR signaling that leads to the activation to TAK1 through the Interleukin-1 receptor-associated kinase (IRAK) pathway, which ultimately triggers the NF-κB pathway [180] (see Figure 11 of Chapter 1). The establishment of this positive feedback loop contributes to the polarization of macrophages towards the pro-inflammatory phenotype (M1) (Figure 1).

Depending on the nature and intensity of pro-inflammatory stimuli, native immune cells might embark upon pyroptosis [181], a cell death program which, unlike apoptosis, carries out an intense inflammatory response due to release of pro-inflammatory molecules through the induction of specific membrane pores [182]. Pyroptosis depends on the formation of a cytosolic protein complex named inflammasome (Figure 2A), whose assembly depends on cytosolic PRRs such as NOD-like receptors (NLRs). Activation of the inflammasome in innate immune cells requires two steps. A primary signal, generally represented by TLR ligands, cytokines or other primitive non-self antigens, activates the NF-κB pathway, resulting in the expression of pro-IL-1β, pro-IL-18 and PRRs such as NLRP3. The secondary signal consists of PAMPs, DAMPs, pore-inducing toxins, viral genetic material, damaged mitochondria and lysosomes related molecules such as ATP, mtDNA and ROS [183]. These pro-inflammatory factors induce the oligomerization of NLRP3 with subsequent recruitment and activation of procaspase 1 through the Apoptosis-associated speck-like protein containing a CARD (ASC) and the essential regulatory protein NIMA-related kinase 7 (NEK7). Caspase 1 processes pro-IL-1b and pro-IL-18 to produce their active forms, IL-1β and IL-18, respectively [184]. Caspase 1 also targets the Gasdermin-D which is formed by a pore forming domain (PFD) and an inhibitory regulatory domain (RD). The Caspase 1 cleaves the linker between the domains, producing a N-terminal peptide which starts the pyroptosis through the induction of specific membrane pores (Figure 2B).

NF-κB exerts also anti-inflammatory roles. As a matter of fact, LPS stimulation of IKKs knock-out macrophages induces caspase 1 hyperactivation, leading to augmented IL-1β production [185]. Such

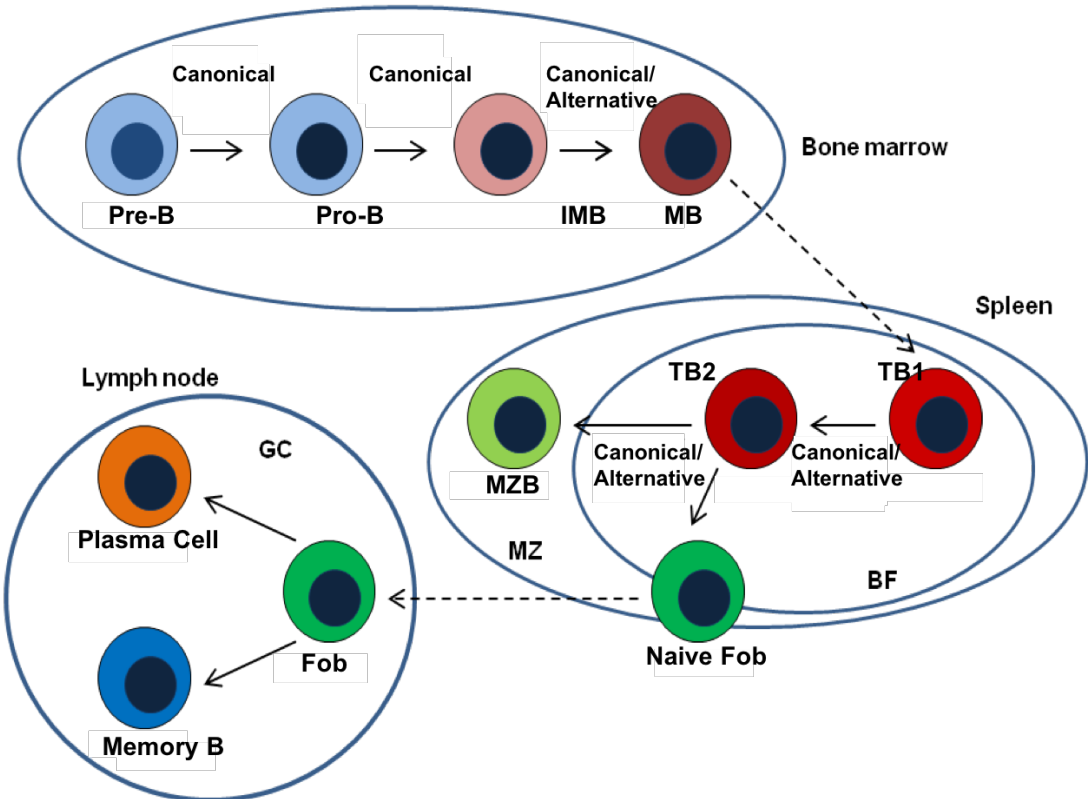



**Figure 3.** B cell maturation process requires NF-κB activation. Pre and pro B cells development is based on the classic NF-κB signaling. IMB and MB cells require both classical and alternative NF-κB signaling for survival, receptor rearrangement and negative selection. Once migrated in the spleen, MB go through the T1/T2 stages under the control of classic and alternative pathways which also induce T2B cell maturation in MZB or FoB cells. Naïve FoB cells reach the lymph nodes where they proliferate creating GC, producing memory B or plasma cells.

anti-inflammatory role of IKK2 is due to the induction of autophagy, through NF-κB dependent expression of the p62 receptor. In fact autophagy allows the cell to get rid of old and damaged organelles, such as mitochondria, thus decreasing the concentration of endogenous pro-inflammatory factors and neutralizing some components of the inflammasome, such as caspases [186].

### 3.1.2 Functions of NF-κB in B cells

NF-κB owes its name after its discovery as a nuclear protein binding to a highly conserved sequence called κB site located at the κ light-chain gene enhancer in B cells. NF-κB signaling is deeply implicated in B cell transitional process and in mature B cell activity. In the bone marrow, the hematopoietic process leads to the formation of pre-B cells, which mature in pro-B cells depending on the activation of the p50/RelA NF-κB heterodimer (Figure 3). After leaving the bone marrow, immature (IMB) and T1-like mature (MB) B cells reach the spleen and go through two stages of transition, T1B and T2B, as a result of which they differentiate in follicular (FoB) or marginal zone (MZB) B cells (Figure 3). The T1-T2 transition is under the control of the alternative NF-κB pathway, which, in this case, is activated by TNF superfamily BAFF ligands binding to BAFFR receptors [187, 188]. However, the canonical pathway seems also required for the transition, since B cells lacking c-Rel and RelA are unable reach stage T2 [189].

BAFFRs are also implicated in survival mature B cells and in the control of the B cell pool size mainly through the activation of NF-κB alternative pathway [190]. The B cell maturation into an antibody-secreting cell upon antigen recognition and selection, is achieved through two distinct processes: thymus dependent (TD) and thymus independent (TI). In the TD process, FoB cells gather in lymph nodes, where they are selected and activated by TCD4+ and follicular dendritic cells, generating a germinal center (Figure 3). Here, B cells proliferate, undergo somatic hypermutation and class switch recombination (CSR), necessary for the antibody affinity maturation, or, alternatively, differentiate into plasma cells. Both B cell proliferation and CSR depend on CD40 signaling, which activates NF-kB signaling leading to the transcription of pro-proliferative genes [191] and of specific $C_H$ genes coding for antibody heavy chains [192].

The TI process regards mainly maturation of MZ B cells, which acquire the ability to produce low affinity IgM antibodies [193]. The TI process is triggered by antigens inducing T helper cell (Th)-independent B cell proliferation. In this case, B cell receptor (BCR) and LPS activated TLR lead to the activation of both canonical and alternative pathways implicated in proper proliferation and accurate LPS-mediated ERK response. TI process does not induce immunological memory [194].

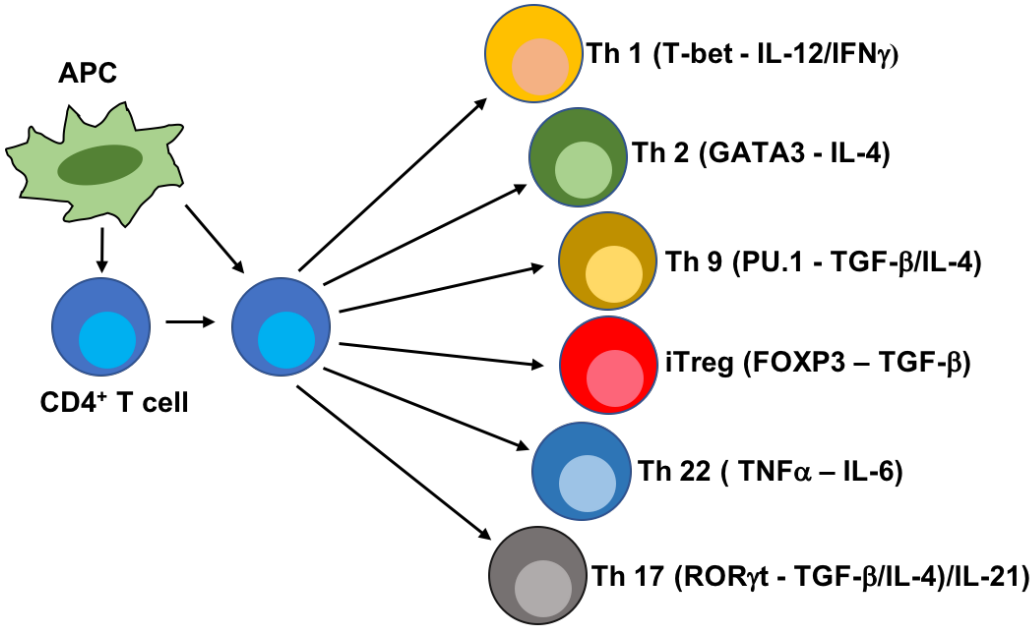



**Figure 4.** CD4+ T cells require NF-κB activation and other specific TFs for differential maturation upon TCR activation. For instance, high pro-inflammatory Th1 and Th17 require NF-κB plus T-bet and RORγt transcriptional activity, respectively (adapted from Hayden and Ghosh, 2011 [179]).

### 3.1.3 Functions of NF-κB in T cells

The NF-κB canonical pathway is directly and indirectly involved in the activation and polarization of T cells, in particular CD4+ T helper cells (Th cells), which belong to the adaptive immune system. The maturation process is based on TCR receptor activation by the antigen-presenting cells, in particular dendritic cells. Th cells can differentiate into pro-inflammatory or immunosuppressive phenotypes based on the composition of the molecular cocktail to which they are exposed [195]. The differentiation of naïve T cells into Th1, Th2, Th9 and Th17 pro-inflammatory phenotypes is regulated at the transcriptional level by NF-κB dimers in cooperation with other transcription factors such T-bet, GATA3, PU1 and RORγt (Figure 4).

NF-κB signaling also promotes naïve T cell polarization indirectly through the expression of pro-inflammatory cytokines, such as IL-2 and IL-12 in dendritic cells [179]. The alternative NF-κB pathway is also implicated in the proper functioning of the T cell pool. In particular, p100 is necessary for the balanced genesis of the Th1 and Th17 phenotypes and for T cell polarization towards the memory phenotype (memory T cell). In addition, the alternative NF-κB pathway is implicated in the activation of Th17 in neuroinflammatory processes [196]. Canonical NF-κB, in contribution with the master TF Foxp3 [197], exerts an immunosuppressive function since it is involved in the generation of the regulatory T cell type (Treg), which controls the intensity of the immune response, thus protecting against autoimmune diseases and inflammation processes chronicization [198].

## 3.2 NF-κB and apoptosis

Apoptosis is a highly regulated cell death process, which promotes the elimination of damaged cells and prevents the inflammation response. This highly controlled process is realized in response to hormone and molecular signaling *via* two pathways, namely the endogenous and exogenous pathways. Apoptosis is deeply implicated in ontogenesis, morphogenesis and postnatal development, in tissue homeostasis and in the elimination of stressed, infected or aberrant T cells.

### 3.2.1 Anti-apoptotic role of NF-κB in the innate immune response

NF-κB plays a role in modulating pro- and anti-apoptotic signals in innate immune cells. NOD1/2 proteins activate NLR receptors, inducing the assembly of the inflammasome and the activation of the NF-κB canonical pathway, with the subsequent expression of pro-inflammatory cytokines and anti-apoptotic genes such as Bcl-2 and Blf-1/1A [199]. On the other hand, the inflammasome inducing proteins NOD1/2 together with CARD and IPAF are also known to trigger cell death programs

pyroptosis and apoptosis through NF-κB-dependent expression of pro-inflammatory genes and pro-caspases. This apparently contradictory role of NF-κB may be explained by the fact that in physiological conditions, the expression of anti-apoptotic genes upon NOD stimulation counterbalances pro-apoptotic signals, generating a checkpoint, which determines the preferential pathway that the cell should undergo based on the nature, the composition and the intensity of the stimuli [200].

In monocytes, the activation of TLR2 by different bacterial peptidoglycans, participates to determine the cell lifespan. In fact, TLR2 induces the NF-κB pathway, which, besides inducing the expression of pro inflammatory cytokines, exerts an anti-apoptotic function against the pro-apoptotic signaling of TLR2-induced Fas associated death domains (FADD) [201].

NF-κB can also protect against apoptosis induced by TNFα. TNFα is a key regulator of immune cell response. Upon TNFR1 receptor activation, two antagonistic signaling pathways are triggered. The first pathway is based on the activation of the IKK canonical complex *via* the TRAFF/RIP1 complex associated to the receptor (sub-membrane) and leads to the expression of anti-apoptotic genes, such as caspase-8 inhibitor FLIP (CASPER), IAPs and Bcl-2 family proteins. The second pathway is triggered by the cytosolic TRAFF/RIP1 complex, which recruits caspase 3. In turn, activated caspase 3 targets IKK2 to proteolysis, thus preventing the expression of anti-apoptotic genes [202]. TNFα also promotes apoptosis by the release of cathepsin B, lysosomal cysteine protease. Cathepsin B is strongly inhibited by Spi2A, which is expressed in a NF-κB-dependent manner [203].

### 3.2.2 Anti-apoptotic roles of NF-κB in the adaptive immune system

NF-κB promotes mature B cells maturation and survival. Upon BCR or BAFFR activation, the NF-κB canonical pathway induces the expression of the anti-apoptotic genes Bcl-xl and Bcl-2. Indeed, knock out of IKK proteins or Rel subunits induces a dramatic increase of the apoptotic ratio [204]. Consistently, lack of anti-apoptotic factors Bcl-2 and Bfl-1/A1 induces premature apoptosis, thus preventing B cells maturation into plasma cells [205].

NF-κB also exerts a protective role against apoptosis during naïve T cells activation. Th cells undergo maturation upon co-stimulation of TCR and CD28, which activate the NF-κB pathway, inducing the expression of anti-apoptotic genes such as Bcl-2, Bcl-xl and Bfl-1/A1 plus the proliferative factors IL-2 and Granulocyte-macrophage colony-stimulating factor (GM-CSF) [206]. Once triggered, T cells need a specific set of pro inflammatory molecules to sustain their survival and activity in the absence of CD28, in a process called adjuvant-induced survival (AIS). This response is correlated with the increase of Bcl-3 nuclear fraction, which is supposed to enhance the expression of anti-apoptotic NF-κB target genes [207].

In naïve T cell, primary stimuli induce the E2F-dependent upregulation of pro-apoptotic p73 and this is antagonized by the simultaneous activation of NF-κB. In certain conditions of stimulation, the NF-κB signaling overwhelms the pro-apoptotic p73 activity, promoting T cell survival and maturation. On the other hand, in absence of NF-κB stimuli, p73 induces apoptosis. In this way, self-reagent T cells are eliminated [208].

### 3.2.3 Crosstalk between NF-κB and tumor suppressors

Inhibition of NF-κB transcriptional activity is a crucial step in the induction of apoptosis. Upon TNFα or mitogenic stimulation, tumor suppressor ARF inhibits NF-κB by inducing the association of RelA NF-κB subunit with histone deacetylase 1 (HDAC1) in a promoter-specific way, downregulating the expression of anti-apoptotic genes. Interestingly, knock out of RelA prevents p53-induced apoptosis [209]. Tumor suppressor p53 promotes instead the association of HDAC1 with the p52 NF-κB subunit, suppressing the expression of Cyclin D1, thus preventing G1-S transition [210].

NF-κB is also known to participate in the transcription of genes such as p53, the death receptor Fas (DR Fas) and its own ligand (FasL), pro-apoptotic Bcl-2 proteins, Bax/Bcl-xS and TNFα [211]. Indeed, TNFα stimulation can promote apoptosis through production of ROS. This occurs through the c-Rel NF-κB subunit, which induces the expression of Manganese Superoxide Dismutase (MnSOD) that converts ROS into hydrogen peroxide $H_2O_2$, thus leading to mitochondrial-dependent apoptosis [212].

## 3.3 Pathologies associated with deregulation of NF-κB signaling

### 3.3.1 Self-immune and inflammatory diseases

Due to the critical role of NF-κB in the regulation of the immune system and the inflammation processes, its deregulation is related to a series of self-immune and inflammatory diseases such as rheumatoid arthritis, inflammatory bowel disease, the multiple sclerosis and atherosclerosis.

*Rheumatoid arthritis.* This is a self-immune disease caused by chronic inflammation and detriment of cartilage and bones. The immune cell infiltration, rich in T and B cells, into the synovium is characterized by a pre-clinical activation of NF-κB, which increases during the development of the disease [213]. NF-κB is responsible for the expression of inflammatory molecules in monocytes/macrophages, which in turn induce the recruitment and the maturation of other immune cells, sustaining and enhancing the inflammation [214]. NF-κB also induces the polarization of monocytes into osteoclasts by the receptor activator of NF-κB (RANK) activation, deregulating the

bone homeostasis and causing bone loss [215]. NF-κB promotes the polarization of CD4 T cells towards the pro-inflammatory T17 phenotype, which is highly related to the genesis of rheumatoid arthritis, by expression of pro-inflammatory cytokines and by direct expression of T cell polarization genes [216]. Also, it promotes survival of self-reactive B cells [217].

*Inflammatory bowel disease.* This group of diseases comprises inflammatory disorders involving the aberrant response of the gastrointestinal immune system to the intestinal microbiome. The chronic inflammation is sustained by the constitutive activation of NF-κB in the mucosal innate immune cells such as the macrophages and neutrophils, which affect the correct functioning of the intestinal epithelium, and that promote the Th1 and Th17 polarization of T cells [218]. On the contrary, in the epithelial cells, NF-κB has a protective role against IBD and is necessary for maintaining the intestinal epithelial and immune homeostasis [219].

*Multiple sclerosis.* This is a neurologic degenerative disease that affects the central nervous system presumably caused by deregulation of T cells, in particular the T1 and T17 lineages as shown in several models of experimental autoimmune encephalomyelitis models (EAE) [220]. IKK-dependent expression of cytokines in T cells induces an EAE-like phenotype, while IKK2 genetic knock-out or pharmacological inhibition reduces the EAE-related symptoms, as well as inhibition of upstream TCR signaling, presumably by preventing T17 polarization [221]. Also p52 in collaboration with c-Rel is found to trigger the expression of EAE-inducing cytokines [196]. Contrariwise, NF-κB is critical for the neural homeostasis as the impairment of NEMO or IKK2 in neurons, astrocytes and oligodendrocites induce neuroinflammation and prevent EAE onset [222].

*Atherosclerosis.* One last example of NF-κB related disease is the atherosclerosis. Indeed, NF-κB is responsible for the expression of genes implicated in several aspects of the pathogenesis and progression of the atherosclerotic lesion [223]. This vascular disorder is characterized by the accumulation against the arterial wall of low-density lipoproteins (LDLs), which induce inflammation of the endothelium by the NF-κB-dependent expression of chemotactic and adhesion factors followed by recruitment of immune cells in the subendothelial space. Monocytes enter and accumulate in the arterial intima and polarize into macrophages, which eventually turn into lipid-laden foam cells, initiating the atherosclerotic plaque formation process [224].

### 3.3.2 SCID

The severe combined immunodeficiency (SCID) is a group of autosomal recessive heritable conditions of humoral and/or cell-mediated immunity deficiencies caused by mutations in genes implicated in the activation and functioning of immune processes. These rare conditions are usually fatal within the first year of life unless they are treated with immune-restoring treatments. Depending on the genetic defect, the immunodeficiency shows different molecular and pathologic features. For example, the X-linked SCID is caused by a mutation in a gene coding for a cytokine receptor sub-unit on the X chromosome, thus affecting only males whose T and NK cell blood count is drastically decreased, preventing the activation of B cells. Another well-known type is the adenosine deaminase (ADA) deficiency, a lethal condition for T cells.

Genetic defects of genes implicated in the NF-κB signaling are also related to different forms of SCID. Hypomorphic mutations in the gene enconing NEMO (*IKBKG*) and IκBα (*NFKBIA*) cause anhidrotic ectodermal dysplasia with immunodeficiency (EDA-ID) or immunodeficiency only [225]. In particular, EDA-ID is caused by short truncations or missense mutations, which abolish the zinc finger or the UBAN motif of NEMO [226].

Other forms of SCID do not impair the number of immune cells but their activation is abolished or largely inhibited [227]. This is the case of a form of SCID with normal count of lymphocytes with a very low immunoglobulin counting and severe defect in immune cell activation, caused by homozigous null mutations in the IKK2 gene (*IKBKB*). This duplication in exon 13 prevents the expression of IKK2, leading to hypogammaglobulinemia or agammaglobulinemia, lack of mature peripheral-blood B cells and T cells and absence of regulatory T cells and γδ T cells [228].

### 3.3.3 Incontinentia pigmenti

This disease is a multisystemic genodermatosis lethal in male fetuses determined by deletions in the IKBKG gene that lead to the expression of a truncated form of NEMO able to dock to IKK but unable to trigger the canonical NF-κB pathway [229].

### 3.3.4 NF-κB-mediated viral pathogenicity

Several viruses evolved to exploit the critical processes regulated by the NF-κB signaling to gain virulence and maximize the chances of infections. Several human pathogenic viruses, such as HIV-1, T cell leukemia virus (HTLV-1), hepatitis B virus (HBV) and different strains of influenza, rely on NF-κB activation to promote cell survival and proliferation, thus increasing the replication potential of the virus. In other cases, some viruses silence the NF-κB signaling to prevent the activation of immune cells increasing their propagation.

NF-κB signaling can be triggered by the presence of viral genetic material or viral proteins in a direct or indirect manner. The accumulation of cytosolic viral dsRNA induces the activation of the dsRNA dependent protein kinase (PKR), which is part of the IKK signalosome and which triggers activation of the IKK complex by direct interaction [230]. An alternative mechanism of NF-κB activation is the one used by several members of the rotavirus *genus*. These viruses use the capsid protein VP4 and its cleavage product VP8 to activate the TRAF2-NIK complex *via* a conserved TRAFF-binding motif, leading to the activation of the NF-κB alternative pathway [231]. Other viruses, such as the influenza virus and the adenovirus, indirectly trigger NF-κB signaling by inducing calcium release from the endoplasmic reticulum (ER)[232].

*HIV*. The HIV-1 virus acts at different steps of NF-κB signaling to sustain the infection. The glycoprotein gp120 binds and activates CD4 signaling, which results in the activation of the T cell-specific tyrK $p56^{lck}$ kinase (p56K) together with Ras/Raf signaling. p56K activates the accessory protein Tat, which increases p50 DNA binding activity by inducing its acetylation *via* the CBP/p300 complex in an IKK-independent manner [233]. In parallel, stimulation of PI3-K signaling by the gp120 HIV-1 protein leads to the activation of Akt and, subsequently, of the IKK complex [233]. Contrariwise, the accessory protein Vpu prevents the expression of NF-κB-regulated pro-apoptotic factors in infected cells by inhibition of the β-TrCP-dependent degradation of IκBα [234].

*HBV and HCV*. The transcriptional activator HBx of HBV induces NF-κB signaling by Ras and PKC dependent activation of MEKK-1 signaling [235]. Specifically, HBx exploits the IκBα nuclear shuttling to enter the nucleus where it sustains the expression of NF-κB genes. Interestingly, the HBV protein ZEBRA inhibits NF-κB activation in T cells, increasing their sensitivity to apoptosis and contributing to decrease the antiviral response [236].

Hepatitis virus C (HCV) is related to hepatocarcinogenesis caused by chronic inflammation. In the case of HCV, the persistent NF-κB activation is promoted by the HCV core protein and the unstructured NS5A *via* direct binding to the TNFR1 death domain [237] and ER calcium-release [238], respectively. On the other hand, HCV increases the apoptotic resistance of infected hepatocytes by impeding NEMO ubiquitination. The viral protease NS3 also binds and inhibits HOIP ubiquitin ligase, preventing the IKK canonical complex activation, thus inhibiting NF-κB transcriptional activity [239].

A number of viruses, including Kaposi's sarcoma-associated herpesvirus (KSHV) and HTLV-1, modulates NF-κB signaling by interacting with NEMO. Chronic infection by KSHV causes Kaposi’s

sarcoma and effusion lymphoma. The KSHV protein vFLIP mimics the cellular cFLIP protein (with a sequence homology of 34% between the two proteins), an interaction partner of NEMO that is required for the correct activation of NF-κB. In B cells, binding of vFLIP to NEMO maintains the latter functional, keeping the NF-κB pathway constitutively active and thus increasing cell survival and resistance to apoptosis [240]. The crystal structure of the complex NEMO/vFLIP reveals that the interaction occurs at the CC1 level [28].

*HTLV-1*. The trans-activator protein Tax of HTLV-1 also binds to NEMO, inducing IKK2 activates the IKK activation by transfosphorylation. Also, Tax recruits the IKK1 subunit to p100, favoring its phosphorylation and inducing its processing to the active form p52 [241]. Ultimately, Tax and NEMO form a stable ternary complex with the ser/thr phosphatase PP2A, preventing IKK dephosphorylation thus maintaining the complex activity [242]. Besides KSHV and HTLV-1, several other pathogens target NEMO to deregulate the immune system. These include the foot-and-mouth disease virus and the molluscum contagious virus [28].

A smaller group of viruses evolved specific mechanisms of inhibition of the NF-κB signaling in order to decrease the infection resistance of the host organism. For instance, the vaccinia virus A52R protein is a dominant negative homolog of MyD88 which prevents the expression of IL-1 [243]. The African swine virus has an IκBα homolog (named A238L) that prevents nuclear translocation of NF-κB proteins [244]. Other strains of vaccinia virus express proteins that interfere the degradation of IκB proteins [245].

*SARS-CoV-2*. The SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) is a *Betacoronavirus* responsible for the COVID-19 syndrome. From a clinical perspective, the COVID-19 syndrome is characterized by a widespread inflammation, which targets several tissues and organs [246]. The molecular features of the syndrome show a typical late INF-I response and a quick and massive cytokine expression induced by aberrant activation of NF-κB in infected cells and which determines the most dangerous clinical traits of the syndrome. In epithelial lung cells, the virus replication relies on NF-κB activation which is presumably induced by the accumulation of dsRNA, misfolded proteins and viral products [22] detected by MDA5 and RIG-I [247]. The INF-I signaling is inhibited by the viral main protein NSP5 which prevents the nuclear translocation of IRF3 by direct binding [248]. Another feature of COVID-19 is its ability to cause brain damage with neurologic symptoms [249]. The presence of the virus in the brain might be explained by the damage of the blood-brain barrier caused by the infection of endothelial cells (endothelitis), a typical damage found in other organs [23]. The brain entothelitis seems to be caused by a clivage of NEMO carried out by the

viral protease Nsp5 which is responsible for the elaboration of the viral polyproteins. Nsp5 cleaves NEMO in different spots completely abolishing NEMO signaling and resulting in microvascular brain pathology [250]. In this case, inhibition of canonical NF-κB signaling results in the increased rate of cell death, which causes the permeabilization of the BBB and allows the entrance of the virus in the brain [251].

### 3.3.5 NF-κB and cancer

In physiological conditions, NF-κB self-regulates its activity through negative feedbacks induced by the expression of IκB proteins and other factors that inhibit the activation of the IKK kinase, such as A20 [19]. Deregulation of NF-κB is associated with Th cell transformation and tumor growth. In aberrant T cells, NF-κB target genes expression is generally altered, thus enhancing survival and proliferative responses. Prolonged activation of NF-κB can increase cell pro-tumor potential, induces proliferation through the expression of proto-oncogens and cyclins, promotes the formation of metastases by altering the expression of metalloproteases and cell adhesion genes, stimulates the angiogenesis and shifts metabolism towards glycolysis (Warburg effect) in response to hypoxia [20]. Furthermore, considering the close correlation between inflammation and cancer, NF-κB can induce tumor transformation *via* the activation of immune cells, lymphocytes and macrophages [21].

The main mechanisms through which the deregulation of NF-κB occurs are by mutations that alter the signaling cascade or indirectly by chronic exposure of the cell to stimuli that activate the pathway, such as cytokines, pro-inflammatory molecules and also chemotherapeutic drugs [252].

Mutation-dependent deregulation affects the pathway's regulatory factors. Gain of function mutations related to NF-κB hyper-activation have been found in the proto-oncogenes HRAS, BCR-ABL1 and NIK. Instead, loss of function mutations in negative regulatory factors such as CYLD and A20 prevent the inhibition or even the termination of the signal once the stimuli have vanished [253]. For example, a recurrent chromosomal translocation in lymphomas causes the over-expression of Bcl-10 and MALT. This event leads to the constitutive activation of the Carma1-Bcl-10-MALT complex followed by the hyper-activation of the IKK complex [254]. A Myd88 mutation promotes TLR signaling even in the absence of ligands, which leads to STAT3 NF-κB-dependent over-expression, thus inducing the production of IL-6, IL-10 and interferon by further support of NF-κB signaling through a positive feedback [255].

On the other hand, mutations in the NF-κB and protein IκB subunits are very rare. These mutations may increase in advanced malignant tumors, where resistance to chemotherapy drugs is observed. For instance, inactivation or loss of IκB proteins can be frequently observed in the Hodgkin

lymphoma [256]. The apparent irrelevance of NF-κB gain of function mutations in tumorigenesis can be explained by the activation mechanism of the subunits. The activity of NF-κB is in fact controlled by several post-transcriptional modifications (PTMs), such as phosphorylation, methylation and acetylation [257]. This means that point mutations mimicking a PTM activation, generally have little impact on the functioning of the total NF-κB population. Furthermore, mutations of this type have poor efficiency in mimicking NF-κB activating PTMs [142].

NEMO is thought to be actively implicated in the development and in the progression of cancer. For instance, in renal cancer, NEMO binds to and stabilizes the hypoxia-inducible factor α (HIF α) 1 and 2 which is considered the master regulator of cellular response to hypoxia. The HIF signaling pathway induces angiogenesis, resistance to apoptosis [258]. In the hepatocellular carcinoma, the loss of NEMO induces resistance to apoptosis in hepatocytes by inhibition of RIP1 kinase activity which prevents the triggering of the NF-κB signaling [259].

The oncogenic activity of NF-κB is also achieved through the cross-talk with other pathways, such as tumor suppressor p53 signaling pathways. Indeed, the relationship between the NF-κB and p53 pathways is complex and regulated at multiple levels [260]. This cross-talk is both competitive and collaborative, as revealed by the existence of promoters recognized by both transcription factors and the presence of p53 dependent genes, whose expression also indirectly depends on NF-κB. The subunits of NF-κB and p53 can interact directly, thus acting as reciprocal co-factors in transcription [261]. Under physiological conditions, p53 inhibits the IKK complex and therefore the activation of NF-κB. The reduced expression of NF-κB target genes, such as Glut3, suppresses glycolysis, promoting oxidative phosphorylation. By contrast, knockout of p53 causes the reduction of mitochondrial metabolism in favor of glycolysis. In this condition, RelA relocates to the mitochondrion where it inhibits gene expression and oxidative phosphorylation. Mitochondrial localization of RelA depends on the phosphorylation of Thr505 of RelA, which creates a binding site for mortaline, a mitochondrial HSP70 chaperone, which competes with p53 to bind to RelA. As a consequence, glycolysis is boosted increasing cell resistance to energetic and hypoxic stress [20]. Thus, p53 is thought to inhibit the IKK-NF-κB pathway directly by p53 target gene products or indirectly by the suppression of glycolysis [262]. However, IKK2 also phosphorylates the C-terminus of p53, inducing its polyubiquitination and consequent degradation [263]. Furthermore, upon stimulation of TNFα, c-Rel moves to the nucleus where it binds to ΔNp63 (a splicing variant of p63, belonging to the p53 family) and induces the translocation of the pro-apoptotic p73 into the cytoplasm, which prevents the expression of cell-cycle arrest and pro-apoptotic genes [264].

NF-κB gene expression can also be deregulated by oncogenic p53 mutants. These mutants gain affinity towards NF-κB transcription factors and, through these interactions, are recruited at the promoter of the NF-κB-dependent genes, increasing NF-κB transactivation and expression of pro-tumor genes [265].

At the level of the tumor microenvironment, the activation of NF-κB promotes the expression of pro-inflammatory cytokines and induces macrophage polarization towards the pro-tumoral phenotype M2 [266]. In fibroblasts, NF-κB promotes tumor progression through the expression of molecules for macrophage recruitment and gene implicated in angiogenesis and extracellular matrix remodeling. Furthermore, NF-κB induces the expression of transcription factors such as twist and snail, implicated in the epithelial-mesenchymal transition and therefore favors the metastatic processes [267].

### 3.3.6 NF-κB in neurological diseases

Given the role important of the NF-κB signaling in the neurodevelopment and neuroprotection, its deregulation is related to a series of heterogenous pathologies targeting the central nervous system.

*Ischemic stroke.* The ischemic stroke is caused by a lack of blood perfusion following a hemorrhage or embolic blockage and which results in irreversible tissue impairment. In this scenario, NF-κB dimers play two different roles in neuroprotection based on their compositions and on the pattern of the epigenetic code. In the brain stroke, NF-κB dimers containing RelA induce the transcription of a pro apoptotic set of genes [268] , while c-Rel dimers enhance apoptotic resistance through the expression of neuroprotective genes. Also, c-Rel dimers induce the expression of enzymes such as the super oxide dismutase (SOD) and the mitochondrial uncoupling proteins 4 (UCP4), which buffer ROS and NOS and Bcl-xL, increasing the resistance to apoptosis [269].

*Neurodegenerative diseases.* NF-κB deregulation is implicated also in neurodegenerative processes and degenerative cognitive diseases such as Parkinson or Alzheimer's disease, which are caused by impaired synaptic function of specific circuits in the central nervous system. The molecular alterations that cause synaptic dysfunction are often accompanied by inflammatory phenomena [270]. It has been observed that lack of c-Rel is related to degeneration of nigral dopaminergic synapses in mice, with the development of a neuropathological syndrome attributable to Parkinson's disease. RelA over expression is associated with senescence-like degeneration in progeroid mice models, while the single knock out delays the central nervous system aging process and mitigates degenerative processes [271]. Activation of RelA is strongly based on the acetylation pattern. In mice model of brain ischemia, RelA subunits, which display a low general level of acetylation and a K310 site-specific

acetylation, express pro-apoptotic genes. Interestingly, in c-Rel knock down mice, RelA acetylation pattern is similar to the one observed in the brain stroke model, inducing the expression of pro-apoptotic genes [272]. Such evidence suggests the existence of a balance between NF-κB subunits, which define a threshold for physiological aging or premature and/or pathological aging. Defects of c-Rel along with over activation of RelA increases the susceptibility to early aging processes [273].

*Depression.* The depressive disorder is a group psychiatric pathologies caused by alterations in neurotransmitter metabolism in the prefrontal cortex and hippocampus [274]. Rats exposed to stress situations, which induced depressive-like symptoms, show decreased levels of NF-κB in the prefrontal cortex and in the hippocampus. Administration of repetitive serotonin reuptake inhibitors recover the physiologic levels of NF-κB [275].

The brain development neurotrophic factor (BDNF) is implicated in a wide range of brain functions, from neurogenesis to neural protection and regeneration. In neurons, upon glutamate stimulation, BDNF is expressed *via* NF-κB, increasing the resistance to oxidative stress and convulsant stimulation. This process is mediated by the X-linked inhibitor of apoptosis protein (XIAP). In astrocyte cultures, TNF exposure causes upregulation of BDNF in an NF-κB dependent manner, involving the MAPK pathway [276]. In rats, BDNF induces NF-κB activation, which results in the expression of Bcl-xl increasing resistance to apoptosis [277].

In the brain, lack of NF-κB signaling generates alterations in neural morphogenesis and development accompanied by cognitive loss [278]. It has been proposed that NF-κB and BDNF could be pharmacological targets in conjunction with conventional therapies for the treatment of depression. For instance, acetyl-L-Carnitine, a protein acetyl group donor, is used in the treatment of depression in older patients. This molecule causes the hyper-acetylation of RelA, which in turn increases the expression and the synthesis of the metabotrobic glutamate receptor 2 (mGlu2) and induces the differentiation of hippocampal progenitor cells ameliorating depression-like symptoms [279].

## 3.4 Inhibitors of IKK

### 3.4.1 The NF-κB as a promising pharmacological target

Given its major role in the activation and hyper-activation of the NF-κB signaling, the IKK complex represents a promising therapeutic target in the treatment of inflammatory diseases [280]. To date, several small molecules have been identified as IKK enzymatic inhibitors [281]. Due to the high level of similarity between IKK1 and IKK2, the inhibitory effect is exerted on both the subunits, generally with a higher affinity for IKK2 [153]. However, IKK1 is slightly more flexible in several regions around

the ATP binding pocket, in particular at the G-loop, and this difference could be expoited to design selective drugs [282]. So far, based on their mechanism of action, at least four categories of IKK inhibitors have been identified.

*ATP analogues.* ATP mimetics represent the broadest category of reversible kinase inhibitors. These molecules compete with ATP for the binding site. Due to the high level of conservation of the kinase domain, ATP analogues are poorly selective molecules. In fact, these molecules indiscriminately inhibit numerous kinases, producing off-target effects even at the concentrations required to obtain the inhibition of the main target [283]. It often happens that these molecules act as inhibitors through alternative or indirect mechanisms [284]. To date, the most promising molecules are MLN-120B and BI605906. These molecules exhibit a high selectivity for IKK1 and IKK2, compared to numerous other kinases [285]. Furthermore, BI605906 has a 300 times higher affinity for IKK2 than IKK1 [174].

*Allosteric inhibitors.* Some molecules inhibit IKK in an allosteric way. These molecules bind to the enzyme at regions other than the active site. The changes induced by this interaction do not necessarily prevent the binding with ATP or with the substrate, but rather alter the dynamic features of the kinase domain by reducing or blocking the enzymatic activity. The compound BSM-345541 acts by inhibiting IKK1 and IKK2 through a non-competitive inhibition mechanism [286].

*Thiol-reactive compounds.* Thiol reactive compounds are a class of molecules having the general formula R-SH. The thiol group oxidizes to form a disulfide bond with the thiol group of a cysteine. Some thiols such as berberine, nimbolide and withaferin A generate a disulfide bridge with a specific cysteine in the activation loop of the kinase, preventing the phosphorylation of the flanking serines and therefore its activation [287]. A second mechanism proposed for these compounds is the formation of a disulfide bond between Cys179 of IKK1 and IKK2 monomers that prevent the interaction with NEMO [288]. The compound ainsliadimer A forms a covalent bond with Cys 46 of both IKK1 and IKK2, which alters the ATP binding pocket [289].

*Indirect inhibition.* Other molecules inhibit IKK through indirect mechanisms. Non-steroidal anti-inflammatory drugs (NSAIDs) reduce inflammation acting primarily by inhibiting cyclooxygenases 1 and 2 (COX1 and COX2). Furthermore, it appears that they may also inhibit IκBα phosphorylation [290]. However, chronic administration of NSAIDs induces the opposite effect, stimulating NF-κB signaling [291]. A promising strategy to inhibit the NF-κB canonical complex by

preventing IKK activation is represented by NEMO-binding domain cell-permeable peptides, which mimic the NBD of IKK2. These peptides compete with NEMO for the NBD of IKK1 and IKK2, thus interrupting signal transduction. This mechanism limits off-target effects due to the high specificity of these peptides and reduces secondary effects as they inhibit the signaling only upon stimulation, without altering the basal activity of the kinase [292].

### 3.4.2 IKK inhibitors clinical trials

The parameters evaluated in a clinical trial are different and include, first of all, the pharmacodynamic and pharmacokinetic features, the biosafety profile and the therapeutic efficacy. To date, despite numerous pre-clinical models have highlighted the therapeutic potential of IKK inhibition, nearly all IKK inhibitory compounds tested so far failed in early clinical trials. A first difficulty lies in the correct selection process of patient cohort for clinical trials. Inflammatory processes are complex and involve several tissue and cell specific signaling pathways. Therefore, it becomes complicated to select those patients with a pathology or clinical condition specifically linked to alterations in NF-κB signaling [293]. Other factors to be considered clinically are the therapeutic dose and the methods of administration. Indeed, many promising molecules fail in the clinical phase due to unfavorable pharmacodynamic or pharmacokinetic features [294].

The main concern related to the biosecurity of these molecules lies in the physiological and pathological features of the NF-κB signaling. The low selectivity for IKK1 or IKK2 with respect to other kinases, may alter other pathways and therefore induce potentially harmful side and secondary effects [295]. Moreover, NF-κB signaling controls the innate immune response and the inflammatory processes in complex ways that vary according to the cell type, the tissue and the pathophysiological conditions considered. In principle, any genetic or environmental alteration that leads to the hyper-activation of NF-κB or its constitutive activation, may lead to the genesis of chronic inflammatory diseases and cancer [296]. Therefore, the targeted inhibition of NF-κB reduces inflammation and the risk of tumorigenesis [297].

However, in some cases, NF-κB inhibition leads to opposite results. Indeed, NF-κB signaling also plays a fundamental role in the control of inflammation through a negative feedback that involve other inflammatory pathways [208]. For example, NF-κB negatively controls the inflammasome activation in macrophages by inducing the elimination of damaged mitochondria, through the production of p62 [298]. In fact, IKK2 chronic inhibition or constitutive ablation induces the production of IL-1β, in neutrophils proliferation and life-span increment and eventually severe tissue damage [299]. This mechanism has a clinical relevance as cancer patients treated with an IKK2 inhibitor developed neutrophilia and enhanced inflammation [300]. Finally, in some tissues NF-κB exerts a major

anti-cancer function. For example, in the intestinal epithelial tissue the IKK2-dependent activation of NF-κB controls cell growth and proliferation by expression of Smad7 and Smurf1 which, in turn, downregulate the transforming growth factor beta (TGF-β) signaling. Constitutive inhibition of IKK2 in cancer associated fibroblasts (CAFs) leads to Smad7/Smurf1 down-expression, TGFβ signaling activation and production of hepatocyte growth factor (HGF), accumulation of T cells and boosting of pro-tumor behaviors in epithelial cells [301].

# PART 2

# METHODS

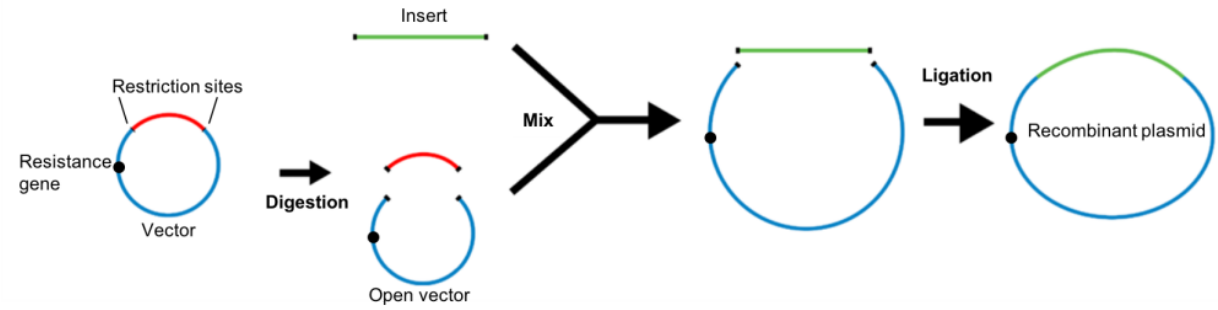



**Figure 1.** Schematic representation of the restriction enzyme cloning. The vector (blue) is linearized by digestion with two different endonucleases. Then, the vector is purified and ligated with the insert (PCR or oligonucleotide annealing product, green), further purified and ready for transformation (adapted from Plasmids 101: Restriction Cloning, Tyler Ford, Addgene).

# Chapter 4
# Materials and methods

In this chapter I review the principles of the methods that I used for my PhD research projects that are described in this dissertation. The precise protocols for the specific experiments are also provided.

## 4.1 DNA cloning

### 4.1.1 Classical restriction cloning

*Principles of classical restriction cloning.* Classical cloning comprises a series of experimental methods used to assemble recombinant DNA (rDNA) by using restriction endonucleases that cut the DNA in correspondence to specific sequences (typically palindromic sequences), creating blunt or sticky ends, which can be joined together (Figure 1).
Classical cloning is carried out in two steps. In the first *in vitro* step the self-replicating DNA plasmid (named vector) is linearized by two specific endonucleases that create unique and non-compatible sticky ends, which are necessary to force the cloning in a directional way in order to maintain the open reading frame of the DNA fragment (a protein or peptide encoding sequence named insert). The vector carries a resistance gene for a specific antibiotic that allows for selection later. In parallel, the insert is prepared by annealing of two synthetic single strand complementary oligonucleotides or by PCR amplification of a specific target sequence followed by DNA purification on agarose gel. Vector and insert are then ligated together using the T4 DNA ligase enzyme.
In the second *in vivo* step the rDNA is amplified by transformation into a living cell (different strains of *E. coli* optimized for DNA amplification), which is able to replicate the rDNA together with its own genetic material. The rDNA insertion in the host organism is obtained by different transformation methods, such as heat shock or electroporation methods. The transformed host is spread on agar plates supplemented with an antibiotic according to the resistance gene encoded by the vector. Colonies grown in the selective medium are collected and amplified in liquid medium supplemented with the same antibiotic. Once the culture reaches saturation, the rDNA is extracted and purified by commercial kits.

*6xHis-IκBα and 6xHis-NEMO.* The DNA sequences of full-length IκBα and NEMO proteins were cloned into the pETM-10 vector (kindly provided by Gunter Stier, Center for Biochemistry,

Heidelberg) that allows for *E. coli* expression of constructs fused to an N-terminal 6xHis tag. The IκBα and NEMO inserts were amplified by PCR and purified on a 1.5% agarose gel followed by extraction using the Nucleospin Gel and PCR cleau-up (Macherey-Nagel). The purified NEMO insert was digested by the NcoI and BamHI enzymes. Since the IκBα sequence contains a NcoI restriction site, the IκBα insert was digested by the restriction endonuclease type II enzyme BbsI (ThermoFisher), which cuts at the +N2 and +N6 nucleotides after the recognition sequence (introduced in the N-terminal PCR oligonucleotide), and by BamHI. The digested fragments were further purified using the same kit and then ligated into the open pETM-10 vector using T4 DNA ligase (ThermoFisher) at 16 °C overnight. Ligation reactions were transformed into *E. coli DH5α* cells and spread out on Luria broth (LB) agar plates supplemented with kanamycin antibiotic. Single colonies were isolated and grown in liquid LB medium for miniprep cultures.

*MBP-IκBα constructs and MBP-peptides.* DNA sequences of IκBα constructs and YDDΦXΦ peptides (from IκBα, IκBβ, p100 and IRF7 proteins) were cloned into the pETM-40 vector (kindly provided by Gunter Stier, Center for Biochemistry, Heidelberg) that allows for *E. coli* expression of constructs fused to an N-terminal MBP (maltose binding protein) tag followed by a Tobacco Etch Virus Protease (TEV) cleavage site. Larger IκBα inserts were amplified by PCR and inserted into the NcoI and BamHI sites of the pETM-40 vector as described above for the 6xHis-IκBα construct. Peptide inserts consisted instead of hybridized DNA oligonucleotides that included NcoI and BamHI recognition sequences at the 5' and 3' ends. Complementary oligonucleotides were phosphorylated by the T4 polynucleotide kinase (ThermoFisher), annealed by cooling down from 95 °C to 25 °C in a water bath and subsequently ligated into the pETM-40 vector using the T4 ligase. Ligation reactions were transformed into *E. coli* DH5α cells and spread out on LB agar plates supplemented with kanamycin antibiotic.

### 4.1.2 Gateway cloning

*Principles of Gateway cloning.* Gateway cloning, originally developed by Invitrogen, is a cloning method that reproduces the integration of the phage lambda λ DNA into the bacterial chromosome. The phage has an attachment sequence (attP), which anneals to the bacterial attachment sequence (attB) and induces its own insertion. The new generated recombinant flanking sequences are named attL and attR (where L and R stay for left and right, respectively).
Gateway cloning is carried out in two *in vitro* steps: the BP and the LR reactions. In the BP reaction, the insert is designed so that it is flanked by attB sequences and it is inserted into an intermediate

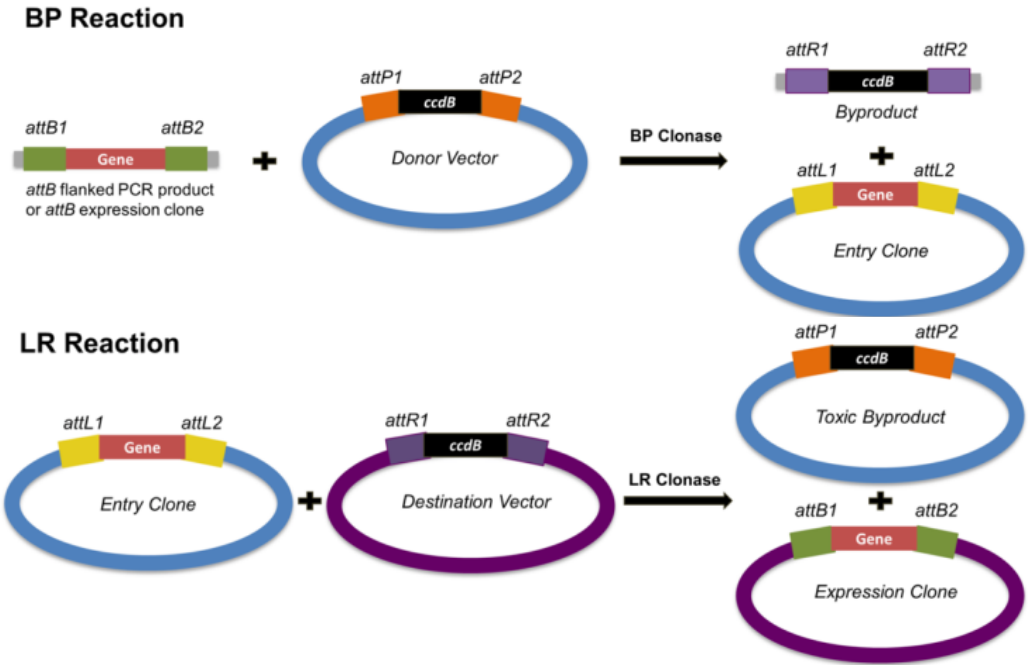



**Figure 2.** Schematic representation of BR and LR reactions of the Gateway cloning (taken from Plasmids 101: Restriction Cloning, Tyler Ford, Addgene).

vector (donor vector) by recombination with attP sequences thanks to the BP clonase enzyme mix, which contains the integrase and the integration host factor. The new entry clone contains the insert flanked by two attL sites. After the plasmid amplification, the insert can be transferred into any other expression clone containing attR sites thanks to the LR reaction, which takes place when the entry clone attL site recombines with the attR site of the destination vector (expression clone) catalyzed by the integrase, the integration host factor and the excisionase enzymes contained in the LR clonase II mix (Figure 2).

Both the donor and the destination vectors contain the *ccdB* gene, which encodes for a gyrase inhibitor, lethal for replicant cells and which is used to prevent the proliferation of bacteria retaining a vector with no insert. The BP and LR reactions cause the excision of the byproduct containing the *ccdB* lethal gene.

*Gluc1(G1)- and Gluc2(G2)-fusion constructs.* Full-length IKK1 and IKK2 proteins, and NEMO and IκBα constructs to be used for the GPCA experiments were cloned by the Gateway approach. Inserts were amplified by PCR using N- and C terminal oligonucleotides comprising the attB recombination sequences. The inserts were purified by the Gel or PCR cleanup kits (Macherey-Nagel). The PCR products were inserted in the pDONR207 vector by the BP reaction, in which 150 ng of attB-PCR product were mixed with 150 ng of Donor vector, brought to a final volume of 8 µl and supplemented with 2 µl of BP Clonase mix (Invitrogen). After 1 hour at 25 °C the reactions were terminated by addition of 1 µl of Proteinase K and incubated at 37 °C for 10 minutes. Inserts were transformed into *E. coli* DH5α cells grown on LB agar plates supplemented with gentamycin antibiotic. The resulting pEntry clones were then transferred into the GPCA destination vectors pSPICA-N1 and pSPICA-N2 (both derived from the pCiNeo mammalian expression vector) by the LR reaction following the same procedure, using LR Clonase mix (Invitrogen). The GPCA vectors enable expression of the Gluc1 (or G1, pSPICA-N1) and Gluc2 (or G2, pSPICA-N2) complementary fragments of the *Gaussia princeps* luciferase, which are linked to the N-terminal ends of the tested proteins by a flexible 20 amino acid linker. The resulting expression clones were selected in *E. coli* DH5α cells grown on LB agar plates supplemented with ampicillin antibiotic.

*6xHis-IKK1 and 6xHis-IKK2 constructs.* The IKK1(10-667) EE and IKK2(1-669) EE constructs were cloned in the pBacPAK8 transfer vector for baculovirus expression. This vector allows for expression of an N-terminal 6xHis tag. In addition, a TEV cleavage site was introduced downstream of the Gateway attB1 sequence, thus allowing for removal of the 6xHis tag and Gateway

sequence from the IKK1 and IKK2 recombinant proteins. Cloning of these constructs was done by another member of the team before my arrival.

### 4.1.3 Site-directed Mutagenesis

Site-directed mutagenesis allows for introduction of residue-specific modifications in virtually every double-stranded DNA (dsDNA) sequence. In our case, we used the QuickChange II site-directed mutagenesis kit (Agilent) to produce site-specific mutants of IκBα (i.e. 6xHis-IκBα and G2-IκBα mutants) and of IKK1/2 (i.e. 6xHis-IKK1(1-669) EE and 6xHis-IKK2(10-667) EE constructs and G1-IKK1 mutants). The desired substitution is introduced at the center of the primers, which anneal on the same complementary sequences. Subsequently, we used high-fidelity DNA polymerase to perform rolling cycle PCR, which results in the amplification of the plasmid containing the substitution. Subsequently, the parental template is eliminated by digestion with Dpn I (1 ml for each amplification reaction, incubated at 37 °C for 1 hour), an endonuclease that targets methylated and hemi-methylated dsDNA (the supercoiled parental dsDNA). Finally, PCR products are transformed into DH5α cells, which are selected on LB agar plates containing the appropriate antibiotic.

## 4.2 Protein expression

Full-length proteins, truncated constructs and peptides were produced by two different heterologous expression systems, namely *E. coli* BL21(DE3) and insect cells using the baculovirus expression vector system (BEVS). For all MBP- and 6xHis-tag constructs expressed in *E. coli,* we used pETM vectors, which contain a resistance gene against kanamycin. This antibiotic is an aminoglycoside that thanks to its resistance to bacterial β-lactamases, enzymes that hydrolyze the β -lactam ring of antibiotics such as penicillins and cephalosporins, maintains the concentration in the culture thus allowing long incubations overnight.

### 4.2.1 Protein expression in *E. coli*

*Principles of E. coli expression.* Protein expression in *E. coli* is the most popular way to produce recombinant proteins thanks to its well-characterized genetic background, the high culture density reachable, the wide number of optimized cloning vectors and strains and also low costs and short times. The bacteriophage T7 RNA polymerase (pol) expression system is one of the numerous *E. coli*-based expression systems available. *E. coli* strains such as BL21DE3 are optimized for such method as they incorporate the T7-gene under the control of the *lac* operon in the bacterial chromosome. The expression vectors contain the gene under the T7 promoter so that the addition of

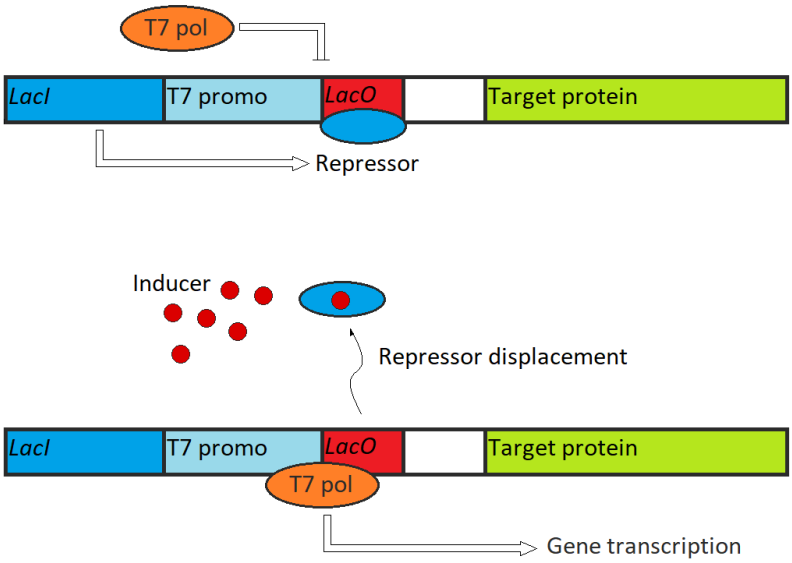



**Figure 3.** Schematic representation of the T7*lac* operon. The repressor binds to the Lac operator (LacO) preventing the T7 RNA polymerase initiation. Addition of inducer displaces the repressor, allowing the gene transcription of the target protein coding sequence.

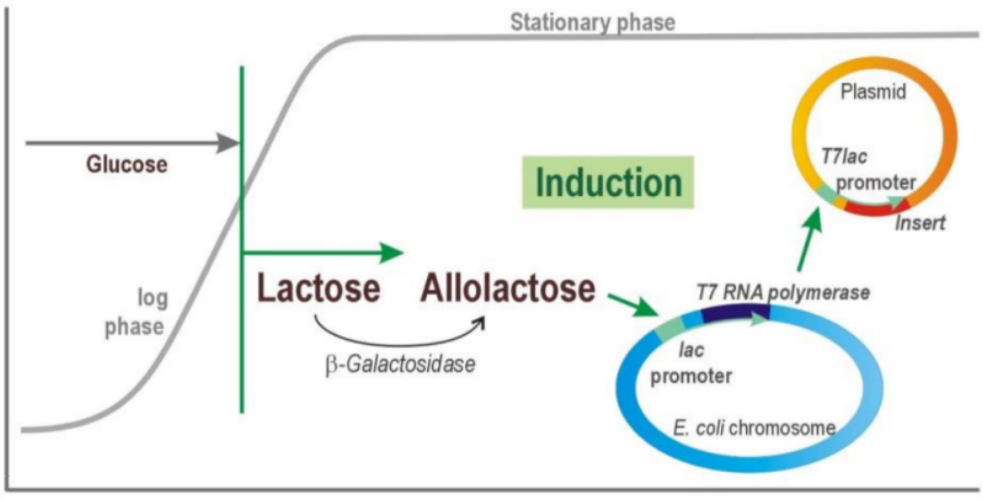



**Figure 4.** Auto-induction medium principles: during the log phase, cells consume preferentially glucose. When glucose is depleted, protein expression is induced by the uptake of lactose, which is converted in allolactose by the bacterial β-galactosidase. In turn, allolactose displaces the T7 pol repressor, allowing the transcription of the *lac* operon containing the insert sequence (adapted from Carl Roth, 2016 [302]).

lactose (or an analogue molecule) induces the expression of the T7 RNA pol, which in turn induces the expression of the recombinant gene. T7 RNA pol has some basal transcription even in absence of inducer, which is prevented by placing the lac operon sequence right downstream of the T7 promoter, generating a T7*lac* so that *lac* repressor hinders the T7 RNA pol elongation complex formation (Figure 3).

To induce the expression of the T7 RNA pol, an auto-induction medium containing glucose and lactose can be used. During the exponential growth phase, the recombinant protein expression is prevented since bacteria intake preferentially glucose rather than lactose. Once the culture reaches saturation, the glucose in the medium runs out, the lactose intake increases, thus generating allolactose and triggering protein expression (Figure 4).

Alternatively, bacterial cultures exponentially growing in LB medium can be supplemented with Isopropyl β-D-1-thiogalactopyranoside (IPTG), a mimic of allolactose able to trigger the *lac* operon. The advantage of using the IPTG is that this molecule is not metabolized unlike lactose, maintaining a constant concentration in the medium and so the recombinant protein expression [303].

*6xHis-IκBα.* His-IκBα full-length was expressed in the BL21(DE3) strain, using auto-induction medium (adapted from the Studier protocol [303]). Cells were transfected by heat shock and incubated in 200 µL of Super Optimal broth with Catabolite repression (SOC) medium for 1 hour at 37° C. 100 µL of cells were inoculated in 5 ml of LB plus 34 µg/ml of kanamycin and incubated overnight at 37 °C. 5 mL of pre-culture were inoculated in 0,5 L of auto-induction and incubated at 37 °C until the optical density (OD600) reached 0.2-0.8 cultures. Cultures were incubated overnight at 18° C. Aliquots of 1 L were prepared and stored at -80° C.

*6xHis-NEMO*. His-NEMO was expressed in *E. coli* Rosetta strain optimized for eukaryotic DNA containing rare codons to avoid the formation of truncated forms of NEMO. The tRNA rare codon enconding plasmid contains a chloramphenicol resistance gene. Chloramphenicol and kanamycin were added at half of the standard usage concentration (17 µg/ml and 12.5 µg/ml respectively). 0.5 L cultures were cultivated in LB plus 0,5 mM $ZnSO_4$ and the protein expression was induced by the addition of IPTG once the OD600 reached 0.2-0.8. Cultures were incubated at 15 °C overnight.

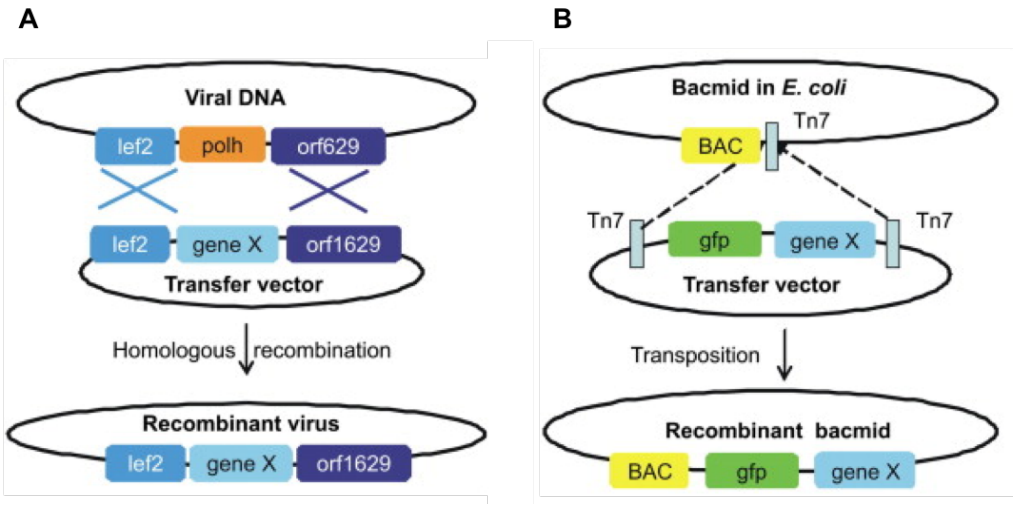



**Figure 5.** Generation of recombinant viruses. **A)** Homologous recombination between a transfer plasmid and the viral genome. The transfer plasmid contains the gene of interest under control of the *p10* or *polh* promoter. This cassette is flanked by baculovirus genome fragments, which allow for homologous recombination once co-transfected. **B)** Bacterial artificial chromosomes containing the baculovirus genome (Bacmid) allow recombination in *E. coli* by site-directed transposition *via* the Tn7 recombinase-mediated transposition. This method improves viral titers shortening the whole process (taken from van Oers, 2011 [304]).

*MBP-fusion constructs.* The MBP-constructs were expressed in BL21(DE3). Cells were transfected by heat shock and incubated in 200 µL of SOC medium for 1 hour at 37 °C. 100 µL of cells were inoculated in 5 ml of LB plus 34 µg/ml of kanamycin and incubated overnight at 37° C. 0.5 mL of pre-culture were inoculated in 50 ml of auto-induction and incubated at 37 °C until the optical density (OD600) reached 0.2-0.8 cultures. Cultures were incubated overnight at 18 °C.

### 4.2.2 Protein expression in insect cells

*Principles of protein expression in insect cells.* Baculoviruses are arthropod-specific viruses that belong to the Baculoviridae family that infect around 600 insect species of several orders. Baculoviruses have been engineered to create a BEVS, which is a powerful method to express a large number of recombinant proteins and genes in insect cells. The BEVS is optimal for the expression of recombinant eukaryotic proteins thanks to the high gene expression levels induced in optimized insect cells, the capacity of cells to provide a correct folding and proper PTMs such as glycosylation, acetylation and phosphorylation, thus ensuring the correct protein folding. BEVS also allows for the expression of multiprotein complexes, which are useful in case of proteins that require scaffolds, chaperons or other interacting partners for their correct folding [305]. Finally, insect cells are relatively easy to grow and to scale up, since they are cultivated in serum-free medium without $CO_2$ at 27-28 °C and they show resistance to bacterial contaminations.

To generate a recombinant virus, cells are co-transfected with a transfer vector containing the sequence of interest and the viral DNA. This step can be carried out in insect cells by host-enzyme-mediated homologous recombination or in *E. coli* by site-specific transposition (Figure 5). Once the recombinant virus is prepared and the expression conditions such as MOI (multiplicity of infection), harvesting time, cell line etc. are optimized, the expression can be scaled up.

It was previously reported that protein expression in insect cells is the best option for expression of the IKK1 and IKK2 catalytic subunits [154, 8]. Given their large size (90 kDa), IKK1 and 2 likely need chaperons for their correct folding, which can't be provided by bacteria.

*Recombinant IKK1 and IKK2 expression in insect cells.* IKK recombinant proteins were expressed in *Spodoptera frugiperda* Sf9 insect cells (Novagen) adapted to a serum-free medium (Sf-900 II SFM). The recombinant virus containing the cDNA of IKK (bacmid) was generated by transfection using lipofectamine. A T25 flask was seeded with $8*10^5$ cells in 3 ml of medium and incubated for at least 20 minutes to let cells adhere to the plastic surface. The transfection mix consisted of 6 µl of lipofectamine, 1 µg of transfer vector and 2 µg of bacmid in 200 µl of Grace's

insect cell medium (G8142, Sigma-Aldrich). After washing the cells, the transfection mix was added in the flask with 3 ml of Grace's insect cell medium and incubated for 3 hours at 27° C. Then, the medium was changed with 5 ml of Sf-900 with 10% SMF supplemented with 10% FCS. Cells were incubated for 1 week at 27° C.

Large scale productions of 1 or 3 L cultures wer performed in Sf21 insect cells (Novagen), which replicate faster than Sf9. The MOI (multiplicity of infection) was set on 1, assuming a titer of $5*10^7$ pfu/ml (60 µl of virus at MOI=1 for a well containing $10^6$ cells). The Sf2 culture with a cell density of $10^6$ cell/ml was transfected with the proper amount of virus (MOI=1) and incubated for 48 hours at 27° C. Cells were harvested, resuspended in phosphate-buffer saline (PBS) pH 7 and flash frozen in liquid nitrogen.

## 4.3 Protein purification

All proteins were purified by affinity chromagraphy (amylose resin for MBP-constructs and $Ni^{2+}$-NTA resin for His-constructs) usually followed by gel filtration chromatography using Superdex 75 or Superdex 200 columns, depending on the size of the constructs. In some cases, the fusion tag was cleaved off by TEV (Tobacco Etch Virus) protease digestion.

*6xHis-IκBα constructs.* Proteins were purified by affinity chromatography using $Ni^{2+}$-NTA resin (Qiagen). 250 ml pellets were resuspended in 25 ml of lysis buffer (200 mM NaCl, 25 mM Bis-Tris(hydroxymethyl)aminomethane hydrochloride (Bis-Tris) pH 8, 10 mM imidazole, 100% Roche Complete™ protease inhibitors cocktail (PIC), 2 mM 1,4-dithiotheritol (DTT), DNase and RNase, lysozyme, 10% glycerol, 0.2% NP-40) and sonicated for 90 seconds (1 second ON and 0.5 seconds OFF) 4 times at 30% amplitude. Clarified lysates were centrifuged 45 minutes at 36000 rpm in a Ti70 ultracentrifuge rotor (Beckman). The supernatant was loaded in a 20 ml gravity column filled with 0.5 ml (dry bead volume) of pre-equilibrated Ni-NTA agarose resin. After the first loading of the clarified supernatant, the flow-through was re-loaded two more times to maximize protein binding to the resin. The resin was extensively washed in four steps: 10 column volumes of wash buffer 1 (200 mM NaCl, 25 mM Bis-Tris pH 8, 10 mM imidazole, 2 mM DTT), 10 column volumes of wash buffer 2 (200 mM NaCl, 25 mM Bis-Tris pH 8, 25 mM imidazole, 2 mM DTT), 10 column of wash buffer 3 (1 M NaCl, 25 mM Bis-Tris pH 8, 2 mM DTT) and 10 column volumes of wash buffer1. The protein was eluted with 2.5 ml of elution buffer (200 mM NaCl, 25 mM Bis-Tris pH 8, 250 mM imidazole, 10% PIC, 2 mM DTT).

For 6xHis-IκBα samples used in the kinase activity experiments, imidazole was eliminated using a PD-10 desalting column (GE Healthcare) in final buffer (150 mM NaCl, 20 mM Bis-Tris pH8, 2 mM DTT). For 6xHis-IκBα samples used for interaction experiments with NEMO an additional gel filtrations step using a Superdex 20 10/60 column equilibrated in final buffer was performed.
6xHis-IκBα samples were aliquoted and flash-frozen in liquid nitrogen for long-term storage.

*6xHis-NEMO*. 1 L bacterial expressions of 6xHis-NEMO were lysed and purified by Ni-NTA following a protocol identical to that described for *6xHis-IκBα constructs*, except that equilibration/wash 1, wash 2, wash 3 and elution buffers included 10% glycerol. Following the affinity chromatography step, the NEMO samples were concentrated and loaded on a Superdex 200 10/30 column equilibrated in gel filtration buffer (400mM NaCl, 20mM Bis-Tris pH 8.0, 2mM DTT, 5% Glycerol).

*MBP-fusion constructs.* The MBP-constructs were coupled to the amylose resins prior usage for the pulldown experiments. Pellets of 10-25 ml of bacterial culture were resuspended in 2.5 ml of lysis buffer (250 mM NaCl, 20 mM Bis-Tris pH 8, 100% PIC, 2 mM DTT, DNase and RNase, lysozyme, 10% glycerol, 0.2% NP40) and sonicated for 1 minute, 3 times at 30% amplitude. Sonicated samples were clarified by high-speed centrifugation at 4° C. 1.5 ml of clarified samples were incubated with 100 µL (dry bead volume) of previously equilibrated dry amylose resin (BioLabs) for 45 minutes at 4 °C on rotating wheel. Resins were washed with equilibration/wash buffer 1 (250 mM NaCl, 20 mM Bis-Tris pH 8, 10% PIC, 2 mM DTT) and wash buffer 2 (150 mM NaCl, 20 mM Bis-Tris pH 8, 10% PIC, 2 mM DTT) and stocked in 200 µL of wash buffer 2 at 4° C. In the case of MBP-IRF7 peptide constructs, to avoid digestion of the peptides, ethylenediamenetetraacetic acid (EDTA) 5 mM and PMFS 0.5 mM were used in addition to 100% PIC in the lysis buffer. For these peptides, sonication was reduced to 3 cycles of 30 seconds and incubation with amylose resin was performed keeping the sample always on ice.

*6xHis-IKK1(10-667) EE and 6xHis-IKK2(1-669) EE.* IKK1 and IKK2 recombinant samples used for the *in vitro* pulldown and kinase assays and pulldown assays were expressed and purified using the same protocol. For these assays, one pellet of 0.5 L of insect cell culture was resuspended in 50 mL of lysis buffer (200 mM NaCl, 25 mM Bis-Tris pH 8, 100% PIC, 10 mM imidazole, 2 mM DTT, DNase and RNase, 10% glycerol) and sonicated for 1 minute (1 second ON and 1 second OFF), 2 times at 20% amplitude. The lysate was centrifuged for 45 minutes at 14000 rpm using a Ti70 rotor. The Ni-NTA affinity chromatography purification protocol was similar to the one described for

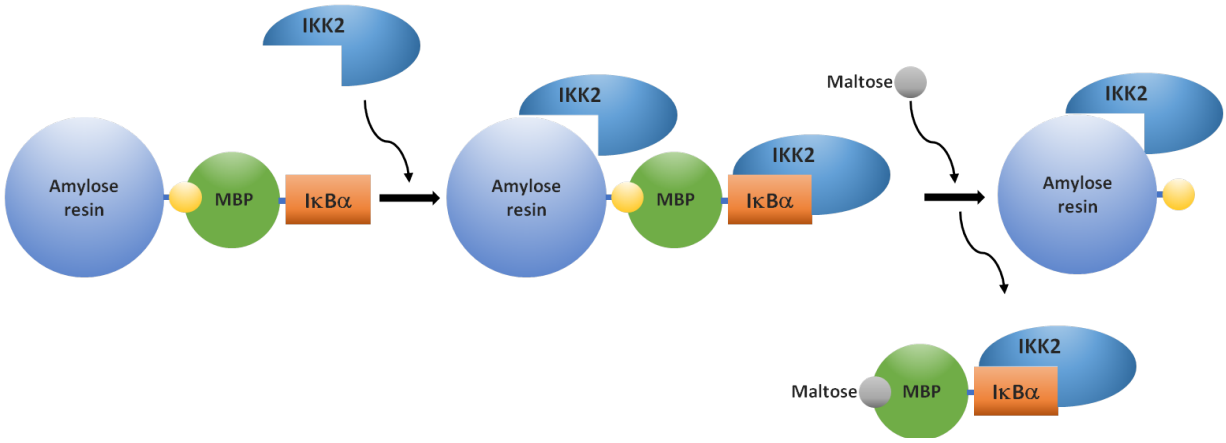



**Figure 6.** Schematic representation of a pulldown experiment with maltose elution. MBP fusion constructs are coupled to the amylose resin and incubated with the bait protein. After washing, the complexes are eluted from the resin by addition of maltose, reducing the background signal due to nonspecific binding with the resin.

6xHis-IκBα with some modifications. All buffers were supplemented with 10% glycerol and extensively degassed before addition of additives (DTT, PIC etc.). A higher volume of Ni-NTA resin was used (1 ml). After elution from the Ni-NTA resin, the sample was concentrated to 0.5 ml and incubated 1h at 4 °C in presence of 1 mM ATP, 20mM $MgCl_2$, 20mM β-Glycerophosphate, 10mM NaF, and 1mM NaOrthovanadate. Then, the sample was loaded on a Superdex 200 10/30 column pre-equilibrated in gel filtration buffer (20mM Tris pH 8.0, 150mM NaCl, 5% Glycerol, 2mM DTT). 6xHis-IKK2(1-669) EE samples for crystallization was expressed with the same protocol. Pellets of 1 to 3 L of insect cell culture were purified for the different crystallization experiments. The lysis and Ni-NTA purification steps were performed as described above. For these larger preps both lysis buffer volumes and Ni-NTA resin volumes were scaled up. After the affinity chromatography step, the 6xHis tag was removed by digestion with TEV protease. For this, the sample was placed in a Geboflex device in the presence of TEV protease and dialyzed overnight against 2 L of dialysis buffer (20 mM Tris pH 8, 150 mM NaCl, 2 mM DTT, 5% glycerol). The sample was loaded onto a column filled with 1 ml Ni-NTA resin pre-equilibrated in dialysis buffer and the flow-through containing the cleaved IKK2 protein collected.

The sample was concentrated incubated for 1 hour on ice with 20 mM $MgCl_2$, 20 mM β-Glycerophosphate, 10 mM NaF and 1mM $Na_3VO_4$ and MNL120B added at a 1:20 stoichiometric ratio (protein/inhibitor). Glycerol concentration was increased up to 10% to avoid precipitation due to the DMSO present in the MNL120B stock solution. The sample was loaded into a Superdex 200 16/60 column equilibrated in gel filtration buffer (20 mM Hepes pH 8.0, 250 mM NaCl, 2 mM TCEP). Collected fractions were pulled together, the salt concentration adjusted to 150 mM and the protein concentrated to 15 mg/ml.

## 4.4 Protein interaction methods

*Principles of the protein-protein interaction assays used in this work.* Throughout this work, protein-protein interactions were investigated mainly using an *in vitro* pulldown assay and the *in vivo Gaussia princeps* complementation assay (GPCA). The pulldown assay was used to test *in vitro* interactions between purified proteins. Here, the bait protein is fused to the MBP tag and immobilized by affinity capture to amylose resin. Resins are then incubated with the purified prey protein. After the incubation, the resin is washed and then the complex (bait + prey) eluted by buffer containing the maltose competitor (Figure 6). This elution method drastically reduces the background due to non-specific binding of prey protein to the beads. The pulldown samples are analyzed by SDS-PAGE.

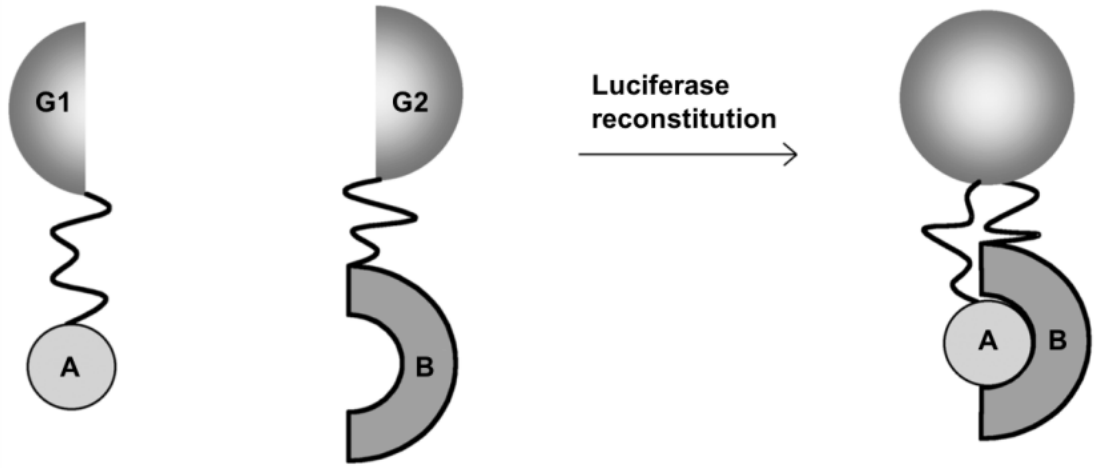



**Figure 7.** Schematic representation of the Luciferase reconstitution mediated by protein interaction. Proteins A and B bring G1 and G2 complementary fragments of luciferase, allowing the enzyme reconstitution. Addition of a proper substrate generates a luminescent signal (Adapted from Muller and Demeret, 2014 [306]).

The GPCA approach is a protein complementation assay used to investigate protein-protein interactions *in cellulo*. The *Gaussia princeps* luciferase is a small monomeric protein of 19.9 kDa that catalyzes the ATP-independent oxidation of the substrate coelenterazine, which emits blue light at 480 nm. The luciferase can be split in two fragments (Gluc1 or G1 and Gluc2 or G2). In the GPCA assay, these two fragments are fused to two proteins of interest. If the target proteins interact, the two luciferase fragments are brought in close contact and the native structure of the luciferase enzyme is reconstituted. In this case, the addition of coelenterazine generates a luminescent signal (Figure 7). The intensity of the interaction is extrapolated by the normalized luminescence ratio (NLR) [307] which is the ratio calculated between the luminescent signal given by the co-transfected sample and the sum of signals given by the controls (cells transfected with G1 or G2 only) according to the formula:

*Luminescence G1 X G2 / (Luminescence G1 + Luminescence G2)*

Quantitative binding parameters were instead obtained by isothermal titration calorimetry (ITC). With this methodology it is possible to determine the binding affinities, the stoichiometries, and enthalpy and entropy changes (ΔH and ΔS) of the binding reaction. The ITC experiment measures the heat changes upon interaction between two proteins. The sample solution contains the main protein at a known and fixed concentration. Small aliquots of ligand are titrated in the solution and the heat released by the binding reaction is measured until the equilibrium is reached. The ligand is titrated until the reaction reaches saturation. With increasing ligand concentration, the heat released decreases. The peaks obtained at each titration point are integrated and the values plotted against ligand concentration to generate a binding isotherm, which is used to extrapolate the affinity ($K_D$) and thermodynamic values of a reaction. These last experiments were performed by a post-doc from our team (Dr. Changquing Li) in collaboration with Dr. Valdimir Torbeev from our institute.

*MBP-pulldown protocol*. For each pulldown experiment 10 µL amylose resin beads (dry volume) coupled to of MBP-fused constructs (IκBα, IκBβ, IRF7, p100, GFP and GSS peptide - see section 4.3 for purification of MBP-fusion constructs) were mixed with 50 µL of purified 6xHis-IKK1 (10-667) EE (at 1.5 µM) , 6xHis-IKK2 (1-669) EE (at 1.5 µM) or 6xHis-NEMO (at 1.0 µM) in a 0.5 ml Eppendorf. After adjusting the volume to 0.5 ml with PD buffer (150 mM NaCl, 20 mM Bis-Tris pH 8, 10% PIC, 2 mM DTT), samples were incubated for 2 hours at 4 °C on a rotating wheel. Then resins were decanted by low speed centrifugation and were washed 2 times with 0.4 ml of PD buffer. Complexes eluted with 20 µL of elution buffer (150 mM NaCl, 20 mM Bis-Tris pH 8, 10% PIC, 2 mM DTT, 200 mM maltose). The pulldown reactions were migrated on a 10% polyacrylamide SDS-PAGE, which was either colored by Coomassie stain (in case of IKK1 pulldowns) or analyzed by Western blotting (in case of IKK2 pulldowns) using an anti 6xHis-tag murine monoclonal primary

antibody (H1029, Sigma-Aldrich) and anti-mouse polyclonal secondary antibody coupled to the horseradish peroxidase (NXA931, GE Healthcare). Detection was done by enhanced chemiluminescence (WesternBright ECL/Sirius, Advansta).

*GPCA assay protocol*. For protein-protein interaction studies in mammalian cells, a codon-optimized luciferase has been generated [308]. Experiments were performed in 96-well plates. The DNA mix was prepared with 100 ng of G1 and G2 constructs plus 0.6 µL of PEI max in 28 µl of Dulbecco's modified essential medium (DMEM), a minimal synthetic cell culture medium used to maintain cells in tissue cultures and based on glucose, salts, vitamins and essential amino acids, and incubated 30 minutes at room temperature. 100 µl of HEK293T cells at a concentration of $40*10^4$ cell/mL were transfected with 28 µl of transfection mix and incubated 48 hours at 37 °C temperature, 5% $CO_2$. Cells were washed with PBS and incubated for 30 minutes with 40 µl of lysis buffer. Then 50 µL of substrate (*Renilla* Luciferase Assay System, Promega) were added to the lysed cells. Luminescence was detected by luminometer (Centro LB Microplate 960, Berthold technologies).

## 4.5 Kinase activity assays

*Kinase assay using purified recombinant 6xHis-IKK2 (1-669) EE.* 6xHis-IKK2 (1-669) EE and 6xHis-IκBα wild type, C308L and Y305S/D307S mutant constructs constructs were purified as described above. 0.5 nM of purified 6xHis-IKK2 EE (1-669) was incubated with 1 µM IκBα constructs in the reaction buffer (50 mM Bis-Tris-HCl pH 8, 100 mM NaCl, 10 mM $MgCl_2$, 2 mM $MnCl_2$, 0.5mg/ml bovine serum albumin (BSA), 20 mM β-glycerol phosphate, 400 µM $Na_3VO_4$, 2 mM DTT, and 400µM ATP) at 30 °C. Reactions were sampled at different times (0, 10 and 30 minutes). The addition of BSA was included to simulate crowding of the cytoplasmic environment. The reaction samples were analyzed by SDS-PAGE and by Western blotting. Phosphorylated IκBα was detected with a murine primary antibody anti P-S32/P-S36 (Cell Signaling Technology) at a dilution of 1:1000 and a polyclonal secondary antibody anti-mouse coupled to the horseradish peroxidase (NXA931, GE Healthcare).

*Kinase assay using immunoprecipitated endogenous IKK2.* For immunoprecipitation of endogenous IKK2 from HEK293T cells, cells were exposed to TNFα and collected after 10 minutes. Pellets were resuspended in lysis buffer (50 mM mM hepes pH 8, 200 mM NaCl, 0.5% NP-40, 1 mM EDTA, 10 mM KCl, 1 mM DTT, 100% PIC, 50 mM NaF, 50 mM β-glycerophosphate and 1 mM $Na_3VO_4$) and

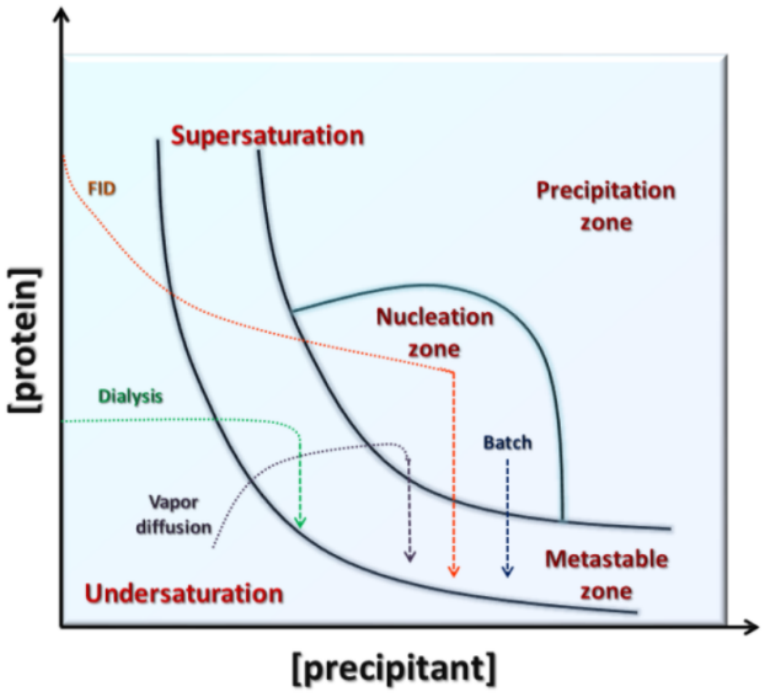



**Figure 8.** A simplified protein crystallization phase diagram showing the different routes to the nucleation and metastable zones for the four main crystallization techniques (taken from Russo Krauss et al., 2013 [309]).

incubated in ice for 30 minutes. The whole-cell extracts were immunoprecipitated with anti-NEMO antibody (Santa Cruz Biotech) bound to protein G Sepharose beads (17-0618-01, GE Healthcare). Beads were washed first with wash buffer 1 (50 mM Bis-Tris, pH 8, 150 mM NaCl, 5 mM β-glycerophosphate, 0.1 mM $Na_3VO_4$, 10 mM $MgCl_2$, 2 mM DTT) and then with wash buffer 2 (20 mM Bis-Tris pH 8, 150 mM NaCl and 2 mM DTT). To perform the kinase assay, the beads were incubated with 1 μM full-length IκBα wild type or mutants in reaction buffer (50 mM Bis-Tris-HCl pH 8, 100 mM NaCl, 10 mM MgCl2, 2 mM $MnCl_2$, 0,5 mg/mL BSA, 20 mM b-glycerol phosphate, 400 μM $Na_2VO_4$, 2 mM DTT, and 400 μM ATP) at 30 °C. Aliquots were taken at time 0, 10 and 30 minutes. Samples were migrated in 10% SDS-PAGE gel analysed by Western blotting. Phosphorylation of IκBα constructs was detected using an anti-phospho-IκBα antibody (9246, Cell Signaling Technology). Endogenous IKK2 was detected by anti-IKK2 antibody (2684S, Cell signaling Technology) whereas IκBα substrates with an anti-His antibody (Sigma).

*Kinase assay using $^{32}P$-labeled ATP.* 50 nM of purified 6xHis-IKK2 EE (1-669) EE were incubated with 5 μM of 6xHis-IκBα constructs (WT, S32A/S35A, Y305E/D307E, S32A/S35A/Y305E/D307E). The kinase reaction buffer consisted of 20 mM Bis-Tris-HCl pH 8, 100 mM NaCl, 2 mM DTT, 10 mM $MgCl_2$, 2 mM $MnCl_2$, 0.2 mM β-glycerol phosphate, 400 μM $Na_3VO_4$, 400 μM ATP. 1 μCi of γ-$^{32}$P-ATP was added to the reaction mixtures, which were then incubated for 2 hours at 30 °C. Reactions were stopped by addition of Laemmli sample buffer and migrated on a 10% SDS-PAGE gel. Gels were stained by Commassie, dried on Waltman paper and analysed by phosphoimaging using a Typhoon FLA 9500 instrument (GE Healthcare).

## 4.6 Crystallization

*Principles of protein crystallization.* Protein crystallization is the process of generation of an ordered crystal structure by aggregation of individual proteins or protein complexes so that its diffraction patterns can be used to determine the tridimensional structures (Figure 8). To form a crystal, proteins must be brought to a supersaturated state, where the concentration of the solute exceeds the saturation concentration. This metastable condition is characterized by an excess of solute dissolved in the solvent (over the saturation concentration), which may lead to solvent aggregation and crystallization by shifting the equilibrium of the system, according to the thermodynamic principles (enthalpy, entropy and temperature). Crystallization requires a first nucleation step in which individual molecules pack together in a translational symmetric manner forming a stable nucleus which grows by attachment of other molecules following the initial spatial disposition. So

far, different techniques have been used for protein crystallization such as vapor diffusion, mircobatch, microdialysis and free-interface diffusion.
Several parameters are implicated in the crystal formations and oftentimes there is no way to predict which ones have an impact on the process. Parameters such temperature, pH, solution composition (buffers, ions, small molecules) need to be tested and eventually fine-tuned.

*Crystallization of IKK2(1-669) EE bound to the IκBα docking peptide.* Crystallization experiments were performed at the macromolecular crystallization platform of IGBMC (Strasbourg) with the help of Pierre Poussin Courmontagne (platform engineer). Purified IKK2 (1-669) EE at a concentration of 15 mg/ml was mixed with a 1.5-fold excess of synthetic peptide comprising IκBα residues 297-317 and incubated one hour on ice. The initial crystallization screen was carried out with six commercial kits: Berkeley, JCSG+, ProComplex, Classics, The PEGs, Wizard I&II, by the sitting drop vapor diffusion method at 277° K using a Mosquito Crystal nano-liter dispensing robot (SPT Labtech). Drops of 0.1 µl protein sample were mixed with 0.1 µl of crystallization mix and equilibrated against 50 µl of reservoir solution in 96-well MRC2 sitting drop vapour diffusion crystallisation plates (Molecular Dimensions). Plates were stored at 277° K and automatically imaged using a RockImager 1000 (Formulatrix). Small initial crystals (dimensions of 50x15x15 µm approximately) appeared after 2 days in condition 4 of the Berkeley screen (1.8 M $(NH_4)_2SO_4$, 0.1 M Bis-Tris HCl pH 6.5, 10% 2-propanol) but appeared multiple. We therefore set up a second screen, which was optimized around this condition. The second screen consisted of 25 conditions, with $(NH_4)_2SO_4$ concentrations ranging between 1.6 M and 2.2 M, pH ranging between 6 and 7, and +/- isopropanol. After eight days two large single crystals (around 500x80x40 µm) grew in one of the drops of the second screen (1.6 M $(NH_4)_2SO_4$, 0.1 M Bis-Tris HCl pH 6.5, 10% 2-propanol). The crystals were transferred to a 2 M $Li_2SO_4$, 0.1 M Bis-Tris HCl pH 6.5 cryoprotectant solution before flash cooling in liquid nitrogen.

### 4.7 Structure determination of IKK2(1-669) EE bound to the IκBα docking peptide by X-ray crystallography

Our collaborator Alistair McEwen from the macromolecular crystallography platform at IGBMC performed x-ray diffraction data collection and structure determination. Data were collected at 100° K on an EIGER X 9M detector (Dectris) at the Proxima 2A beamline of Synchrotron SOLEIL. 360° of data were collected using 0.1° rotation and 0.025s exposure per image at 100% beam intensity at an energy of 12.65 keV (wavelength = 0.98 Å). The data were processed and scaled using XDS [310] before anisotropic truncation and correction was performed using the STARANISO server (Tickle *et*

*al*., 2018). The crystal diffracted anisotropically to 4.2 Å (6.8 Å in the worst direction) and belonged to the C-centered monoclinic space group C2 with unit cell dimensions a = 226.3 Å, b = 136.8 Å, c = 204.4 Å, β = 91.45°. The structure of the IKK2/IκBα peptide complex was solved by molecular replacement using PHASER [311] in the PHENIX suite [312] using a monomer extracted from the structure of the asymmetric dimer of human IKK2 (PDB: 4KIK) as a search model. The asymmetric unit contains five copies of the protein (two dimers and one monomer that forms a dimer with itself over a crystallographic 2-fold rotation axis), with a corresponding Matthews' coefficient of 4.1 $Å^3$ $Da^{-1}$ and a solvent content of 69.9 %. Refinement of the structures was performed using PHENIX [313] and BUSTER followed by iterative model building in COOT [314].

# Part 3

# RESULTS AND DISCUSSIONS

# Objectives

The role of the IKK complex in the activation and regulation of the NF-κB signaling pathway has been widely studied. Despite the three-dimensional structures of the IKK1 and IKK2 catalytic subunits have been already resolved, as well as those of the IκB and the NF-κB protein family members, little is known about the mechanisms of the interactions between the IKK kinase and their substrates.

The general aim of this thesis project is to characterize the interactions between the IKK kinase complex and its partner proteins with a main focus on the substrates acting in NF-κB signaling. In particular, the project is articulated around three objectives that aimed to:

1. identify of the regions of the core IKK complex subunits (IKK1, IKK2 and NEMO) and their substrates involved in the interaction;
2. investigate the impact of the interactions on the activity and functions of the IKK kinase;
3. provide structural data of these interactions.

Work on these objectives involved an integrated approach that combined Biochemistry and Structural Biology methods with Cellular and Functional Biology approaches.

# Chapter 5

# Identification and characterization of the IKK1 and IKK2 interactions with a novel docking motif

The NF-κB signaling pathway is known for its fundamental role in the activation and regulation of the inflammatory processes and in the innate immune response. In absence of stimuli, the NF-κB dimers are retained in the cytoplasm by interaction with either IκB proteins or with the C-terminal auto-inhibitory tails of p100 or p105 precursors. The IKK complex is responsible for the activation of NF-κB transcriptional activity. Upon proper stimuli, the assembly of a complex signaling platform ultimately results in the activation of the IKK kinase complex, which, in turn, targets and phosphorylates IκB proteins and p105 or p100 at specific serine residues. This phosphorylation generates a degron motif recognized by the E3 ubiquitin ligase SCF-bTrCP. The resulting polyubiquitination leads to the degradation of the IκB proteins or to partial proteolysis of the C-terminal region of p100 and p105, thus allowing for translocation of the nuclear factors into the nucleus and activation of transcription.

The IKK canonical complex consists of the two homologous catalytic subunits IKK1 and IKK2 and of the scaffold/dimerization NEMO subunit at a stoichiometric ratio of 1:1:2. In this case, only the IKK2 subunit is able to phosphorylate IκB substrates. The IKK species regulating the alternative NF-kB pathway consists instead of an IKK1 homodimer, which targets p100.

In this chapter I describe our work that led to identification of a novel docking motif and the structure-function characterization of its interactions with the IKK1 and IKK2 catalytic subunits. The results presented here are described in the following manuscript, which is currently in preparation:

*Changqing Li[#], Stefano Moro[#], Kataryna Shostak, Francis O'Reilly, Andrea Graziadei, Alistair Mc-Ewen, Pierre Poussin-Courmontagne, Thomas Bachelart, Denise Martinez-Zapien, Mert Fiskin, Marie Laure Straub, Georges Orfanoudakis, Massimo Tommasino, Arnaud Poterszman, Vladimir Torbeev, Juri Rappsilber, Norman Davy, Alain Chariot and Katia Zanier.* A novel YDDΦXΦ linear motif mediates docking of the IKK kinase to its substrates: structural and functional insights.

[#] *co-first authors*

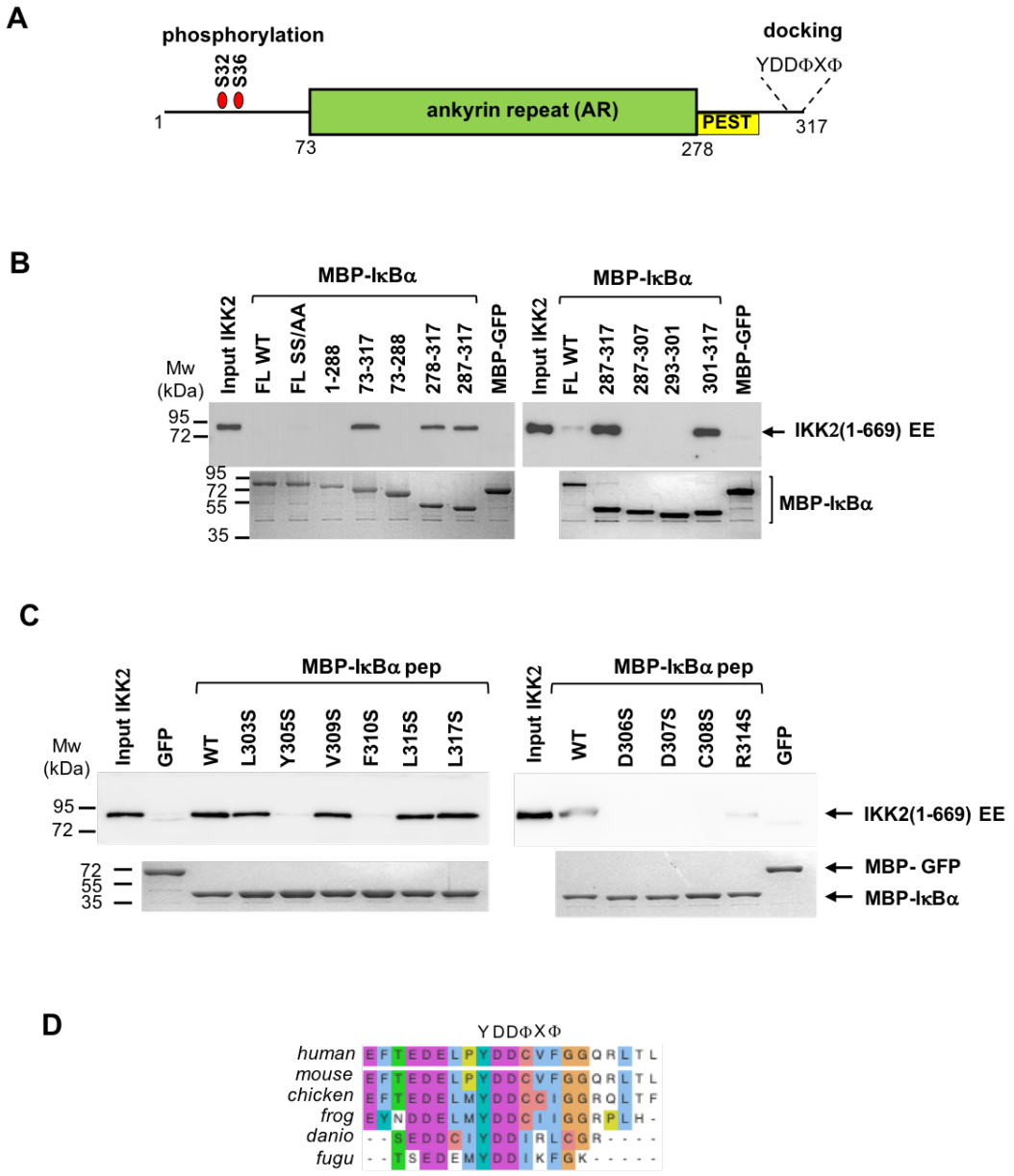



**Figure 1.** IKK2 interactions with a docking site within IκBα. **A)** Schematic representation of the domain architecture of the IκBα substrate showing the disordered N-terminal region containing the phospho-acceptor residues (Ser32 and Ser36) (residues 1-72), the ankyrin repeat domain (AR) (residues 73-278) and the disordered C-terminal region containing the PEST rich sequence followed by the YDDΦXΦ docking site (residues 278-317). **B)** *(Left panel)* Pulldown analysis of the interactions between IκBα full-length or IκBα deletion constructs with purified recombinant IKK2. MBP-IκBα constructs were coupled to amylose resin and incubated with the purified 6xHis-IKK2(1-669) EE construct (0.4 μM concentration). Samples were migrated on 2 separate 10% SDS-PAGE gels. IKK2 was detected by Western blotting using an anti-His antibody, whereas the MBP-IκBα constructs were detected by Coomassie staining. MBP-GFP was used as negative control. (*Right panel*) Pulldown experiment showing the interaction with full-length IκBα at higher concentrations of the 6xHis-IKK2(1-669) EE construct (5 μM concentration). IKK2 is visualized by Western blotting using an anti-His antibody, whereas MBP-IκBα and the negative control MBP-GST by Ponceau staining of the same Western blot membrane. **C)** Pulldown analysis of the 6xHis-IKK2(1-669) interactions with MBP-IκBα(301-317) peptides (MBP-IκBα pep) containing the indicated serine substitutions. Substitution of residues Y305, D306, D307, C308 and F310 prevent the interaction. **D)** Sequence alignment of the IκBα C-terminal region containing the docking site for IKK2. The YDDΦXΦ consensus is conserved in IκBα orthologues.

## 5.1 Results

### 5.1.1 Identification of a docking motif within the canonical IκBα substrate

The IκBα protein acting in the canonical NF-κB pathway is the most studied substrate of IKK. IκBα comprises a disordered N-terminal region (residues 1-72), hosting the residues Ser32 and Ser36, which are phosphorylated by the IKK2 catalytic subunit of the canonical IKK complex, an ankyrin repeat domain (residues 73-278) and a disordered C-terminal region containing a PEST sequence (residues 278-317) (Figure 1A). In the canonical NF-κB pathway, IKK2 phosphorylation of residues Ser32 and Ser36 of IκBα substrate generates a phospho-degron motif that leads to substrate polyubiquitination and subsequent degradation *via* the proteasome. Previous work suggested the existence of a docking site for IKK2 located somewhere within the ankyrin repeat domain and C-terminal regions of IκBα [8].

To further characterize the IKK2-IκBα interaction, we performed *in vitro* pulldown experiments employing a panel of IκBα deletion constructs and purified recombinant IKK2 protein. The IKK2 construct used corresponds to the large fragment previously crystallized by Huxford and collaborators [7]. This construct, named 6xHis-IKK2 (1-669) EE, comprised residues 1-669 of IKK2 and contains two phosphomimic mutations in the activation loop that make the kinase constitutively active. 6xHis-IKK2 (1-669) EE was produced in insect cells by BEVS and purified by $Ni^{2+}$-NTA affinity chromatography followed by gel filtration chromatography. IκBα constructs were produced as N-terminal fusions to the MBP in *E. coli* and coupled to amylose resin. After incubation with purified IKK2, the complexes were eluted by the addition of maltose competitor, reducing in this way the background signals due to unspecific interaction between IKK2 and the amylose resin. Samples were run on two different 12% SDS-page gels. IKK2 was detected by Western blotting using an anti-6xHis tag antibody, whereas the MPB-IκBα deletion constructs were detected by Coomassie stain. In this way we were able to properly quantify the IKK2 band even in the case of comigration with MPB-IκBα constructs of similar size. The results from these experiments show that the IKK2 interaction with full-length IκBα wild type or S32A/S36A (lacking the IKK1 phospho-acceptor residues) mutant is hardly detectable by this binding assay when relatively low concentrations of the IKK2 protein (0.4 μM) are used (Figure 1B *left panel*). However, this interaction becomes visible at higher concentrations of the kinase (about 5 μM) (Figure 1B *right panel)*, thus confirming the functionality of the 6xHis-IKK2 (1-669) EE construct already demonstrated by Liu *et al.* [154]. Unexpectedly, deletion of the N-terminal disordered region of IκBα enhances the interaction with IKK2 (Figure 1B

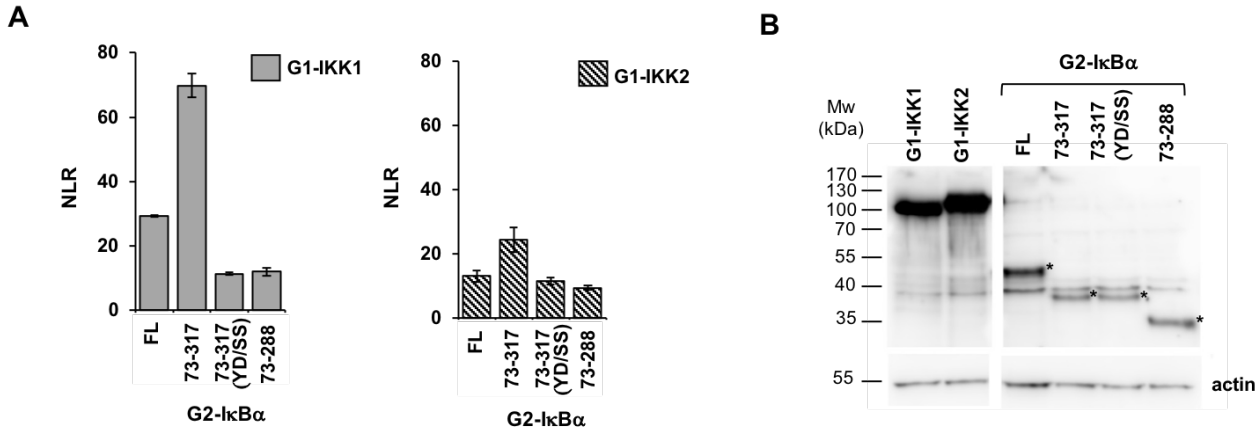



**Figure 2.** IKK1 and IKK2 interactions with the YDDΦXΦ docking site. **A)** GPCA analysis of the interactions between full-length IKK1 (FL IKK1) or full-length IKK2 (FL IKK2) and IκBα constructs. IKK1 and IKK2 were fused to Gluc1 (G1) fragment of the luciferase whereas the IκBα constructs were fused to Gluc2 (G2) luciferase fragment. HEK293T cells were co-transfected with G1-IKK1 or G1-IKK2 and G2-IκBα vectors. The IκBα constructs tested included the full-length protein (FL), the 73-317 construct lacking the N-terminal disordered region and the 73-288 construct corresponding to the AR domain. YD/SS corresponds to the double Y305S/D307 mutation suppressing the YDDΦXΦ docking site. The NLR values was calculated according to the formula: *(Luminescence G1 * G2) / (Luminescence G1 + Luminescence G2.)* **B)** Expression levels of G1-IKK1, G1-IKK2 and G2-IκBα in HEK293T cells. G1-IKK1 and G1-IKK2 fusions were detected by Western blotting using an antibody recognizing the *Gaussia princeps* luciferase. The signal for the G2-IκBα constructs (indicated by an asterisk *) appear lower than that for the G1-IKK1/2 due to a lower affinity of the antibody for the G2 fragment.

*left panel*, compared FL IκBα with IκBα (73-317) construct). Furthermore, a 20 amino acid peptide comprising residues 301 and 317 within the disordered C-terminus of IκBα (i.e. IκBα (301-317)) is sufficient to interact with IKK2 at levels comparable with the IκBα (73-317) construct lacking the N-terminal region (Figure 1B, *left panel*).

We reasoned that the IκBα (301-317) region might represent a conserved docking site for IKK2. Hence, I performed a serine mutagenesis scan of the IκBα (301-317) peptide and tested the interactions of the resulting peptide mutants with recombinant purified IKK2 (1-669) EE by pulldown. This experiment showed that residues Tyr305, Asp306, Asp307, Cys308 and Phe310 are essential for the interaction with IKK2 (Figure 1C). Consistently, sequence alignments revealed that these same residues are conserved in IκBα proteins from different species and are part of a conserved YDDΦXΦ consensus (where Y is a tyrosine, D is an aspartic acid, Φ is a hydrophobic amino acid and X is any amino acid) (Figure 1D).

We also tested the interactions of the IKK1 and IKK2 catalytic subunits with the IκBα substrate *in cellulo* using the *Gaussia princeps* protein complementation assay (GPCA) [24]. This assay, which does not use washing steps, enables detection of protein–protein interactions with a broad range of affinities, including transient interactions. Here, full-length IKK1 and IKK2 proteins were N-terminally fused to the Gluc1 (G1) fragment of the luciferase, whereas IκBα to the Gluc2 (G2) fragment. We created four constructs of IκBα that included the full-length protein, the IκBα (73-317) construct lacking the N-terminal region and the IκBα (73-288) lacking both the N-terminal ad C-terminal region hosting the docking site. In addition, we introduced the double Y305S/D307S mutation within the YDDΦXΦ consensus in the IκBα (73-317) construct. Pairwise combinations of G1-IKK1/2 and G2-IκBα constructs were co-transfected in HEK293T cells. Results show that both IKK1 and IKK2 subunits interact with IκBα, with IKK1 showing slightly higher binding responses (Figure 2). This finding is coherent with the high sequence homology between IKK1 and IKK2. In agreement with the results of the pulldown analyses, truncation of the N-terminal disordered region of IκBα (IκBα (73-317)) increases binding to both IKK1 and IKK2, compared to IκBα full-length. In both cases, further deletion of the disordered C-terminal region of IκBα or mutation of the YDDΦXΦ consensus strongly decrease binding (compared IκBα (73-317) with IκBα (73-288) and with IκBα (73-317) YD/SS). Together these results show that both the IKK1 and IKK2 proteins recognize the YDDΦXΦ docking site within the IκBα substrate both *in cellulo* and *in vitro*.

To further characterize the contributions of the YDDΦXΦ consensus, I cloned fifteen new mutants of the IκBα (301-317) peptide (3 mutations per conserved position of the consensus motif) and evaluated their impact on the interactions with recombinant purified IKK1 and IKK2 proteins by

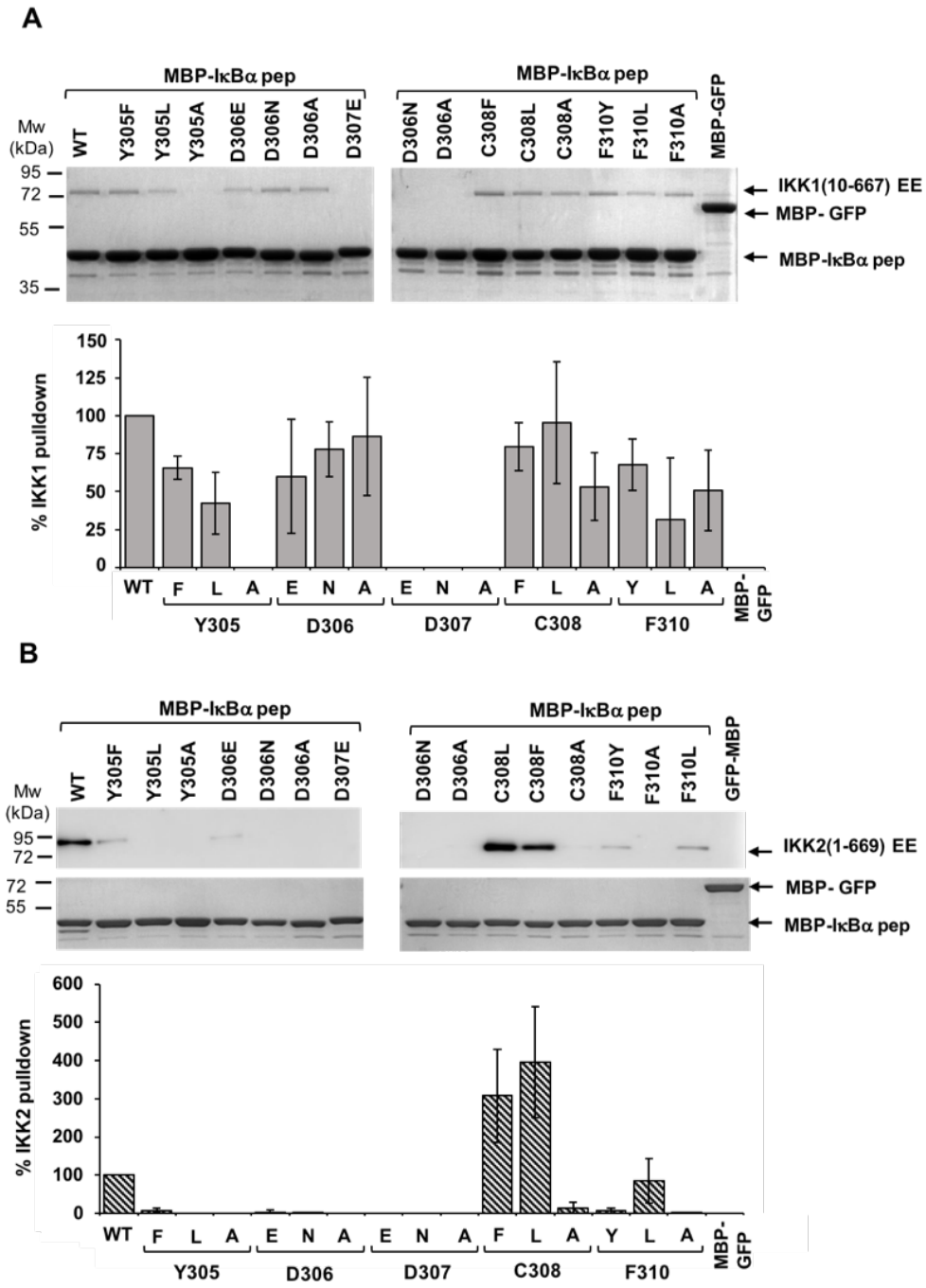



**Figure 3.** IKK1 and IKK2 interactions with IκBα pep mutants. Characterization of the YDDΦXΦ linear motif in the interaction with IKK proteins. Three mutants were produced for each conserved position of the YDDΦXΦ consensus, and tested against purified IKK1 and IKK2 proteins. **A)** *(Upper panel)* Representative image of the pulldown analysis of the interactions between 6xHis-IKK1 and IκBα pep constructs. Samples were migrated on a 10% SDS-PAGE gel. IKK1 and IκBα pep mutants were visualized by Coomassie staining. MBP-GFP was used as negative control. *(Lower panel)* Quantification of the interaction levels between purified recombinant 6xHis-IKK1 and MBP-IκBα pep mutants. The IKK1 bands from the Coomassie SDS-PAGE gel analysis were quantified using ImageJ and normalized to the IKK1/IκBα pep wild type interaction. The data are averages of three independent experiments with error bars representing standard deviation values. **B)** *(Upper panel)* Representative image of the pulldown analysis of the interactions between IKK2 and IκBα pep mutants. Samples were migrated on 10% SDS-PAGE gels. IKK2 was detected by Western blotting using an anti-His antibody, whereas the IκBα pep constructs were detected by Coomassie staining. MBP-GFP was used as negative control. *(Lower panel)* Quantification of the interaction levels between purified recombinant 6xHis-IKK2 and MBP-IκBα pep mutants. IKK2 bands from the Western blotting analysis were quantified using ImageJ and normalized to the IKK2/IκBα pep wild type interaction. The data are averages of three independent experiments with error bars representing standard deviation values.

pulldown experiments. In these experiments the IKK constructs used were IKK2 (1-669) EE and a similar construct of IKK1, named IKK1 (10-667) EE, which was previously crystallized by Gosh and co-workers [9], and comprises residues 10 to 667 of IKK1 plus the double phospho-mimic mutation S176E/S180E in the activation loop that makes the kinase constitutively active. The results from these experiments are shown in Figure 3. Most of the mutations have a negative impact on the interactions with IKK1 and IKK2. Interestingly, IKK1 seems to tolerate a larger number of substitutions within the docking motif (Figure 3A), whereas IKK2 displays higher sensitivity to mutations (Figure 3B). Most importantly, I could identify two mutations at the first hydrophobic (Φ) position of the motif, namely C308L and C308F, that displayed increased binding activity towards both IKK1 and IKK2.
We also measured the IKK1 and IKK2 interactions with synthetic peptides comprising the wild type and mutant sequence of the YDDΦXΦ docking site within IκBα by isothermal titration calorimetry (ITC). The peptide construct used in these experiments comprises residues 297-317 of IκBα and includes four additional amino acids compared to the peptide analysed in the pulldown experiments (i.e. IκBα (301-317)). In fact, I previously observed that the addition of these four N-terminal residues (EFTE) moderately increases the binding activity in pulldown experiments (data not shown). In the ITC experiments used purified IKK1 (10-667) EE and IKK2 (1-669) EE samples and were performed by my co-worker Changqing Li (post-doc) in collaboration with Professor Vladimir Torbeev from our institute (BSC-UMR7242, Strasbourg). The results, summarized in Table 1, show that IKK1 binds to the IκBα (297-317) peptide with a slightly higher equilibrium binding affinity constant ($K_D$) compared to IKK2 ($K_D$=9 μM *versus* $K_D$=40 μM). The introduction of the C308 mutation in the IκBα (297-317) peptide leads to a 30-fold increase in the affinity for IKK1 ($K_D$=0.35 μM) and a 20-fold increase in the affinity for IKK2 ($K_D$=2.6 μM). Together these results show that both IKK1 and IKK2 catalytic subunits contribute to docking to a specific site of the IκBα substrate that contains a conserved YDDΦXΦ consensus. Our mutagenesis studies allowed us to identify mutations within this consensus that modulate the affinities of the interactions with IKK1 and IKK2 both in a positive and negative manner. Interestingly, IKK1 seems to have slightly higher affinity for the IκBα docking site and to tolerate a larger number of substitutions within the YDDΦXΦ consensus.

### 5.1.2 Activity and functional analyses on the YDDΦXΦ docking interaction

To determine the impact that the docking interaction has on the catalytic activity of IKK2, I expressed and purified full-length IκBα proteins fused to an N-terminal 6xHis tag, either wild type or containing the double suppressive Y305S/D307S mutation or the C308L mutation that increases binding affinity. IκBα proteins were expressed in *E. coli* and purified by $Ni^{2+}$-NTA affinity chromatography. These

| *interaction* | *N* | $K_D$ *(μM)* | *ΔH (kcal/mol)* | *TΔS (kcal/mol)* | *ΔG (kcal/mol)* |
|---|---|---|---|---|---|
| IKK1/IκBα pep WT | 0.62 | 9.0 | -15.2 | -8.35 | -6.88 |
| IKK1/IκBα pep C308L | 0.65 | 0.35 | -13.5 | -4.71 | -8.80 |
| IKK2/IκBα pep WT | 0.58 | 40 | -15.8 | -9.78 | -6.00 |
| IKK2/IκBα pep C308L | 0.63 | 2.6 | -12.2 | -4.59 | -7.61 |
| IKK1/p100 pep | 0.60 | 11 | -12.3 | -5.56 | -6.77 |
| [b]IKK1/IRF7 pep | 0.6 | 90 | -6.7 | -1.2 | -5.5 |

**Table 1** Thermodynamic parameters of YDDΦXΦ peptides to IKK1 or IKK2. All ITC experiments were performed at 25 °C. The data were fitted to a 1:1 binding model. N refers to the molar ratio of IKK1-2: pep complex. [b] Values for the IKK1 interaction with IRF7 peptide should be regarded as estimates due to low affinity and weakly exothermic binding, and therefore high experimental error and fitting uncertainty. All data were processed using NITPIC software [315] and global fitting of two datasets was carried out using SEDPHAD [316]. The ITC experiments were carried out using a PEAK ITC microcalorimeter (Malvern Instruments). The titration parameters were set as following, temperature 25 °C, reference power 10 μcal/s, feedback=high, stirring speed 750 RPM, initial delay 60 s, first injection 0.4 μL in 0.8 s and the remaining 19 injections were 2 μL in 4 s with 120 s of spacing.

three IκBα proteins were evaluated for phosphorylation at residues Ser32 and Ser36 by a kinase assays that used the recombinant and purified IKK2 (1-669) EE construct. In this assay, I kept a low kinase concentrations (0.5 nM) to avoid phosphorylation saturation of the substrate and added BSA to the reaction buffer to mimic the cytoplasmic crowding as previously described by Hoffmann and collaborators [27]. In addition, the substrate concentrations were set to 1 μM to avoid signal saturation. Phosphorylation was detected using a phospho-specific P-S32/P-S36 primary antibody. Results show that the IκBα C308L displays the highest phosphorylation levels. By contrast, phosphorylation of the Y305AS/D307S mutant is lower than that of wild type IκBα (Figure 4A).

The kinase activity experiments were repeated by replacing recombinant IKK2 with the endogenous canonical IKK complex immunoprecipitated from HEK293T cells by an anti-NEMO antibody. These experiments, which were performed by Changqing Li, confirmed higher phosphorylation levels for the IκBα C308L and as compared to wild type IκBα and Y305AS/D307S IκBα constructs (Figure 4B).

We also wanted to verify whether the mutations in the docking site would redirect IKK2 kinase activity to other residues within the IκBα substrate. For this, we performed an *in vitro* kinase activity assay that used recombinant purified IKK2 (1-669) EE kinase, radiolabelled γ-$^{32}$P-ATP and four recombinant purified full-length constructs of IκBα, that included wild type protein, the S32A/S36A and Y305A/D307A and double mutants and the S32A/S36A/Y305A/D307A quadruple mutant. The Y305A/D307A mutation was used here instead of Y305S/D307S to suppress the YDDΦXΦ consensus without introducing potentially new phosphorylation sites. The Y305A/D307A abolishes binding to IKK2 as efficiently as the Y305S/D307S mutation in pulldown assays. In this experiment we wanted to make sure to detect even low amounts of phosphorylation. Therefore, assay conditions included concentrations of both substrate and kinase that were higher compare to the experiments described above (50 nM of kinase and 5 μM substrate) and long incubation times (2 hours). Results are shown in Figure 4C. As expected, the double S32A/S36A mutant does not display any phosphorylation. In these assay conditions, the Y305A/D307A mutant shows similar phosphorylation levels than wild type IκBα. Instead, the quadruple S32A/S36A/Y305A/D307A mutant does not display any phosphorylation, indicating that the suppression of the docking site by the Y305A/D307A mutation does not redirect kinase activity to residues other then Ser32 and Ser36.

In agreement with the findings described above, I observed that the recombinant purified IκBα (297-317) peptide containing the C308L mutation and N-terminally fused to a double 6xHis-MBP tag (6xHis-MBP-IκBα (297-317) C308L) was able to specifically capture the endogenous IKK2 kinase from clarified extracts of HEK293T cells (Figure 5A). Hence, we supposed that this peptide could be used as an inhibitor of IKK2 kinase activity. To test this hypothesis, we performed *in vitro* kinase

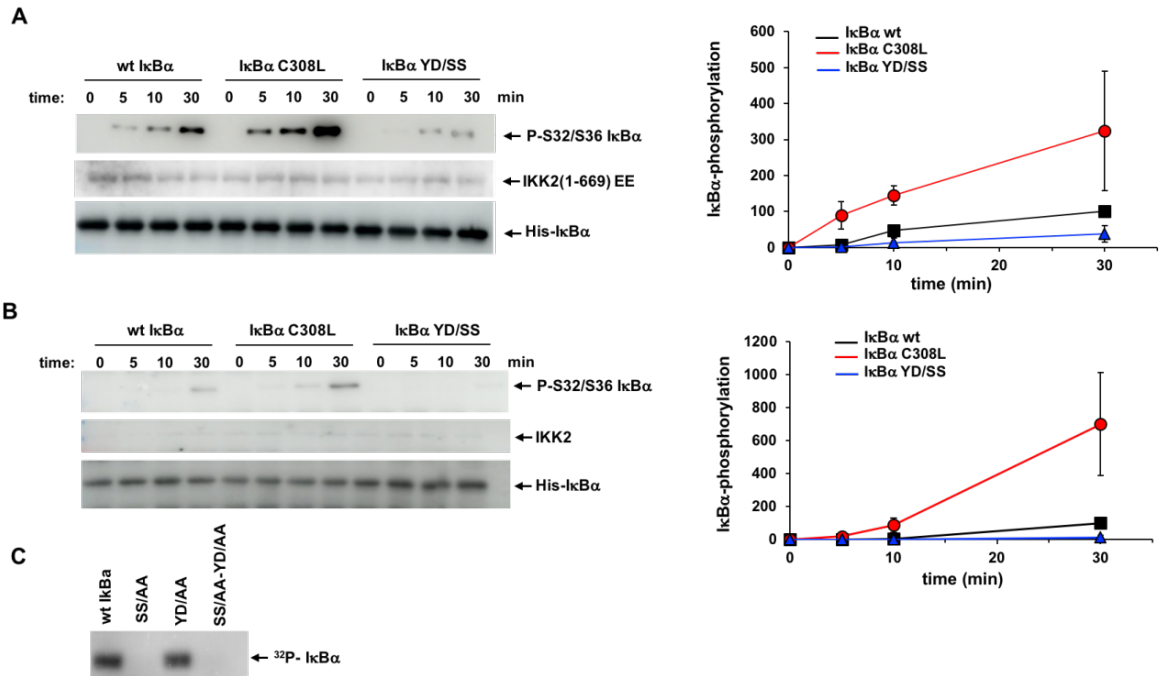



**Figure 4.** Kinase activity analyses on the YDDΦXΦ docking interaction. **A)** (*Left panel*) Representative Western blot images for the kinase experiment using purified recombinant 6xHis-IKK2 (1-669) EE (0.5 nM) and full-length IκBα wild type (wt), C308L and Y305S/D307S (YD/SS) proteins (1 μM). Reaction samples were incubated at 30° C and aliquots were taken at time 0, time 5, time 10 and time 30 (minutes). Reactions were migrated on a 10% SDS-PAGE gel and phosphorylation at Ser32 and Ser36 of IκBα was detected by Western blotting using an anti P-S32/S36-IκBα antibody. 6xHis-IKK2 (1-669) EE and 6xHis-IκBα were detected using an anti-His antibody. (*Right panel*) Quantification of the phosphorylation levels of full-length IκBα wild type, C308L and YD/SS mutants. The phospho-IκBα bands from the Western blotting analysis were quantified using ImageJ. The phospho-IκBα values are normalized to wild type IκBα at 30 min incubation (100%). The data are averages of three independent experiments with error bars representing standard deviation values. **B)** Representative Western blot images for the kinase experiment using immunoprecipitated endogenous IKK2 and purified full-length IκBα wild type (wt), C308L and Y305S/D307S (YD/SS) proteins. HEK293T cells were harvested after stimulation with TNFα. The canonical IKK complex was immunoprecipitated using an anti-NEMO antibody fused to Protein G Sepharose beads. Immunoprecipitated IKK2 was incubated with 1 μM of purified IκBα proteins. The reaction samples were incubated at 30° C and aliquots were taken at time 0, time 5, time 10 and time 30 (minutes). Reactions were migrated on a 10% SDS-PAGE gel and phosphorylation was detected using an anti P-S32/S36 antibody as described for panel A. 6xHis-IκBα proteins were detected using an anti-His antibody, whereas endogenous IKK2 was detected using an anti IKK2 antibody. (*Right panel*) Quantification of the phosphorylation levels of full-length IκBα wt, C308L and YD/SS constructs. The phospho-IκBα bands from the Western blotting analysis were quantified using ImageJ and normalized to the values of wild type IκBα at 30 min (100%). The data are averages of three independent experiments with error bars representing standard deviation values. **C)** Kinase assay using radiolabeled ATP. Purified recombinant 6xHis-IKK2 (1-669) EE (50 nM) was incubated with 5 μM of full-length IκBα wild type (wt), S32A/S36A (SS/AA), Y305A/D307A (YD/AA) and S32A/S36A/Y305A/D307A (SS/AA-YD/AA) mutants constructs in the presence of γ-$^{32}$P-ATP. Samples were incubated at 30° C for 2 hours and migrated on a 10% SDS-PAGE gel and phosphorylation levels detected using a phosphoimager.

activity experiments in the presence of an excess of synthetic IκBα (297-317) peptides. These experiments were performed together with Changqing Li using endogenous IKK complex and low concentrations of full-length purified 6xHis-IκBα protein substrate. Results indicate that both the wild type and C308L IκBα (297-317) peptides inhibit phosphorylation of IκBα, whereas the IκBα (297-317) peptide containing the suppressive Y305S/D307S mutation is less efficient in inhibiting kinase activity (Figure 5B).

Since IKK2 phosphorylation of IκBα leads to its degradation and the downstream triggering of the NF-κB transcriptional activity, we collaborated with Professor Alan Chariot (University of Liège, Belgium) to assess degradation of wild type and docking site mutant IκBα proteins. In these experiments, wild type IκBα, and the Y305S/Y307S and C308F (this latter mutation was better expressed in MEF cells compared to the C308L mutant) mutant constructs were transduced in mouse embryonic fibroblasts (MEFs) lacking the IκB family genes (MEF $^{IκBα-, IκBβ-, IκBε-}$). Upon activation of the canonical NF-κB pathway by TNFα, the Y305S/Y307S mutant is degraded at a lower extent compared to wild type IκBα. By contrast, degradation is enhanced for C308F IκBα (Figure 6).

Taken together the data show that the affinity of the docking interaction correlates with both phosphorylation and degradation of the IκBα substrate *in vitro* and *in cellulo*. In addition, our results show that a synthetic peptide containing the YDDΦXΦ docking site sequence can inhibit IKK2 catalytic activity.

### 5.1.3 Structural studies on the IKK2/IκBα peptide complex

To obtain structural information on the YDDΦXΦ motif interaction, we carried out several crystallization screenings on both *human* IKK1 and IKK2 bound to the synthetic IκBα (297-317) peptide. More precisely, I carried out crystallization of the complex consisting of IKK2 (1-669) EE and the wild type IκBα (297-317) peptide. The crystallization work was performed with the help of Pierre Poussin Courmontagne of the crystallization platform at IGBMC (Strasbourg). The 6xHis-IKK2 (1-669) EE construct was purified by $Ni^{2+}$-NTA affinity chromatography and then separated by the 6His tag by cleavage using the TEV protease. Subsequently, IKK2 samples were incubated with the specific MLN120B inhibitor to stabilize the conformation of the kinase domain and further purified by preparative gel filtration chromatography. The concentrated IKK2 sample was then mixed to the IκBα peptide at a stoichiometric ratio of 1:1.5 and subject to crystallization screenings. Several crystals grew in conditions based on ammonium sulphate. After optimization of the crystallization conditions (see Materials and Methods), we could identify one crystal which diffracted anisotropically to 4.2 Å (6.8 Å in the worst direction) using synchrotron radiation. Our collaborator

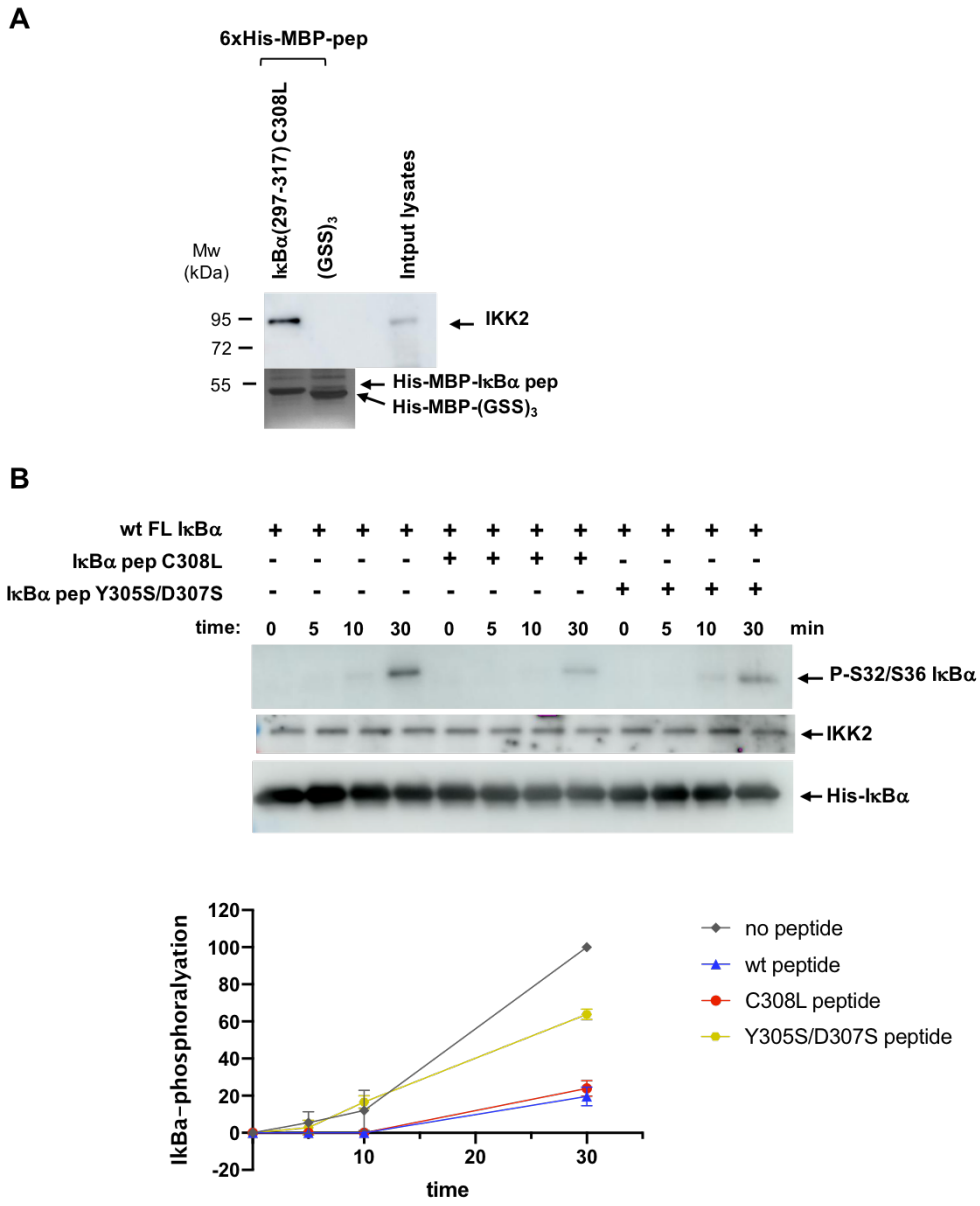



**Figure 5.** Inhibition assay of endogenous IKK2 activity by IκBα synthetic peptides **A)** Pulldown analysis of the interaction between the IκBα(297-317) C308L peptide and endogenous IKK2. The IκBα (297-317) C308L peptide was fused to a double 6xHis-MBP tag was expressed in *E. coli* and purified by Ni2+-NTA followed by gel filtration chromatography (on a Superdex 200 column). The peptide fusion was then coupled to amylose resin and incubated with clarified extracts of unstimulated HEK cells clarified supernatant. Pulldown samples were migrated on two separate 10% SDS-PAGE gels. Endogenous IKK2 was detected by Western blotting using an antibody anti IKK2 whereas IκBα by silver staining. A 6xHis-MBP fusion to a GSS rich peptide was used a negative control. **B)** Kinase activity experiments using endogenous IKK kinase immunoprecipitated from stimulated HEK293T cells, purified 6His-IκBα protein and synthetic peptides encompassing the YDDΦXΦ docking site of IκBα (residues 297-317). The peptides tested included the wt IκBα sequence and the C308L and Y305S/D307S mutations (IκBα pep C308L and IκBα pep Y305S/D307S). A 100-fold excess of synthetic peptide with respect of the IκBα substrate concentration (1 μM) was used in each reaction. (*Upper panel*) Representative Western blot images of the kinase activity experiments. Reactions were migrated on a 10% SDS-PAGE gel and phosphorylated and total IκBα proteins were detected using anti P-S32/S36-IκBα and anti-His antibodies, respectively. IKK2 was detected using an anti-IKK2 antibody. (*Lower panel*) Quantification of the phosphorylation levels of 6xHis-IκBα synthetic peptides. The phospho-IκBα bands from the Western blotting analysis were quantified using ImageJ. The phospho-IκBα values are normalized to the values obtained in the absence of peptide (no peptide) at 30 min incubation (100%). The data are averages of three independent experiments with error bars representing standard deviation values.

Alistair McEwen from the macromolecular crystallography platform at IGBMC used this dataset to solve the structure of the IKK2/IκBα peptide complex by molecular replacement, using a monomer extracted from the structure of the asymmetric dimer of human IKK2 (PDB: 4KIK) as a search model. In the resulting electron density map, the IKK2 homodimer is clearly visible together with density of the peptide binding at the interface between the two SDD domains of the IKK2 dimer. Due to the medium resolution of the X-ray dataset, the map lacks side chain information. This was problematic for the accurate fitting of the peptide in the electron density.

For a long time, we tried to obtain crystals with improved diffraction. My co-worker Changqing Li performed several crystallization trials of complexes of both IKK1 (the IKK1 (10-667) EE construct) and IKK2 (the IKK2 (1-669) EE construct) bound to either the wild-type or C308L mutant IκBα (297-317) peptide. Although several crystals were obtained, none displayed better diffraction than the initial crystal described above. For this reason, we opted cross-linking mass spectrometry (x-link ms) [317] to independently map the peptide binding site on IKK2. These experiments, which were performed by Changqing Li in collaboration with the group of Juri Rappsilber (Technisches Universität Berlin, Germany), used two cross-linking agents, namely EDC/sulfo-NHS and BS3. EDC sulfo-NHS is a short cross-linker (called zero length spacer), which conjugates aspartic acid or glutamic acid side-chains with lysine side chains or amino groups from the N-terminus of the peptide. Using EDC/sulfo-NHS we could identify by trypsin cleavage followed by mass spectrometry analysis several cross-links involving Glu300, Asp301 and Glu302 in the N-terminal region of the IκBα peptide and several lysine residues at the dimer interface of IKK2. Using BS3, which is a longer cross linker (10 Å spacer arm) that bridges primary amino groups from lysine residues or the N-terminus, we could get cross-links between the N-terminus of the peptide and the same lysine residues at the dimer interface of IKK2 identified with EDC/sulfo-NHS. These cross-linking data were translated into distance restraints and used together with the x-ray data to calculate a structural model for the IKK2/IκBα peptide complex using the Haddock software [26].

In the resulting 3D structure, the IκBα peptide adopts an extended conformation (Figure 7A). Whereas the cross-linking restraints hold in place residues 297-302 from the N-terminus of the peptide, residues 303-314 containing the YDDΦXΦ consensus are fitted to the electron density from the x-ray data (Figure 7B-C). Based on this structure, the YDDΦXΦ consensus lies in a pocket made by the two SDD domains of the interacting IKK2 subunits of the dimer (Figure 7D). This pocket is characterized by several conserved lysine and arginine residues, which provide an electrostatically favourable environment for the negatively charged residues of the YDDΦXΦ sequence. In addition, we can also observe several conserved methionine residues.

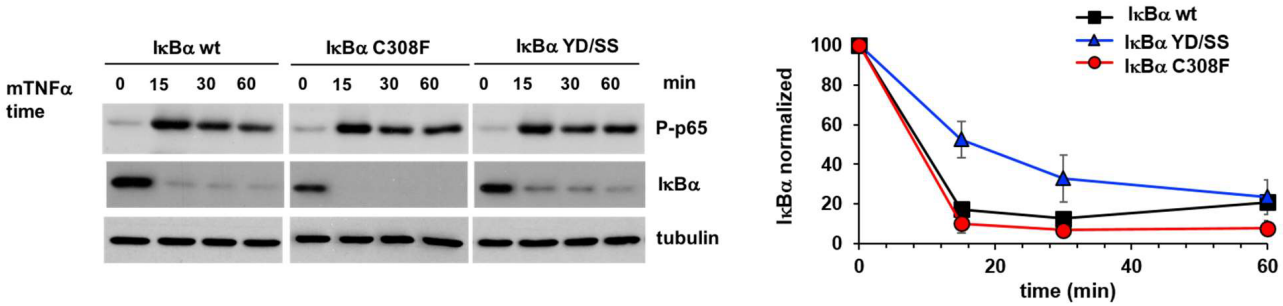



**Figure 6**. *In cellulo* degradation of IκBα. (*Left panel*) Representative Western blot images of the degradation experiment using MEF $^{IκBα-, IκBβ-, IκBε-}$ cells, which were retro-transfected with pBabe plasmids coding for full-length IκBαwild type (wt IκBα) or mutant (IκBαC308F and IκB α Y305S/D307S). MEF cells were stimulated with TNFα. Phosphorylated p65 (P-p65) was detected using phospho-specific anti-p65 antibody, whereas IκB α proteins were detected with an anti-IκBα antibody. Samples were taken at 0, 15, 30 and 60 minutes. (*Right panel*) Quantification of IκBα levels. IκB α bands were quantified by ImageJ. The data are normalized to tubulin and then to the IκBα values at time 0 (100%). The data are averages of three independent experiments with error bars representing standard deviation values.

To validate this model, we introduced site specific mutations at conserved residues within the YDDΦXΦ binding pocket and the dimer interface contacting the N-terminal region of the peptide (Figure 8A). In addition, we also included a previously characterized mutant (L654D/W655D) that prevents homodimerization of IKK2 [8]. The mutations were introduced in the G1-IKK2 full-length construct and the interaction with the G2-IκBα (73-317) construct measured by GPCA. These experiments were performed by Changqing Li.

Results show that the L654D/W655D mutation strongly reduced binding to IκBα, showing that an intact homodimer of IKK2 is required for the interaction (Figure 8B-C). The R460E, K469E, K480E and M475A mutations in the YDDΦXΦ binding pocket have smaller but significant effects. However, the combination of these substitutions in a single R460E/K469E/K480E triple mutant shows a higher impact on binding (Figure 8B-C). In addition, the K641E, R645E, K659E and K664E mutations targeting the region contacting the N-terminus of the peptide also displayed mild but significant effects.

**5.1.4 Identification of other substrates interacting with IKK1 or IKK2 via the YDDΦXΦ motif.**

Bioinformatic analyses detected the YDDΦXΦ consensus also in other substrates of IKK2, namely IκBβ from the IκB family, and of IKK1, namely the non-canonical p100 substrate and IRF7 [318]. This finding suggests that YDDΦXΦ is a novel short linear motif conserved in substrates of the IKK kinase (Figure 9A). Hence, I designed peptides containing the YDDΦXΦ motif from IκBβ (315-340), p100 (6-29) and IRF7 (482-500) substrates and tested their interactions with recombinant IKK1 and IKK2 proteins by pulldown analysis. Peptides were C-terminally fused to the MBP tag and expressed in *E. coli*, then coupled to amylose resin and incubated with recombinant purified IKK1(1-667) EE or IKK2(1-669) EE constructs. The results show that IKK2 recognizes only the motifs of IκBα and IκBβ, while IKK1 recognizes the motifs from all the substrates analyzed (IκBα, IκBβ, p100 and IRF7) (Figure 9B). Then, we measured by ITC the affinities of the interactions between IKK1 and the YDDΦXΦ peptides from p100 and IRF7 (p100 (6-29) pep and IRF7 (482-500) pep). These experiments were performed using purified recombinant IKK1 (10-667) EE and synthetic peptides. The results show that the equilibrium binding affinity constant for the IKK1/p100 pep interaction is very similar to that of the IKK1/IκBα pep interaction ($K_D$=11 μM) (Table 1). By contrast, the IKK1/IRF7 pep interaction displayed a much weaker affinity ($K_D$=90 μM).

p100 consists of an N-terminal disordered region that harbors the YDDΦXΦ motif (residues 15-20), followed by a Rel homology domain (RHD, residues 38-343), an ankyrin repeat domain (AR, residues 487-758) and a death domain (DD, residues 764-851). In the alternative NF-κB pathway,

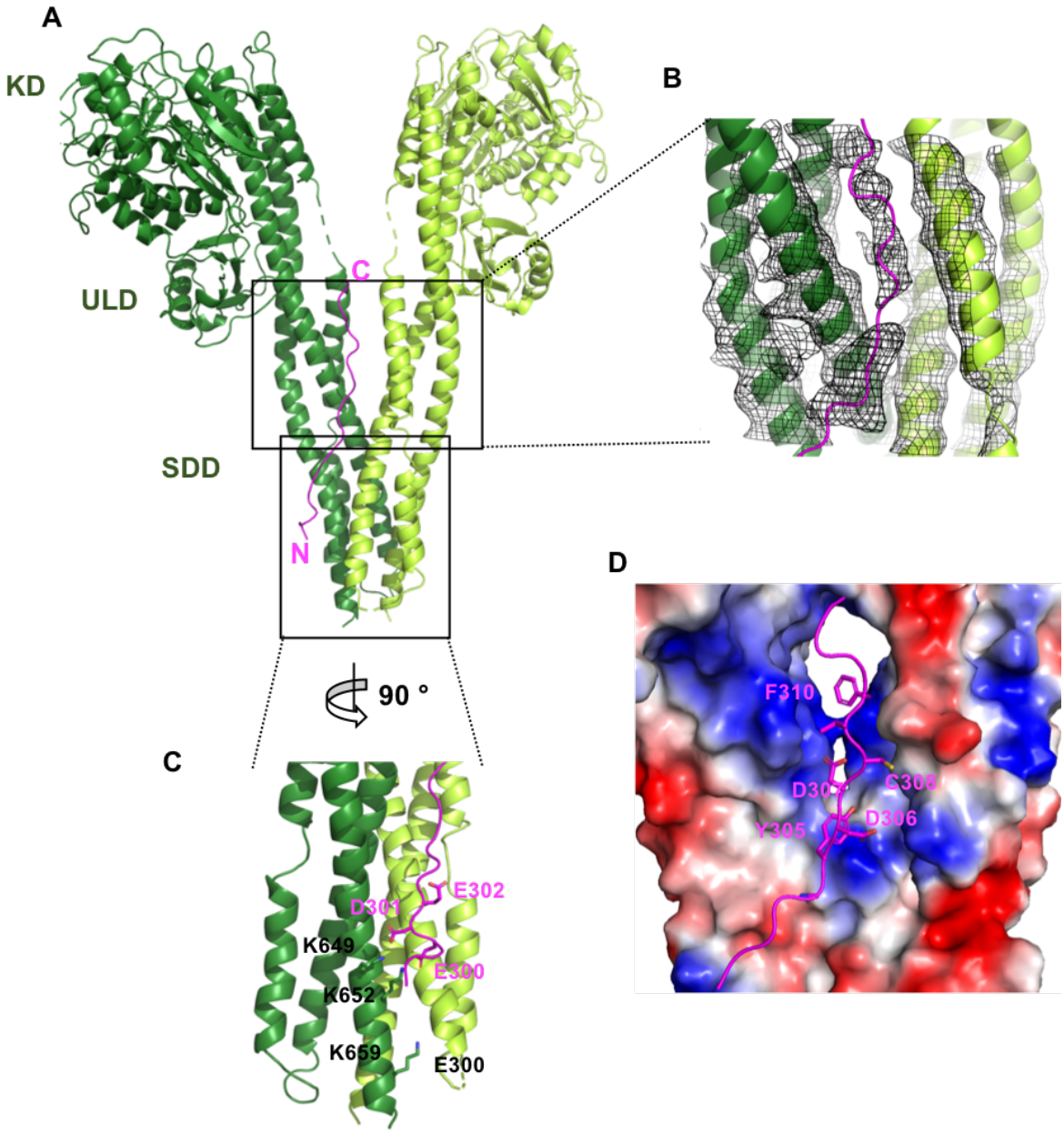



**Figure 7.** Structure of the IKK2/IκBα peptide complex. **A)** Ribbon representation of the structural model. The two subunits of the IKK2 dimer are shown in dark and light green. The IκBα peptide is shown in magenta. The kinase (KD), ubiquitin-like (ULD) and scaffold dimerization (SDD) domains of IKK2 are indicated. The structural model was calculated from the X-ray (for the IKK2 homodimer and YDDΦXΦ motif region of the IκBα peptide) and x-link ms data (N-ter region of the IκBα peptide) using the Haddock software. Boxes show position of the enlarged views shown in panels B and C. **B)** Enlarged view of the YDDΦXΦ binding pocket showing the electron density map calculated from the X-ray data. **C)** Enlarged view of the dimer interface contacting the N-ter region of the peptide. The side chains of IKK2 and IκBα residues involved in the x-links catalysed by EDC/sulfo-NHS are displayed. **D)** View of the YDDΦXΦ binding pocket. IKK2 is shown in the surface representation and coloured according to the electrostatic potential (blue: positive; red: negative). The IκBα peptide is shown in magenta. Sidechains of residues belonging to the YDDΦXΦ consensus are displayed.

IKK1 phosphorylates p100 at specific serine residues (Ser99, Ser108, Ser115, Ser123, Ser866, Ser870 and Ser872) inducing its poly-ubiquitination and subsequent partial proteolysis *via* the proteasome, thereby unmasking the NLS sequence of p100 [319] (Figure 10A, *left panel*).

We investigated the interaction between full-length p100 and IKK1 *in cellulo* by the GPCA approach. IKK1 and p100 were fused to the G1 and G2 fragments of the luciferase, respectively. We created three constructs of p100, namely the full-length protein, the 1-852 construct lacking the C-terminal disordered region hosting the Ser866, Ser870 and Ser872 phosphorylation sites, and the 39-852 construct lacking both the N-terminal region hosting the YDDΦXΦ motif and the C-terminal region. In addition, we generated a mutant of full-length p100 with a double substitution in the YDDΦXΦ motif (FL p100 Y15S/D17S). Binary combinations of G1-IKK1 and G2-p100 constructs were co-transfected in HEK293T cells. Results from the GPCA analyses show that the deletion or the suppression of the motif by double Y15S/D17S mutation strongly reduces the interaction between IKK1 and p100 (Figure 10B-C, *left panels*). By contrast, deletion of the C-terminal region containing the phosphorylation sites did not have any significant effect on the interaction (Figure 10B, *left panel*).

A similar analysis was performed for the IRF7-IKK1 interaction. IRF7 consists of a DNA binding domain (DBD, residues 10-127), followed by the IRF association domain (IAD, 237-423) and a disordered C-terminal region containing the YDDΦXΦ motif (residues 489-494) (Figure 10A, *right panel*). IRF7 is targeted and phosphorylated by IKK1, inducing its activation. To date, there is no precise information about the residues that are phosphorylated by IKK1 [320]. Here, G1-IKK1 was co-transfected with G2-IRF7 full-length wild-type or containing the double Y489S/D491S substitution in the YDDΦXΦ motif. Results from the GPCA experiments show that the suppression of the YDDΦXΦ motif does not have a significant impact on IKK1 binding IKK1, suggesting that additional mechanisms exist for this interaction (Figure 10B-C, *right panels*).

IRF7 activation in the interferon pathway depends on phosphorylation by the TBK1 kinase, which induces IRF7 dimerization and nuclear translocation [321]. TBK1 targets a cluster of serine/threonine residues within the C-terminal region of IRF7 (Thr465, Ser483, Ser484 [322], Ser477 and Ser479 [323]), which is closely upstream of the YDDΦXΦ motif [324] (Figures 9A-B). Interestingly, the YDDΦXΦ motifs from IκBα, IκBβ and p100 are preceded by a cluster of negatively charged residues that is missing in IRF7 (Figure 7A). In the case of IκBα this cluster corresponds to residues E300, D301 and E302, which contact the dimerization region of the IKK2 dimer (Figure 7C) and provide some small contributions to the interaction. We reasoned that in IRF7 these negatively charged residues are replaced by phospho-serines deriving from TBK1 phosphorylation, which would increase the affinity to IKK1 (Figure 11A-B).

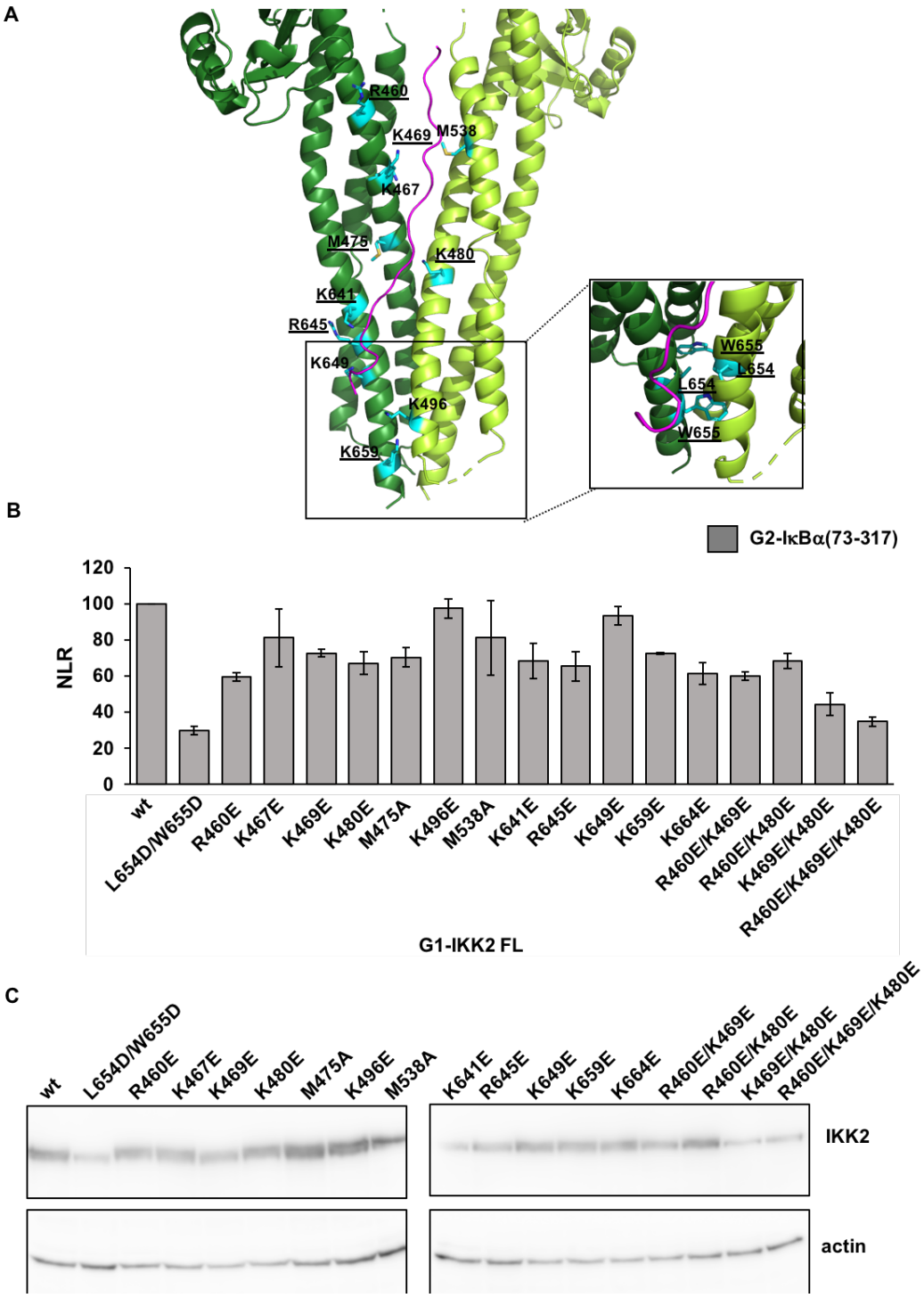



**Figure 8.** Binding studies using mutants of IKK2 residues at the interface with the IκBα peptide. **A)** Mutated residues displayed on the structural model of the complex. (*Right panel*) Mutated residues at the interface with the peptide are shown in cyan. Mutations of underlined residues have a significant impact on IκBα binding (see panel B). (*Left panel*) Enlarged view of the dimer interface showing the position of residues Leu654 and Trp655 that are crucial for IKK2 dimerization. **B)** GPCA analysis showing the interactions of the different IKK2 mutants with IκBα (construct 73-317). Mutations were introduced in the context of the full-length IKK2 protein. The data are averages of three independent experiments with error bars representing standard deviation values. The data were normalized to the NLR values obtained with wild-type IKK2 (100%). **C)** Expression levels of the G1-IKK2 FL mutant constructs in HEK293T cells.

Hence, I produced peptides comprising residues 469 to 503 of IRF7, which include the YDDΦXΦ motif, and containing phosphomimic mutations at the sites of TBK1 phosphorylation (Figure 11B). Two peptides contained two double Ser to Asp replacements (IRF7 pep S477D/S477D and IRF7pep S483D/S484D) and one peptide contained a single substitution (IRF7 pep S487D). These peptides were produced in *E. coli* as fusions to MBP and showed extensive proteolysis despite the addition of different protease inhibitors. Wild type and mutant MBP-IRF7 (469-503) peptides were coupled to amylose resin and incubated with the purified 6xHis-IKK1 (10-667) EE construct. A GSS rich peptide was used as negative control. The results of the pulldown analysis show that the S483D/S484D mutation significantly increases binding to IKK1 (Figure 11C).

Then, we inserted phosphomimic mutations in the context of full-length IRF7. We created the 5S/5D mutant bearing five Ser to Asp substitutions (S477D/S477D/S483D/S484D/S487D) and the 5S/5D YD/SS mutant that additionally included the double suppressive Y489S/D491S mutation in the YDDΦXΦ motif. GPCA analyses showed that neither of these two mutants showed significantly different binding responses to full-length IKK1. These results suggest that the YDDΦXΦ docking site in full-length IRF7 may not be available for interaction with IKK1 (Figure 11D).

Taken together, the results from this section show that the interactions of IKK1 and IKK2 with some of their substrates of are mediated by a novel YDDΦXΦ docking motif. The data on the IKK1/IRF7 interaction show that the affinity for the motif might be modulated by phosphorylation. However, in the context of the full-length IRF7, additional mechanisms are also likely to regulate the availability of the motif for interaction with IKK1.

## 5.2 Discussion

To acquire knowledge on IKK-substrate interactions, we initially focused on the IKK2 subunit and the IκBα substrate. In fact, Xu et al. [8] had previously predicted the existence of a docking site for IKK2 somewhere within the region comprising the ankyrin repeat domain and the C-terminal tail of IκBα. By performing pulldown experiments using a series of deletion constructs of IκBα and purified recombinant IKK2, we identified the minimal binding region of IKK2 within IκBα. This region, which corresponds to a sequence of 20 amino acids located in the C-terminal disordered region of IκBα, is necessary and sufficient for the interaction. Serine mutagenesis of all positions of this sequence allowed us to identify the residues critical for the interaction, which are part of a YDDΦXΦ consensus that is conserved in IκBα orthologue proteins. Interestingly, the IKK1 catalytic subunit also binds to IκBα *via* the YDDΦXΦ motif. The requirement of the YDDΦXΦ motif for the

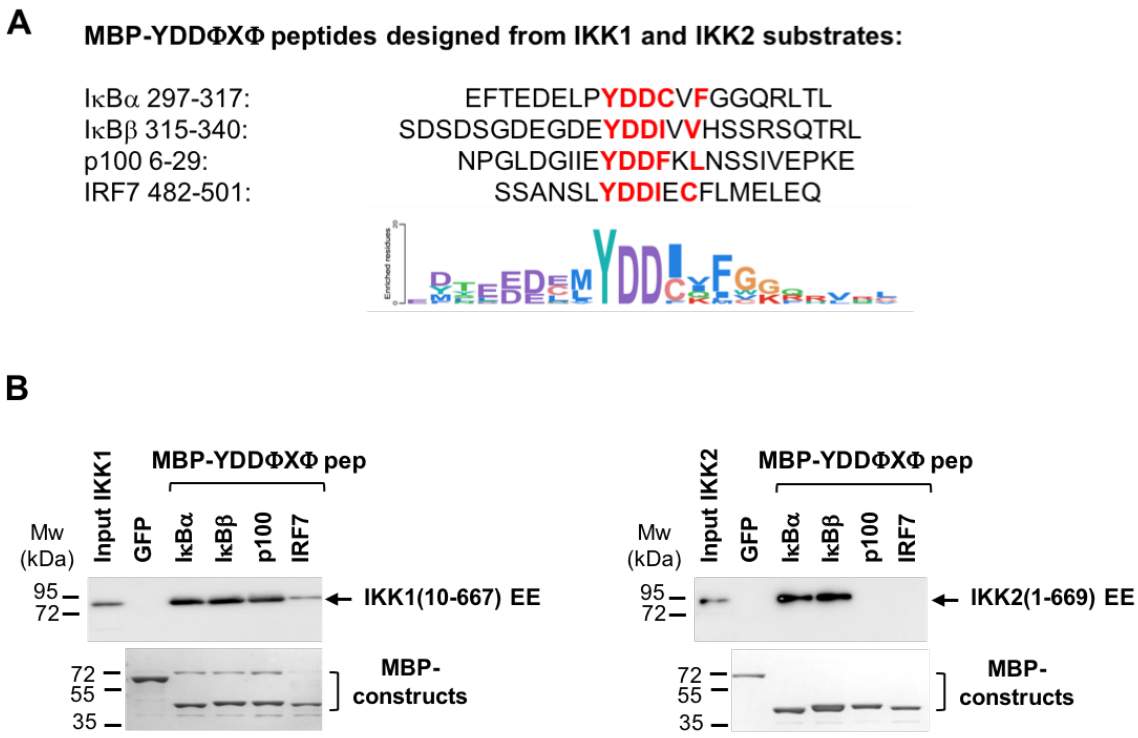



**Figure 9.** Characterization of interactions of IKK1 and IKK2 with YDDΦXΦ peptides from different substrates. **A)** (*Upper panel*) YDDΦXΦ peptides from IκBα, IκBβ, p100 and IRF7 used in the experiments of panel B. (*Lower panel*) Sequence logo calculated from the YDDΦXΦ peptides from IκBα, IκBβ, p100 and IRF7 showing the position-specific frequency of each amino acid composing the motif. **B)** Pulldown analysis of the interactions between IKK1 and IKK2 and the MBP-YDDΦXΦ peptides. Experiments were performed using purified recombinant 6xHis-IKK1 (10-667) EE or 6xHis-IKK2 (1-669) EE and YDDΦXΦ peptides fused to the MBP tag and coupled to amylose resin. Samples were migrated on 2 separate 10% SDS-PAGE gels. IKK1 (*left panel*) and IKK2 (*right panel*) were detected by Western blotting using an anti-His antibody, whereas the MBP-constructs were detected by Coomassie staining. MBP-GFP was used as negative control.

interactions between IKK1 or IKK2 and IκBα was confirmed by results from *in cellulo* experiments (GPCA), which showed that the deletion of the region containing the motif or the suppression of the motif by site directed mutagenesis drastically decreased the interaction levels.

The function of the IKK1 subunit of the canonical complex was thought to be the activation of IKK2 by trans-phosphorylation within the catalytic loop. Here we show for the first time that IKK1 also contributes to binding to IκB substrates. This finding is in agreement with the high sequence homology between the IKK catalytic subunits (70%) and with the results of our structural studies. Indeed, these results show that the YDDΦXΦ motif of IκBα interacts within a pocket formed by the two adjacent SDD domains of the IKK dimer, which in the case of the canonical complex, are provided by the IKK1 and IKK2 subunits.

Data from our ITC experiments show that IKK1 binds to the YDDΦXΦ peptide of IκBα with a slightly higher binding compared to IKK2 ($K_D$=10μM versus $K_D$=30 μM). Moreover, our mutagenesis analysis showed that IKK1 tolerates a larger number of substitutions within the YDDΦXΦ motif than IKK2. In agreement with this finding, we found that IKK1 is able to recognize the YDDΦXΦ motifs also of other proteins, that is p100 and IRF7, which are also its substrates. These results seem to suggest that IKK1 has a broader specificity than IKK2 for the motif. Indeed, our hypothesis is that YDDΦXΦ could be a variant of a more general motif. Proteomics analyses such as phage display should be carried out to identify this more general motif.

Most of the IKK1 and IKK2 interactions with YDDΦXΦ peptides have equilibrium binding affinities in the low micromolar range. These $K_D$ values are characteristic of transient interactions between the enzyme and the substrate, which is an essential feature for those kinases implicated in signaling pathways that govern cell metabolism and cell behavior. Such low affinities are combined with the low catalytic efficiencies of kinases compared to other enzymes. While low affinities impose the achievement of high substrate concentrations, preventing unwanted catalytic reactions, which could evoke a cell response in absence or in presence of weak of stimuli, a low catalytic efficiency allows the kinase to indirectly gain high specificity, thus becoming able to discriminate among apparently similar substrates. These features make that kinases need several mechanisms of regulation. In this work we observed that the truncation of the N-terminal region of IκBα enhances the interaction with both IKK1 and IKK2. This result suggests the existence of a regulatory mechanism involving the IκBα substrate. It is possible that, in the context of full-length IκBα, the N-terminus interacts with the C-terminal region, masking the docking site or sterically preventing the binding to IKK1/2. However, the interaction of IκBα with the NF-κB dimers would induce the displacement of the N-terminus unmasking the docking site, thus priming IκBα for efficient binding and subsequent

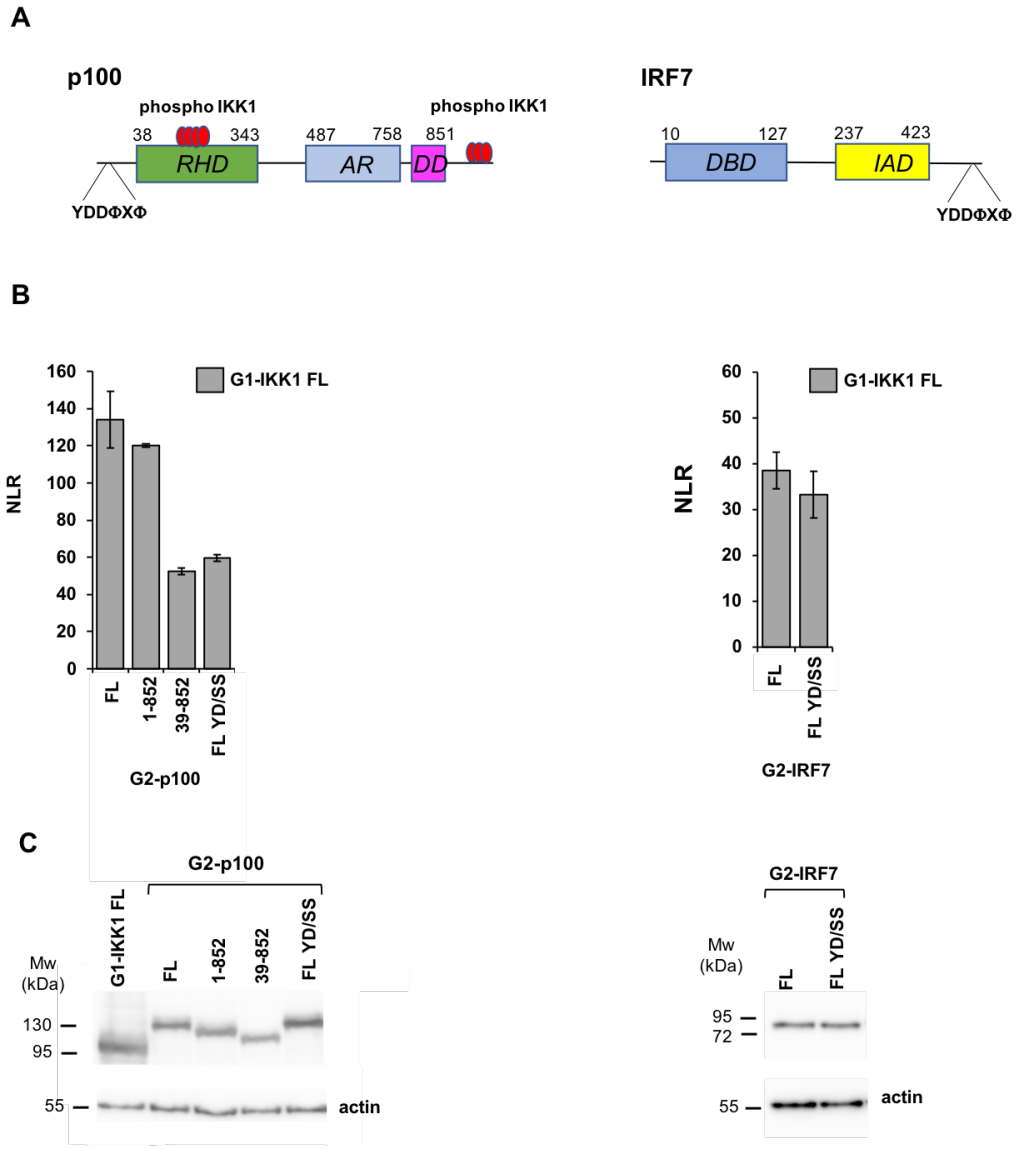



**Figure 10.** *In cellulo* characterization of the interactions of IKK1 with p100 or IRF7. **A)** (*Left panel*) Schematic representation of the domain architecture of the p100 substrate. The N-terminal disordered region contains the YDDΦXΦ motif (15-20), followed by the Rel homology domain (RHD, 38-343), the ankyrin repeat domain (AR, 487-758) and the death domain (DD, 764-851). The AR domain and the C-terminus region contain the phospho-acceptor serine residues targeted by IKK1 indicated by red ovals (S99, S108, S115, S123, S866, S870 and S872). (*Right panel*) Schematic representation of the domain architecture of IRF7 consisting of a DNA binding domain (DBD, 10-127), followed by the IRF association domain (IAD, 237-423) and a disordered C-terminal region containing the YDDΦXΦ motif (489-494). **B)** GPCA analysis of the interactions of full-length IKK1 (FL IKK1) with p100 constructs and full-length IRF7. IKK1 was fused to the Gluc1 (G1) fragment of the luciferase, whereas p100 constructs were fused to Gluc2 (G2) luciferase fragment. (*Left panel*) HEK293T cells were co-transfected with G1-IKK1 and G2-p100 plasmids. The p100 constructs tested include the full-length protein, the 1-852 construct lacking the C-terminal disordered region and the 39-852 lacking both N- and C-terminal disordered regions. YD/SS corresponds to the double Y15S/D17S mutation in the YDDΦXΦ motif. (*Right panel*) HEK293T cells were co-transfected with G1-IKK1 and G2-IRF7 plasmids. The YD/SS mutation corresponds to the double Y502S/D504S mutation in the YDDΦXΦ motif. **C)** Expression levels of G1-IKK1 and G2-p100 (*left panel*) and G2-IRF7 (*right panel*) constructs in HEK293T cells. G1-IKK1 and G2 fusions were detected by Western blotting using an antibody recognizing the *Gaussia princeps* luciferase.

phosphorylation by active the IKK2. In other words, this regulatory mechanism would ensure that only IκBα bound to NF-κB heterodimers is targeted by IKK for phosphorylation and degradation [325]. IKK2 targets two specific serines (Ser32 and Ser36) located in the N-terminal region of IκBα, whereas the docking site is found in the C-terminal region. To analyze the impact that the docking interaction exerts on the kinetic activity of IKK2, we introduced mutations within the YDDΦXΦ motif that modulate the affinity of the interactions with IKK1 and IKK2. In this work, in addition to suppressive mutations, we also identified two mutations in the first hydrophobic position of the YDDΦXΦ motif (C308F and C308L), which increase the affinity of IκBα for both IKK1 and IKK2. The kinase activity experiments were performed using either recombinant purified IKK2 or endogenous IKK canonical complex immunoprecipitated form HEK293T cells extracts upon TNFα stimulation. Results show that the modification of the affinities has a significant impact on the kinase activity. The IκBα C308L mutant displayed the highest phosphorylation levels, whereas the suppressive IκBα YD/SS mutant resulted in lower phosphorylation compared to the wild type protein. In addition, we could also show that the suppression of the docking interaction does not redirect kinase activity to other sites of IκBα. These results are consistent with the findings of Xu *et al.* [8] who derived lower $V_{max}$ values for the IKK2 phosphorylation of the N-terminal region of IκBα (residues 1-54) compared to the full-length IκBα protein, indicating that the docking interaction enhances catalysis. Consistently, the results from the experiments performed by our collaborator Professor Alan Chariot (University of Liège, Belgium) show that the *in cellulo* degradation of the suppressive YD/SS IκBα mutant is slower than that of wild-type or C308F IκBα.

Given that the increased affinity, determined by the mutation of the YDDΦXΦ motif, correlates with a higher phosphorylation level *in vitro* and to a quicker degradation *in cellulo*, we wondered if we could use a peptide containing the YDDΦXΦ motif to inhibit the activity of the IKK kinase. As a matter of fact, several enzymatic inhibitors against IKK2 have been tested but most of them failed at the clinical trials due to important side effects. This happened because of the relatively high degree of conservation of the kinase domain, leading to the inhibition of off-target enzymes and also because of the many functions of the IKK kinase in the cell. In addition, a brand-new category of biological inhibitors was designed to target the interaction between IKK2 and NEMO. Small peptides derived from the IKK binding domain of NEMO have been tested with promising results. These peptides would compete for the NEMO binding domain of IKK2 with NEMO itself, blocking the assembly of the IKK canonical complex and consequently the signal transduction. Likewise, we thought that a peptide comprising the YDDΦXΦ motif could prevent the phosphorylation of IκBα by direct competition with the IKK at the docking site. Results show that the presence of an excess of the

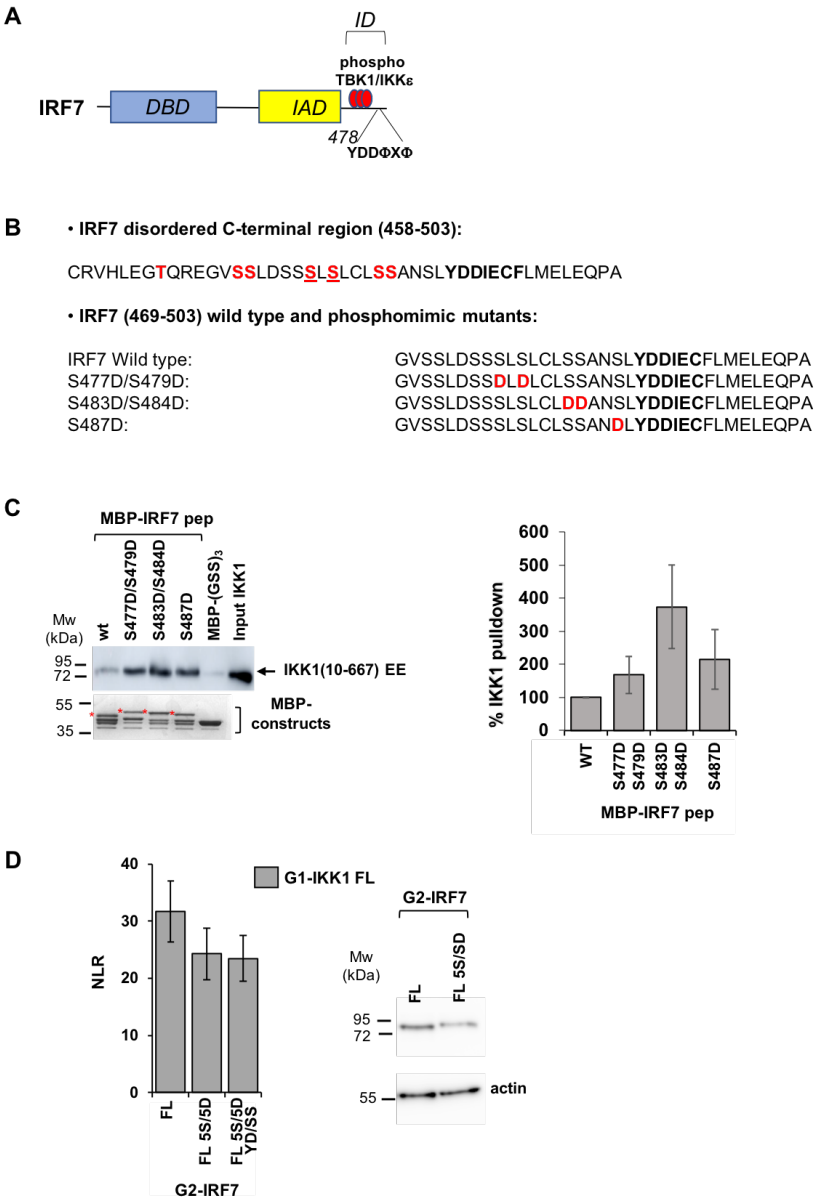



**Figure 11.** Effects of phosphorylation on recognition of the YDDΦXΦ motif of IRF7 by IKK1 **A)** Domain architecture of IRF7 indicating the phosphorylation cluster (red ovals) targeted by the TBK1 kinase and adjacent to the YDDΦXΦ motif within the C-terminal disordered region. **B)** (*Top panel*) Sequence of the IRF7 C-terminal region (residues 458-503). Residues targeted by TBK1 targets according Caillaud *et al.* [322] and tenOever *et al.* [323] are indicated in red. The YDDΦXΦ motif is also indicated (bold). (*Lower panel*) wild-type and phosphomimic IRF7 peptides IRF7 substrate, used for the pulldown experiments shown in panel B. Phosphomimic substitutions (D) are indicated in red. **B)** (*Left panel*) Representative image of a pulldown experiment between for the IKK1/IRF7 interactions. Experiments were performed using purified recombinant 6xHis-IKK1 (10-667) EE and IRF7 peptides fused to MBP and coupled to amylose resin. Samples were migrated on 2 separate 10% SDS-PAGE gels. IKK1 was detected by Western blotting using an anti-His antibody, whereas the MBP-constructs were detected by Coomassie staining. An MBP-$GSS_3$ peptide was used as negative control. Red asterisks indicate full-length MBP-IRF7 peptide fusions, whereas lower bands correspond to degradation products. *(Right panel)* Quantification of the interaction levels between purified recombinant 6xHis-IKK1 and MBP-IRF7 peptides. The IKK1 bands were quantified using ImageJ. The data were normalized to the IKK1/IRF7 wild type peptide interaction. The data are averages of three independent experiments with error bars representing standard deviation values. **D**) *In cellulo* interaction between IKK1 and phosphomimic mutants of IRF7. (*Left panel*) Results from GPCA analysis between the full-length (FL) IKK1 and IRF7 proteins, fused to the G1 and G2 fragments of the luciferase, respectively. The 5S/5D corresponds to the phosphomimic mutant S477D/S479D/S483D/S484D/S487D. The 5S/5D YD/SS mutant contains the five phosphomimic mutations plus the Y502S/D504S mutation in the YDDΦXΦ motif. The data are averages of three independent experiments with error bars representing standard deviation values. (*Right panel*) Expression levels of FL IRF7 wild-type and 5S/5D mutant fused to the G2 fragment in HEK293T cells. G2-IRF7 proteins were detected by Western blotting using an antibody recognizing the *Gaussia princeps* luciferase.

C308L IκBα (297-317) peptide decrease phosphorylation of full-length IκBα by the endogenous IKK complex. Therefore, these results, which are very promising, open the way to development of a new class of inhibitors of IKK. The next step will be to further improve the binding affinities of YDDΦXΦ peptides either by mutagenesis or by chemical modifications in order to obtain more efficient peptides.

In this work we could obtain, for the first time, structural information on the IKK-IκBα interaction. Here, we could not use a strategy based on deletion mutagenesis to define the region of IKK1 or IKK2 binding to the docking motif. In fact, the architecture of the IKK1 and IKK2 subunits is such that the KD, the ULD and the SDD domains are intimately connected and require each other for their correct folding. For this reason, in our structural studies we used the previously crystallized IKK1 (10-667) and IKK2 (1-669) constructs, which comprise all three domains and exclude the unfolded NBD region. All complexes tested gave crystals with poor resolution. Even complexes that involved the higher affinity C308L IκBα peptide mutant did not show improved resolution. We think that this poor resolution is related to the flexibility of the IKK homodimers that display different degrees of opening. The only crystal that reached a sufficient resolution for structure resolution was obtained with IKK2 and the IκBα wild type peptide. However, whereas the electron density of the IKK2 homodimer was easy to interpret (thanks to the high content of α-helices), the electron density of the peptide could not be interpreted due to the lack of side-chain information. For this reason, we performed x-linking ms experiments, which provided data that allowed us to determine the N-to-C-ter directionality of the peptide in the complex. In particular, the N-terminal region of the peptide involved in the x-links is not visible in the electron density map, suggesting that this part of the peptide is more flexible than the YDDΦXΦ region. In the resulting model structure calculated using X-ray and x-linking ms data, the IκBα peptide binds in an extended conformation to the two SDD domains of the IKK2 dimer. In this model the precise positioning of the side chains of the YDDΦXΦ motif is not precise. To improve the quality of the model and increase the details regarding the positioning of the side chains implicated in the interaction between the kinase and the substrate, we are currently repeat x-linking ms experiments using the sulfo-SDA agent [326], as UV photo-activatable cross-inker, which cross-links lysine side chains (and also tyrosine, serine and threonine residues in presence of a neighboring histidine [327]) with any other primary N-terminal amino group .

To validate the structural model, we mutated several conserved IKK2 residues at the interface with the IκBα peptide. Generally speaking, single point mutations had mild effects on the interaction with IκBα. These mild effects could be explained by the fact that the peptide binding pocket of IKK2 is largely dominated by electrostatic and polar interactions, with little hydrophobic interactions mostly contributed by conserved methionine residues. In addition, this pocket is rich in lysine residues that

are quite mobile and can adapt their conformation to compensate a mutation at a different residue. In fact, it seems that IKK2 provides several residues for the interaction with the peptide, but with each of these residues having small contributions. As a matter of fact, to obtain a strong decrease of the interaction, it was necessary to combine three mutations in the same construct (R460E/K469E/K480E). By contrast, a double point mutation that prevents IKK2 dimerization (L654D/W655D) had a much stronger effect on the interaction with IκBα. This result shows that IκBα binds only to the IKK2 homodimer form and not to an IKK2 monomer. To corroborate these findings, we are currently preparing recombinant IKK2 constructs carrying selected mutations at the binding interface that will be tested by *in vitro* pulldown assay with purified IκBα.

Regarding the role of the YDDΦXΦ motif in the interactions of IKK1 with substrates other than IκB proteins, our results show the deletion of the p100 N-terminal region containing the YDDΦXΦ motif or the suppression of the motif by double point mutation, suppressed binding IKK1, demonstrating that the motif is critical for the interaction. Concerning IRF7, the affinity for the interaction is low, with a $K_D$ in the high micromolar range. Consistently, the suppression of the YDDΦXΦ motif of IRF7 by a double mutation does not reduce the IKK1-IRF7 interaction in a significant way. We thought that the affinity YDDΦXΦ docking site of IRF7 for IKK1 could be increased by phosphorylation. In fact, a remarkable difference between the other proteins hosting the YDDΦXΦ motif (IκB and p100) and IRF7 lies in the region immediately upstream of the motif. Whereas the formers show an acidic patch, IRF7 lacks these negative charges. Moreover, IRF7 is an IRF transcription factor implicated in the INF signaling and besides IKK1, it is targeted by different kinases such as IKK-related kinases TBK1/IKKE, which phosphorylate a specific cluster immediately upstream of the YDDΦXΦ motif, thus inducing the dimerization and subsequent translocation into the nucleus. Our results from pulldown experiments using phosphomimic mutants do show that the phosphorylation of TBK1 target residues significantly increases the interaction between YDDΦXΦ peptides from IRF7 and IKK1. However, when the same phosphomimic mutations are introduced in the full-length IRF7 protein we could not observe any significant increase of binding to IRF7. Altogether, these results suggest that additional regulatory mechanisms exist for this interaction, that could implicate either the presence of additional docking sites within IRF7 or of conformations that prevent the accessibility to the YDDΦXΦ motif. In fact, the YDDΦXΦ motif is localized in the C-terminal mobile helix of the protein. Based on the structural data on the IRF3 homologue, this helix undergoes rearrangements upon IRF3 homodimerization following TBK1/IKKε activation. Therefore, further work will be needed to identify the precise mechanisms of the IKK1-IRF7 interaction.

Together, our results clearly show that the YDDΦXΦ motif mediates the interactions of the IKK1 and IKK2 subunits with their key substrates acting in NF-κB signaling (IκB proteins and p100). However, additional mechanisms are likely to exist for the interactions with other substrates participating in other signaling pathways.

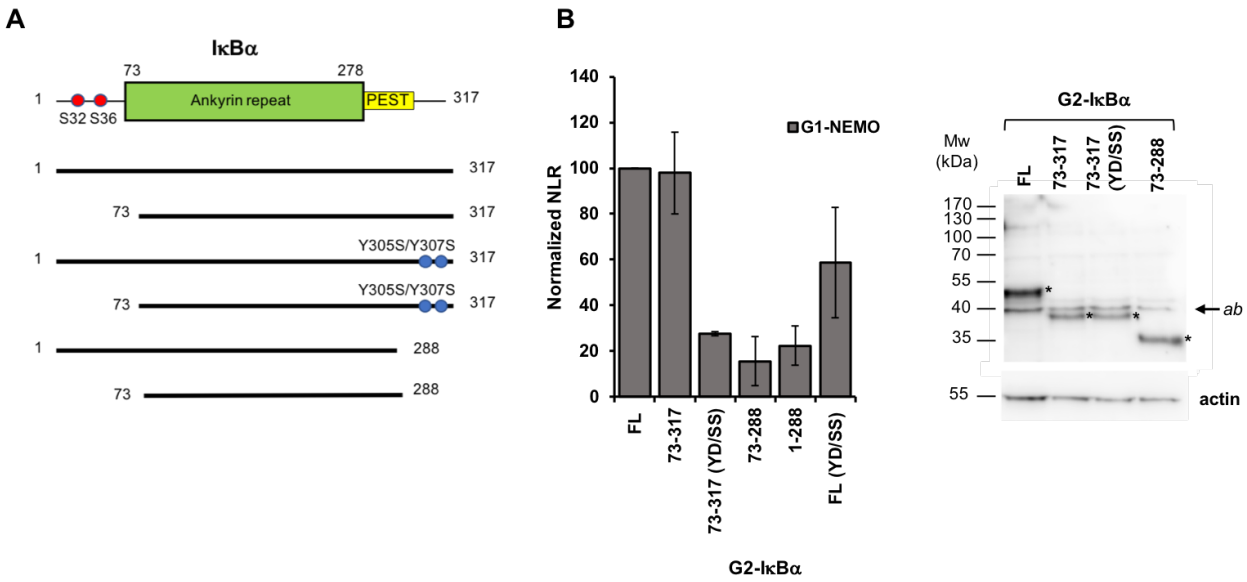



**Figure 1.** NEMO interaction with the C-terminal region of IκBα containing the YDDΦXΦ docking motif. **A**) G1-IκBα constructs used in GPCA experiments shown in panel B. **B**) GPCA analysis of the interactions between full-length NEMO (G1-NEMO) and G2-IκBα constructs. HEK293T cells were co-transfected with G1-NEMO and G2-IκBα vectors. The IκBα constructs tested included the full-length wild type and YD/SS mutant proteins (FL and FL(YD/SS), respectively), the 73-317 wild type and YD/SS mutant constructs, the 73-288 construct, the 1-288 construct. YD/SS corresponds to the double Y305S/D307S mutation suppressing the YDDΦXΦ docking site. **C**) Expression levels of G2-IκBα in HEK293T cells. G2- IκBα fusions were detected by Western blotting using an antibody recognizing the *Gaussia princeps* luciferase. IκBα constructs are indicated by an asterisk (*). Other aspecific bands are visible (ab).

# Chapter 6

# Characterization of the interaction between NEMO and IκBα

NEMO plays a pivotal role in the NF-κB canonical signaling. In fact, upon proper stimulation, NEMO ensures the correct targeting of IκBα by the IKK2 subunit of the IKK canonical complex. It appears that *in vivo* the intrinsic substrate specificity of IKK2 is improved and optimized by NEMO. Despite a wide body of evidences supporting the critical role of NEMO for correct NF-kB canonical signaling, the precise mechanism that NEMO utilizes for substrate recognition is still unclear. Hoffman and coworkers [27] propose that NEMO could act as a scaffold, recruiting the substrate to the IKK complex by direct binding. In this chapter I present preliminary results from *in cellulo* and *in vitro* interaction analyses that were performed to identify the minimal binding regions of NEMO on IκBα and *vice versa*.

## 6.1 Results

### 6.1.1 The NEMO binding region within IκBα involves the YDDΦXΦ motif

In order to identify the minimal binding region of NEMO within IκBα, we performed GPCA interaction experiments, in which IκBα deletion constructs or mutants N-terminally fused to the Gluc2 fragment of the luciferase (G2- IκBα) were tested against full-length NEMO N-terminally fused to the Gluc1 fragment (G1-NEMO). G2-IκBα and G1-NEMO constructs were co-expressed in HEK293T cells. IκBα FL WT was found to interact robustly with NEMO with elevated NLR values ranging between 100 and 200 depending on the experiment (Figure 1). Truncation of the N-terminus of IκBα (IκBα (73-317)) had no significative effect on binding. By contrast, truncation of the C-terminal portion containing the PEST region and the YDDΦXΦ motif (compare FL IκBα with IκBα (1-288) and IκBα (73-288)) strongly reduced the interaction. Also, the introduction of the Y305S/D307S mutation within the YDDΦXΦ motif in the context of the N-terminally truncated IκBα construct showed a decrease of the interaction (compare FL IκBα with IκBα (73-317) YD/SS). Interestingly, the same mutation introduced in the background of the full-length IκBα protein has a weaker impact (compare FL IκBα wt with FL IκBαYD/SS). These results indicate that IκBα docks to NEMO *via* the C-terminal region and that the N-terminus partially re-establishes the interaction. This is in line with the findings of Hoffman and collaborators [27], who identified a NEMO interaction

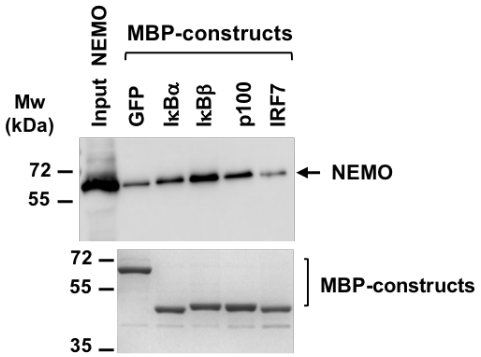



**Figure 2.** Pulldown experiment using purified MBP-YDDΦXΦ peptides and 3xFlag-NEMO overexpressed in HEK293T cells. Samples were migrated on 2 separate 10% SDS-PAGE gels. NEMO was detected by Western blotting using an anti-Flag antibody, whereas the MBP-YDDΦXΦ peptide constructs were detected by Coomassie staining. MBP-GFP was used as negative control.

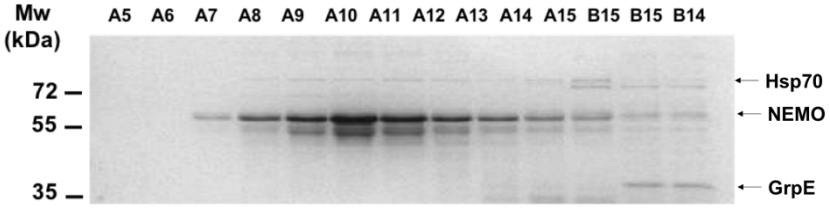



**Figure 3.** Analysis of the purification of 6xHis-NEMO by $Ni^{2+}$-NTA affinity chromatography and gel filtration chromatography. SDS-PAGE analysis (10% gel) of the elution fractions from the gel filtration column (Superdex 200 10/30) covering the elution volume from 6 ml to 13 ml (corresponding to A5 to B13 fractions). NEMO co-purified with the chaperons Hsp70 and GrpE.

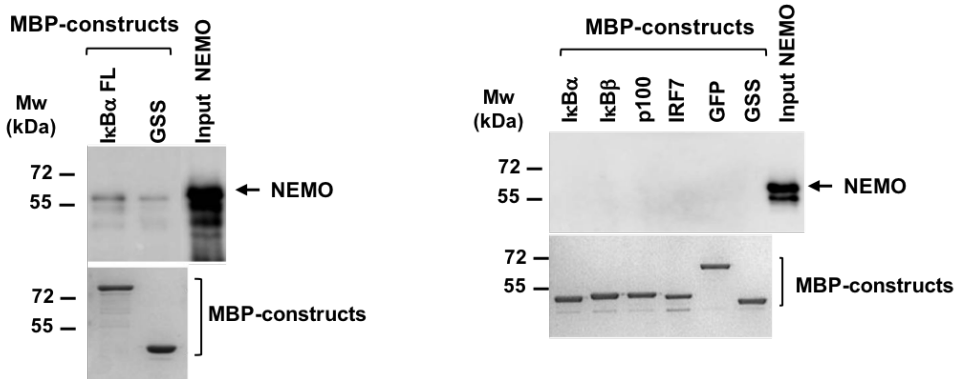



**Figure 4**. Pulldown experiments using recombinant purified 6His-NEMO and MBP-IκBα (*left panel*) or MBP-YDDΦXΦ peptides (*right panel*). Samples were migrated on 2 separate 10% SDS-PAGE gels. NEMO was detected by Western blotting using an anti-His antibody, whereas the MBP-IκBα and MBP-YDDΦXΦ peptide constructs were detected by Coomassie staining. MBP-GFP and MBP fused to a glycine serine rich peptide (GSS) was used as negative controls.

site within the N-terminal region of IκBα, with two aspartic acid residues (Asp27 and Asp28) being critical for the binding.

Since the GPCA assay is not very well suited for protein-peptide interactions, I over-expressed 3xFlag-NEMO full-length in HEK293T cells and tested its interactions with the YDDΦXΦ peptides from IκBα, IκBβ, p100 and IRF7 that were previously used for the interactions with IKK1 and IKK2 (see Chapter 5, Figure 9). The clarified lysate of NEMO expressing cells was incubated with MBP-YDDΦXΦ peptides (IκBα (301-317), IκBβ (315-340), p100 (6-29) and IRF7 (482-500)) coupled to amylose resin. In this experiment, NEMO displays higher binding for the IκBα, IκBβ and p100 peptides as compared to the IRF7 peptide and GFP (Figure 2). However, a high background is clearly shown by the non-specific interaction with the MBP-GFP. The same interaction pattern was reproduced in three independent experiments.

To further evaluate the role of the YDDΦXΦ motif in the interaction between NEMO and IκBα, we decided to perform binding experiments using recombinant purified full-length NEMO. Initially, I produced 6xHis-NEMO in the *E. coli.* DE3 strain and purified it by $Ni^{2+}$-NTA affinity chromatography followed by gel filtration chromatography. Using this protocol, we obtained several truncated forms of NEMO that were related to the numerous rare codons present within the NEMO sequence. So, we chose to express NEMO using the *E. coli* Rosetta strain, optimized for rare codon-containing sequences, in LB medium supplemented with zinc sulfate for correct folding of the C-terminal zinc finger domain. In this way, we could remove in good part the truncated products (Figure 3). However, we noticed that NEMO co-purifies with chaperons, possibly Hsp70 and GrpE. The presence of different chaperons could be explained by the structural features of NEMO, which displays a high degree of disorder.

Next, we tested the interactions of recombinant purified NEMO with full-length IκBα and YDDΦXΦ peptides from IκBα, IκBβ, p100 and IRF7 by pulldown. MBP- IκBα and MBP-peptides were incubated with NEMO adjusted at a concentration of 10 μM. Results show very weak binding with full-length IκBα and some tendency of NEMO to bind non-specifically to the negative control (Figure 4, *left panel*) No interaction between NEMO and any of the YDDΦXΦ peptides was observed (Figure 4, *right panel*).

Hence, we wondered whether the use of an MBP tag could prevent the interaction of the constructs tested with NEMO. Hoffmann and collaborators [27] had shown an interaction between recombinant full-length IκBα and NEMO proteins using gel filtration chromatography. I therefore repeated this same experiment. The two 6xHis-IκBα and 6xHis-NEMO proteins were expressed in bacteria and purified separately by $Ni^{2+}$-NTA affinity and gel filtration chromatography. Then, the proteins were

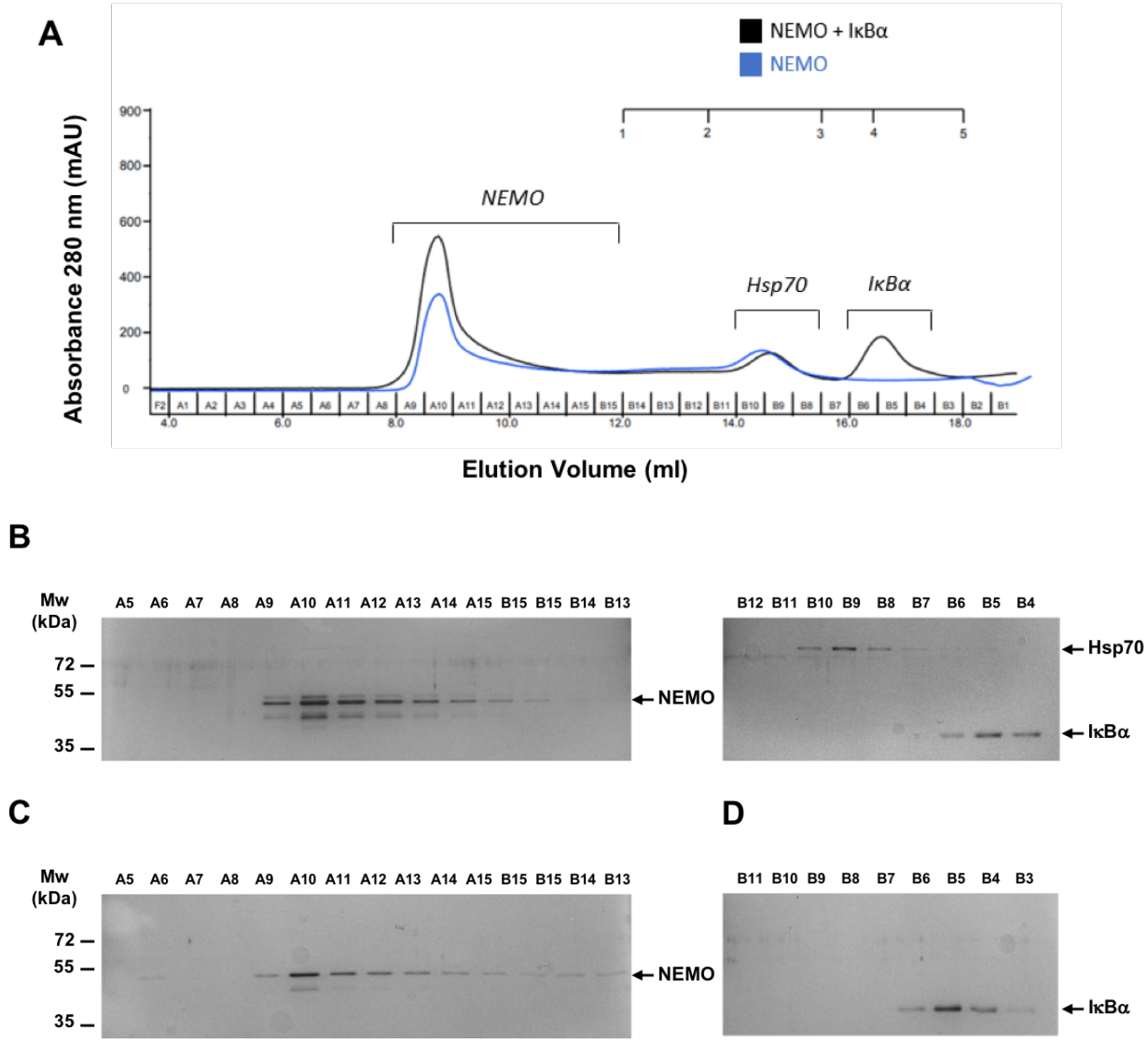



**Figure 5.** Interaction studies between 6xHis-NEMO and 6xHis-IκBα by gel filtration chromatography. **A)** Overlay of chromatograms of the NEMO/IκBα 1:1 mixture (black) and of NEMO alone (blue) obtained using a Superose 6 10/300 column. Size markers: 1 - Thyroglobulin (669 kDa); 2 - Ferritin (440 kDa); 3 - Aldolase (158 kDa); 4 - Ovalbumin (44 kDa); 5 - Ribonuclease A (13.7 kDa); 6 - Aprotinin (6.5 kDa). **B)** 10% SDS-PAGE gel analysis followed by silver nitrate staining of the fractions between 6 and 17.5 ml of the elution volume (A5 to B4) for the NEMO/IκBα mix. **C)** 10% SDS-PAGE gel analysis followed by silver staining of the 6 to 13 ml elution fractions (A5 to B4) for NEMO alone. **D)** 10% SDS-PAGE gel analysis followed by silver staining of the 13.5 to 18 ml fractions (B11 to B3) for IκBα alone.

mixed at a 1:1 stoichiometric ratio at concentrations of 10 μM, incubated on ice and loaded onto the gel filtration column. SDS-PAGE analysis of the elution fractions shows that the IκBα and NEMO proteins elute in separate fractions (Figure 5). This suggest that the two proteins either do not interact or their interaction is not strong enough to persist during the migration through the column.

To conclude, results from *in cellulo* GPCA experiments clearly indicate that the YDDΦXΦ motif contributes to the NEMO-IκBα. However, whether the motif is sufficient for the interaction remains to be established.

### 6.1.2 Characterization of the IκBα binding region within NEMO

In parallel, we performed *in cellulo* GPCA interaction experiments to identify the minimal binding region of IκBα on NEMO. This work was started by a master student of the team (Suresh Sivakumar) and I recently completed it by testing a number of second generation deletion mutants designed from the initial results. At the beginning, we focused on the C-terminal region of NEMO since the work of Hoffman and collaborators [27] indicated that the zinc finger domain is implicated in the interaction. Our results showed that the isolated C-terminal region of NEMO is not sufficient for the interaction with IκBα (Figure 6B, see NEMO constructs 344-419, 360-419, 394-419). A deletion construct of NEMO lacking the C-terminal region (NEMO (1-344)) interacts with IκBα at levels similar to full-length NEMO in most GPCA experiments. However, we observed that the NEMO (1-344) construct occasionally shows lower binding responses. We think that this might be related to a role of the zinc finger domain in stabilizing NEMO structure rather than a direct contribution to binding.

In the next step we created NEMO constructs comprising progressive C-terminal truncations to test the role of the different functional regions of NEMO towards binding. Results from the GPCA experiments show that deletion of the region between the CC2 domain and C-terminus reduces the interaction with IκBα (compare FL NEMO with NEMO (1-255), Figure 6C). Interestingly, the UBAN motif does not seem to be involved in the interaction since its deletion does not reduce binding (compare FL NEMO with NEMO (1-288)) to IκBα.

Together, these results suggest that NEMO may bind to IκBα *via* the CC2 region, in the proximity of the helix hinge region (HLX2).

## 6.2 Discussion

Results from *in cellulo* GPCA analyses show that the YDDΦXΦ motif sequence is implicated in binding to IκBα. However, whether this motif is sufficient for the interaction with IκBα remains to

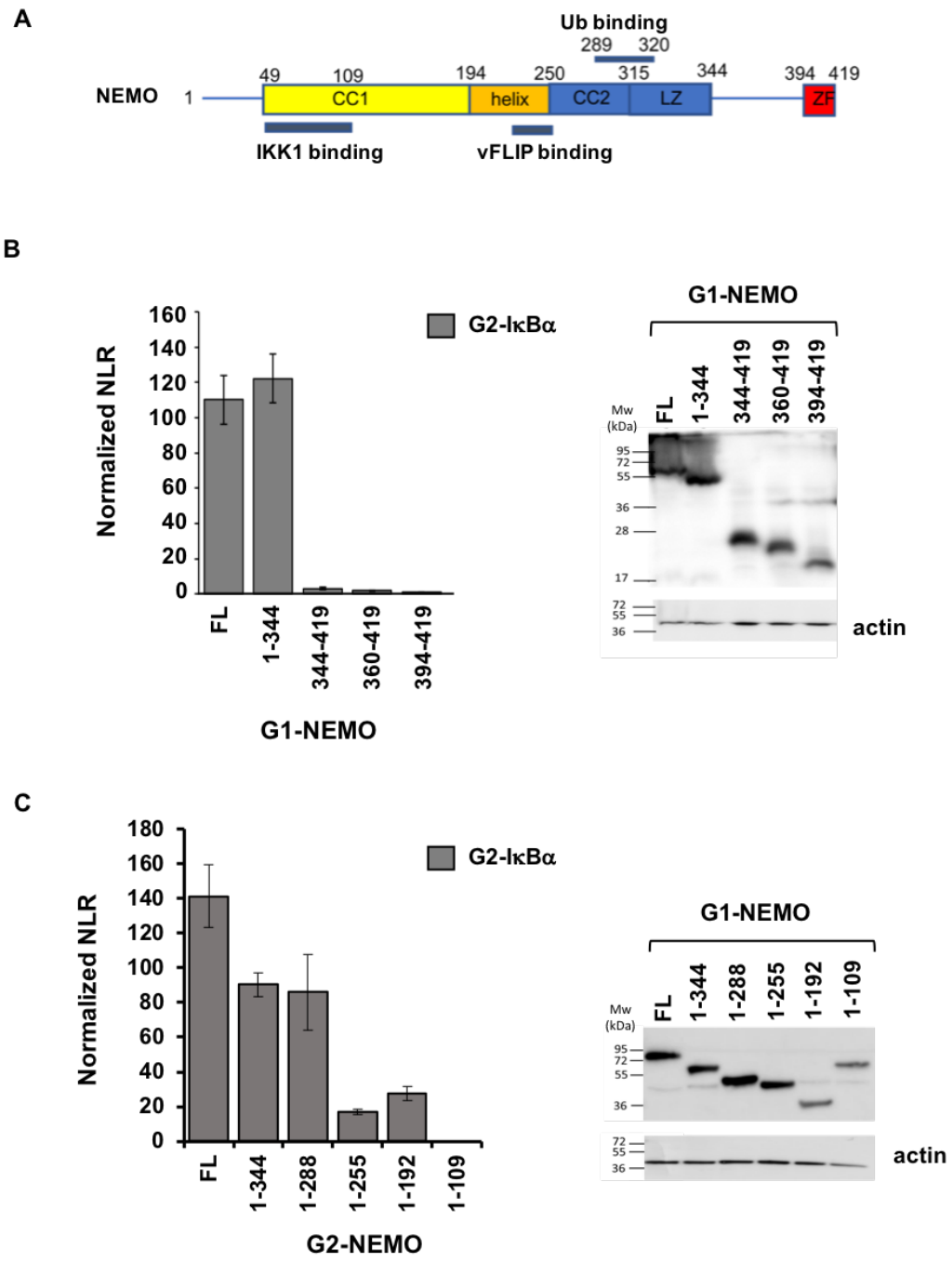



**Figure 6.** Analysis of the interactions of NEMO deletion constructs and full-length IκBα. **A)** Domain architecture of NEMO. The coiled-coil domain 1 (CC1), harboring the IKK2 binding domain, spans from residue 49 to 194. The CC1 is followed by an α-helix (195-250), a second coiled-coil domain (CC2) and a leucin-zipper domain. The central portion of the protein contain the vFLIP and the Ub binding domains (see Figure 9, chapter 2 for more details). At the C-terminus a zinc-finger is found. **B)** Results from an initial GPCA analysis ruling out the implication of the C-terminal region of NEMO comprising the zinc finger domain. HEK293T cells were co-transfected with G2-NEMO and G1-IκBα vectors. The NEMO constructs tested included the full-length wild type (FL) and the C-ter deletion constructs (1-344, 344-419, 360-419, and 394-419) (*left panel*). Expression levels of G2-NEMO in HEK293T cells. G2-NEMO fusions were detected by Western blotting using an antibody recognizing the *Gaussia princeps* luciferase (*right panel*). **C)** GPCA analysis of the C-terminal truncations of NEMO. (*Left panel*) The NEMO constructs tested included the full-length wild-type (FL) and the C-ter deletion constructs (1-344, 1-288, 1-255, 1-192 and 1-109). (*Right panel*) Expression levels of G2-NEMO in HEK293T cells.

be determined, since our pulldown analyses using YDDΦXΦ peptides and NEMO over-expressed in mammalian cells gave high background signals. Moreover, our GPCA data suggest that the YDDΦXΦ motif might cooperate with the previously identified N-terminal region of IκBα identified by Hoffman and collaborators [27]. It is in fact not unusual to observe different regions of a protein contributing to the interaction with a single partner.

In an attempt to decrease the background signal of cellular lysates overexpressing NEMO, we performed interaction analyses using purified recombinant NEMO. These analyses did not enable the detection of significant interactions even with the full-length IκBα protein. This lack of interaction between recombinant proteins could be explained by three different reasons. The first one is that the interaction between IκBα and NEMO may require a third bridging protein absent in bacteria. The second reason could be that the interaction depends on specific PTMs on NEMO, which bacteria are not able to provide. The third reason could be that recombinant NEMO produced in bacteria is misfolded, thus altering the IκBα docking site. Previous studies have shown that polyUb chains are required for the correct folding of NEMO [11, 13]. Hence, it will be important to test this hypothesis and perform interaction assays using purified recombinant IκBα and NEMO proteins in the presence of polyUb chains. For these experiments it would also be good to improve the level of purification of recombinant NEMO samples. In fact, we observed that recombinant NEMO co-purifies with *E. coli* chaperons. Chaperons associate to the disordered regions of NEMO acting as holdases, stabilizing the protein in an ATP-dependent manner. Hence, the NEMO purification protocol could be improved by addition of ATP/$Mg^{2+}$ to eliminate chaperon copurification.

Concerning the IκBα binding region on NEMO, Hoffman and collaborators [27] suggested that this corresponds to the C-terminal region, in particular the zinc-binding domain. Our GPCA data clearly indicate the C-terminal region of NEMO is not sufficient for binding to IκBα. However it seems that the C-terminal zinc-binding domain of NEMO confers stability to NEMO by affecting oligomerization processes that depend on the expression levels of the protein (Fabrice Agou, personal communication). This could explain the lower binding properties occasionally observed for the NEMO deletion constructs lacking the C-terminal region.

Given the implication of the region comprising the YDDΦXΦ motif in the interaction, we think that the motif might be recognized by a region of NEMO that shares sequence homology with IKK1 and IKK2. However, we were unable to identify this region by alignment of the NEMO sequence with the sequences of IKK1 and IKK2 (Laurent Bianchetti, unpublished data). Results from our GPCA experiments using deletion constructs of NEMO indicate that the IκBα binding site on NEMO is located within the CC2 domain, probably at the border with the helix hinge region (HLX2), and does

not involve the UBAN motif. Interestingly this region is known to interact with other protein partners, such as the vFLIP viral protein. In the next step, we will express a construct comprising HLX2 and CC2 regions of NEMO (residues 194-289) and test its interaction with IκBα by *in cellulo* GPCA analyses. This experiment will enable us to understand if the identified region is sufficient for the interaction or not.

# Conclusions and perspectives

The work described in this dissertation deals with the characterization of the interaction between the core subunits of the IKK complex and the IκBα substrate, the best studied inhibitor of NF-κB transcription factors. We initially focused on the interaction between IκBα and the catalytic subunit IKK2, which is responsible for the phosphorylation and subsequent degradation of the IκB proteins in the context of the canonical NF-κB signaling. In this way, we could identify the minimal binding region of IKK2 within IκBα, located within the disordered C-terminal region, right downstream of the PEST sequence. This region harbors a conserved linear motif with the sequence YDDΦXΦ (where Φ is a hydrophobic residue), which is necessary and sufficient for the docking interaction with both IKK1 and IKK2 subunits. In addition, this motif is also present in the IκBβ protein, orthologue of IκBα and in the p100 and IRF7 substrates of IKK1. Whereas the requirement for the YDDΦXΦ motif in the IKK1/p100 interaction is clearly shown by our results, the role of the motif in the IKK1/IRF7 interaction remains to be established.

The YDDΦXΦ motif contributes to the catalytic activity of the IKK complex. As a matter of fact, the modulation of the affinity of the IκBα interaction with IKK1 and IKK2 by mutations in the YDDΦXΦ consensus affects the phosphorylation levels of IκBα, with higher affinities correlating with higher phosphorylation levels. This, in turn, is reflected in the rates of IκBα degradation *in cellulo*, which are faster for the higher affinity interactions.

Results from our structural studies demonstrated for the first time that the IKK1 and IKK2 interactions with the YDDΦXΦ motif require stable IKK dimers. Indeed, the motif-binding pocket consists of the two SDDs contributed by each monomer subunit. In the context of the canonical complex the SDD domains belong to the IKK1 and IKK2 subunits of the heterodimer.

Our preliminary studies on the IκBα interaction with the NEMO regulatory subunit indicate that that the YDDΦXΦ motif is also involved in this interaction. It therefore appears that the IKK1/2 heterodimer and NEMO subunits of the canonical IKK complex recognize the same sequence element within the IκBα substrate, leading to an enhancement of the interaction *via* avidity effects. In line with these findings, we observed that a synthetic peptide containing the YDDΦXΦ motif is able to reduce the phosphorylation of IκBα by the endogenous IKK kinase.

This work lays the foundations for further developments. Our work shows that IKK1 tolerates a larger number of amino acid substitutions within the motif, suggesting that YDDΦXΦ might be a specific sub-motif of a wider consensus. Also, IKK proteins display a wide cohort of substantially different

substrates so it is reasonable to think that more than motif consensus can be recognized. For instance, despite being a target of IKK2, p105 does not have the YDDΦXΦ motif. Hence, it will be interesting to proceed with a broader characterization of the linear motifs recognized by IKK1 and IKK2 by proteomics approaches, such as phage display. This latter technique is a high throughput approach that allows for the study of protein-protein, protein-peptide and protein-DNA interactions based on engineered bacteriophages. Here, the sequence encoding for a protein or peptide of interest is inserted into the viral genome and expressed as a fusion product with the viral coating protein. In this way, the protein of interest is exposed on the viral surface and can be used as bait to identify possible partners (other proteins or DNA sequences) [328]. In our case, the most suited approach would be a phage display library resource, which was created by Dr. Ylva Ivarsson (Uppsala University, Sweden), that consists of overlapping peptides issued from the disordered regions of the human proteome[60]. By this approach, we will be able to better define the IKK1 docking motif, identify other motifs or sub-motifs recognized by IKK1 and IKK2, and, at the same time, identify potentially novel binding partners of IKK1 and IKK2.

Concerning the structural studies on the IKK2/IκBα peptide complex, current experiments based on x-linking mass spectrometry aim at improving the structural model of the IKK-IκBα complex by introducing additional constraints for the side-chains at the interface between the docking motif and IKK pocket. The details at atomic level will allow the design and the development of high affinity peptides that could eventually be proposed as biological inhibitors highly specific for the NF-κB related interactions of the IKK kinase complex.

Concerning the NEMO-IκBα interaction, results from *in cellulo* experiments suggest that IκBα interacts with NEMO primarily *via* its C-terminal YDDΦXΦ motif and that the N-terminus participates in the stabilization of the interaction. By contrast, the IκBα interacting region within NEMO seems to be located in the central portion of the protein, somewhere around the vFLIP binding domain. However, we were unable to reproduce the NEMO-IκBα interaction *in vitro* using recombinant proteins, suggesting that additional factors may be required. In the future it will be important to design an appropriate *in vitro* assay to further study this interaction. In this assay we could, for example, use recombinant NEMO produced in insect cells, to ensure post-translational modifications and correct folding, and add poly-ubiquitin chains that, according to the literature, seem to be required for folding of NEMO. Later on, x-linking mass spectrometry could be used to identify with high precision the regions of IκBα and NEMO that mediate the interaction.

# Chapter 7

# Side project: biochemical characterization of the interaction of the oncogenic ΔNp73α isoform with the E2F4/p130 repressor complex

The p53 protein family of proteins comprises three transcription factors, namely p53, p63 and p73, which exert fundamental anti-tumor activity. ΔNp73α is an oncogenic isoform of p73. This isoform is transcribed from an alternative promoter (P2) that localizes within intron 3 of the TP73 gene and, as a consequence, lacks the N-terminal transactivation domain (TAD) present in all p53 proteins. The TAD is the key element for the transcription activity as it permits the recruitment of transcriptional co-factors to the RNA polymerase II complex. However, ΔNp73α can still bind to p53 response elements (RE) within the regulatory regions of genes, thereby acting as an inhibitor of p53 signaling. In particular expression of ΔNp73α is upregulated in several types of cancer (such as breast, prostate and liver cancers) that are resistant to chemotherapy.

Our team in collaboration with the team of Massimo Tommasino of the International Agency for Cancer Research in Lyon performed proteomic analyses to identify proteins binding to and regulating the functions of ΔNp73α. Analyses by affinity purification - mass spectrometry (AP-MS) revealed that ΔNp73α interacts with the E2F4/DP1/p130 complex. Further functional studies showed that this interaction leads to the inhibition of specific genes including negative regulators of cellular proliferation.

E2F4 belongs to the E2F protein family of transcription factors, which act both as activators (E2F1/2/3/6) and repressors (E2F4/5) of transcription and which play a key role in the control of cell cycle. E2F proteins associate with DP1–4 proteins through the dimerization domain and interact with the retinoblastoma (RB) family proteins (pRb, p130, p107) *via* the transactivation domain (TAD). This interaction with RB proteins prevents recruitment of the transcriptional machinery, thereby inhibiting transcriptional activity. In particular the E2F4/DP1/p130 complex acts mainly as a repressive complex, inhibiting the transcription of genes implicated in cell proliferation.

In this work I contributed to a series of protein-protein interaction analyses that led to the identification and characterization of a direct interaction between ΔNp73α and the E2F4 protein. My contribution consisted in exploring different interaction assays for this interaction. In particular, I tested both pulldown and co-immunoprecipitation (co-IP) assay set-ups using ΔNp73α, E2F4 and DP1 proteins over-expressed in HEK293T cells and fused to different tags (3xFlag and HA).

Comparison of the results from these experiments with those obtained using the *in cellulo* GPCA assay, showed that this latter approach remained the most sensitive one for this interaction. We therefore used the GPCA assay coupled to deletion mutagenesis to map the regions of ΔNp73α and E2F4 involved in the interaction, which involve the N-terminal region of ΔNp73α and the C-terminal region (dimerization domain plus TAD) of E2F4.

The recombinant expression and purification of the full-length ΔNp73α and E2F4 proteins is very difficult. In addition, E2F4 requires DP partner proteins for stability. Hence, we designed a minimal construct of ΔNp73α and a minimal E2F4/DP1 heterodimer comprising the regions involved in the interaction identified in the GPCA experiments. The minimal E2F4/DP1 heterodimer was produce by co-expression in bacteria, with the E2F4 and DP1 proteins fused to a 6xHis and GST tag, respectively. The ΔNp73α construct was also produced in bacteria as fusion to the MBP tag. With these proteins, I could perform a pulldown experiment whose results showed that ΔNp73α interacts specifically with the E2F4/DP1 heterodimer.

Results from these experiments are described in a manuscript provided in annex, which is currently under revision (Taverniti V / Krynska H. *et al.* The E2F4/p130 repressor complex cooperates with oncogenic ΔNp73α to promote cell survival in human papillomavirus 38-transformed keratinocytes and in cancer cells).

# Bibliography


1. Dhanasekaran, N. & Premakumar, R. E. Signaling by dual specificity kinases. *Oncogene* **17**, 1447–1455 (1998).
2. Hanks, S. K. & Hunter, T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification1. *FASEB J.* **9**, 576–596 (1995).
3. Manning, G. *et al*. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
4. Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**, 65–77 (2011).
5. Ahuja, L. G., Taylor, S. S. & Kornev, A. P. Tuning the 'violin' of protein kinases: The role of dynamics-based allostery. *IUBMB Life* **71**, 685–696 (2019).
6. McClendon, L. *et al*. Dynamic architecture of a protein kinase. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E4623–E4631 (2014).
7. Polley, S. *et al.* A Structural Basis for IκB Kinase 2 Activation Via Oligomerization-Dependent Trans Auto-Phosphorylation. *PLoS Biol.* **11**, (2013).
8. Xu, G. *et al.* Crystal structure of inhibitor of κB kinase β. *Nat. 2011 4727343* **472**, 325–330 (2011).
9. Polley, S. *et al.* Structural Basis for the Activation of IKK1/α. *Cell Rep.* **17**, 1907 (2016).
10. Israël, A. The IKK complex, a central regulator of NF-kappaB activation. *Cold Spring Harb. Perspect. Biol.* **2**, (2010).
11. Catici, D. A. *et al*. Polyubiquitin Drives the Molecular Interactions of the NF-κB Essential Modulator (NEMO) by Allosteric Regulation. *J. Biol. Chem.* **290**, 14130–14139 (2015).
12. Cordier, F. *et al*. Solution Structure of NEMO Zinc Finger and Impact of an Anhidrotic Ectodermal Dysplasia with Immunodeficiency-related Point Mutation. *J. Mol. Biol.* **377**, 1419–1432 (2008).
13. Scholefield, J. *et al.* Super-resolution microscopy reveals a preformed NEMO lattice structure that is collapsed in incontinentia pigmenti. *Nat. Commun.* **7**, (2016).
14. Rahigi, S. *et al.* Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* **136**, 1098–1109 (2009).
15. Clark, K. Nanda, S. & Cohen, P. Molecular control of the NEMO family of ubiquitin-binding proteins. *Nat. Rev. Mol. Cell Biol.* **14**, 673–685 (2013).
16. Larabi, A. *et al.* Crystal Structure and Mechanism of Activation of TANK-Binding Kinase 1. *Cell Rep.* **3**, 734–746 (2013).

17. Hayden, M. S. & Ghosh, S. NF-κB in immunobiology. *Cell Res. 2011 212* **21**, 223–244 (2011).
18. Kumar, H., Kawai, T. & Akira, S. Pathogen recognition by the innate immune system. *Int. Rev. Immunol.* **30**, 16–34 (2011).
19. Paszek, P. *et al.* Population robustness arising from cellular heterogeneity. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11644–11649 (2010).
20. Johnson, R. F., Witzel, I. I. & Perkins, N. D. p53-dependent regulation of mitochondrial energy production by the RelA subunit of NF-κB. *Cancer Res.* **71**, 5588–5597 (2011).
21. Huang, S. *et al*. Blockade of NF-kappaB activity in human prostate cancer cells is associated with suppression of angiogenesis, invasion, and metastasis. *Oncogene* **20**, 4188–4197 (2001).
22. Nilsson-Payant, B. E. *et al.* The NF-κB Transcriptional Footprint Is Essential for SARS-CoV-2 Replication. *J. Virol.* **95**, (2021).
23. Varga, Z. *et al.* Endothelial cell infection and endotheliitis in COVID-19. *Lancet (London, England)* **395**, 1417–1418 (2020).
24. Muller, M. *et al*. A comparative approach to characterize the landscape of host-pathogen protein-protein interactions. *J. Vis. Exp.* (2013) doi:10.3791/50404.
25. Cassonnet, P. *et al.* Benchmarking a luciferase complementation assay for detecting protein complexes. *Nat. Methods* **8**, 990–992 (2011).
26. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).
27. Schröfelbauer, B. *et al*. NEMO ensures signaling specificity of the pleiotropic IKKβ by directing its kinase activity toward IκBα. *Mol. Cell* **47**, 111–121 (2012).
28. Bagnéris, C. *et al.* Crystal Structure of a vFlip-IKKγ Complex: Insights into Viral Activation of the IKK Signalosome. *Mol. Cell* **30**, 620–631 (2008).
29. Shaw, A. S. *et al*. Kinases and pseudokinases: lessons from RAF. *Mol. Cell. Biol.* **34**, 1538–1546 (2014).
30. Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification - PubMed. *FASEB J* . (1995).
31. Mora, A. *et al*. PDK1, the master regulator of AGC kinase signal transduction. *Semin. Cell Dev. Biol.* **15**, 161–170 (2004).
32. Nakayama, S., Moncrief, N. D. & Kretsinger R. H.. Evolution of EF-hand calcium-modulated proteins. II. Domains of several subfamilies have diverse evolutionary histories. *J. Mol. Evol.* **34**, 416–448 (1992).

33. Biondi, R. M. & Nebreda, A. R. Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem. J.* **372**, 1–13 (2003).
34. Murray, P. J. The JAK-STAT signaling pathway: input and output integration. *J. Immunol.* **178**, 2623–2629 (2007).
35. Cargnello, M. & Roux, P. P. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol. Mol. Biol. Rev.* **75**, 50–83 (2011).
36. Knippschild, U. *et al.* The casein kinase 1 family: participation in multiple cellular processes in eukaryotes. *Cell. Signal.* **17**, 675–689 (2005).
37. van Dijkman, S. C. *et al*. Effect of Age-Related Factors on the Pharmacokinetics of Lamotrigine and Potential Implications for Maintenance Dose Optimisation in Future Clinical Trials. *Clin. Pharmacokinet.* **57**, 1039–1053 (2018).
38. Aparicio, J. G. & Applebury, M. L. The photoreceptor guanylate cyclase is an autophosphorylating protein kinase. *J. Biol. Chem.* **271**, 27083–27089 (1996).
39. Kwon, A. *et al.* Tracing the origin and evolution of pseudokinases across the tree of life. *Sci. Signal.* **12**, (2019).
40. Eyers, P. A., Keeshan, K. & Kannan, N. Tribbles in the 21st Century: The Evolving Roles of Tribbles Pseudokinases in Biology and Disease. *Trends Cell Biol.* **27**, 284–298 (2017).
41. Mendrola, J. M. *et al*. Receptor tyrosine kinases with intracellular pseudokinase domains. *Biochem. Soc. Trans.* **41**, 1029–1036 (2013).
42. Lopez-Borges, S. & Lazo, P. A. The human vaccinia-related kinase 1 (VRK1) phosphorylates threonine-18 within the mdm-2 binding site of the p53 tumour suppressor protein. *Oncogene* **19**, 3656–3664 (2000).
43. None, M., *et al*. Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat. Struct. Biol.* **9**, 273–277 (2002).
44. Steinberg, S. F. Post-translational modifications at the ATP-positioning G-loop that regulate protein kinase activity. *Pharmacol. Res.* **135**, 181–187 (2018).
45. Ramakrishnan, C., Dani, V. S. & Ramasarma, T. A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. *Protein Eng. Des. Sel.* **15**, 783–798 (2002).
46. Taylor, S. S. *et al.* PKA: a portrait of protein kinase dynamics. *Biochim. Biophys. Acta* **1697**, 259–269 (2004).
47. Iyer, G. H., Moore, M. J. & Taylor, S. S. Consequences of lysine 72 mutation on the phosphorylation and activation state of cAMP-dependent kinase. *J. Biol. Chem.* **280**, 8800–8807 (2005).

48. Song, H. *et al.* Phosphoprotein–Protein Interactions Revealed by the Crystal Structure of Kinase-Associated Phosphatase in Complex with PhosphoCDK2. *Mol. Cell* **7**, 615–626 (2001).
49. Kyriakis, J. M. In the Beginning, There Was Protein Phosphorylation. *J. Biol. Chem.* **289**, 9460 (2014).
50. Modi, V. & Dunbrack, R. L. Defining a new nomenclature for the structures of active and inactive kinases. *Proc. Natl. Acad. Sci.* **116**, 6818–6827 (2019).
51. Moore, M. J., Adams, J. A. & Taylor, S. S. Structural basis for peptide binding in protein kinase A. Role of glutamic acid 203 and tyrosine 204 in the peptide-positioning loop. *J. Biol. Chem.* **278**, 10613–10618 (2002).
52. Yang, J. *et al.* A conserved Glu-Arg salt bridge connects coevolved motifs that define the eukaryotic protein kinase fold. *J. Mol. Biol.* **415**, 666–679 (2012).
53. Shudler, S. & Niv, M. Y. BlockMaster: partitioning protein kinase structures using normal-mode analysis. *J. Phys. Chem. A* **113**, 7528–7534 (2009).
54. Kannan, N. & Neuwald, A. F. Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? *J. Mol. Biol.* **351**, 956–972 (2005).
55. Scheeff, E. D. & Bourne, P. E. Structural evolution of the protein kinase-like superfamily. *PLoS Comput. Biol.* **1**, e49 (2005).
56. Kornev, A. P. & Taylor, S. S. Dynamics-Driven Allostery in Protein Kinases. *Trends Biochem. Sci.* **40**, 628–647 (2015).
57. Kornev, A. P. & Taylor, S. S. Defining the Conserved Internal Architecture of a Protein Kinase. *Biochim. Biophys. Acta* **1804**, 440 (2010).
58. Lochhead, P. A. Protein kinase activation loop autophosphorylation in cis: overcoming a Catch-22 situation. *Sci. Signal.* **2**, (2009).
59. Davey, N. E., Cyert, M. S. & Moses, A. M. Short linear motifs - ex nihilo evolution of protein regulation. *Cell Commun. Signal.* **13**, (2015).
60. Gógl, G. *et al*. Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores. *Trends Biochem. Sci.* **44**, 300–311 (2019).
61. Hari, S. B., Merritt, E. A. & Maly, D. J. Sequence determinants of a specific inactive protein kinase conformation. *Chem. Biol.* **20**, 806–815 (2013).
62. Belham, C., Shilan, W. & Avruch, J. Intracellular signalling: PDK1 – a kinase at the hub of things. *Curr. Biol.* **9**, R93–R96 (1999).
63. Zhu, G. *et al.* Exceptional disfavor for proline at the P + 1 position among AGC and CAMK kinases establishes reciprocal specificity between them and the proline-directed kinases. *J.*

*Biol. Chem.* **280**, 10743–10748 (2005).

64. Canagarajah, B. J. *et al*. Activation Mechanism of the MAP Kinase ERK2 by Dual Phosphorylation. *Cell* **90**, 859–869 (1997).
65. Songyang, Z. *et al.* Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr. Biol.* **4**, 973–982 (1994).
66. Clark-Lewis, I., Sanghera, J. S. & Pelech, S. L. Definition of a consensus sequence for peptide substrate recognition by p44mpk, the meiosis-activated myelin basic protein kinase. *J. Biol. Chem.* **266**, 15180–15184 (1991).
67. Gonzalez, F. A., Raden, D. L. & Davis, R. J. Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *J. Biol. Chem.* **266**, 22159–22163 (1991).
68. Kannan, N. & Neuwald, A. F. Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2alpha. *Protein Sci.* **13**, 2059–2077 (2004).
69. Meggio, F. & Pinna, L. A. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **17**, 349–368 (2003).
70. Miller, C. J. & Turk, B. E. Homing in: Mechanisms of Substrate Targeting by Protein Kinases. *Trends Biochem. Sci.* **43**, 380–394 (2018).
71. Tanoue, T. *et al*. A conserved docking motif in MAP kinases common to substrates, activators and regulators. *Nat. Cell Biol.* **2**, 110–116 (2000).
72. Garai, Á. *et al.* Specificity of linear motifs that bind to a common mitogen-activated protein kinase docking groove. *Sci. Signal.* **5**, (2012).
73. Sheridan D. L. *et al*. Substrate discrimination among mitogen-activated protein kinases through distinct docking sequence motifs. *J. Biol. Chem.* **283**, 19511–19520 (2008).
74. Su, B. & Jacinto, E. Mammalian TOR signaling to the AGC kinases. Crit Rev Biochem Mol Biol. **46**, 527–547 (2011).
75. Aubol, B. E. *et al.* Processive phosphorylation of alternative splicing factor/splicing factor 2. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12601–12606 (2003).
76. Örd, M. *et al.* Proline-Rich Motifs Control G2-CDK Target Phosphorylation and Priming an Anchoring Protein for Polo Kinase Localization. *Cell Rep.* **31**, (2020).
77. Faustova, I. *et al.* A new linear cyclin docking motif that mediates exclusively S-phase CDK-specific signaling. *EMBO J.* **40**, (2021).
78. Pearce, L. R., Komander, D. & Alessi, D. R. The nuts and bolts of AGC protein kinases. *Nat. Rev. Mol. Cell Biol.* **11**, 9–22 (2010).

79. Rellos, P. *et al.* Structure of the CaMKIIdelta/calmodulin complex reveals the molecular mechanism of CaMKII kinase activation. *PLoS Biol.* **8**, (2010).
80. Hamil, S. *et al*. Structural Basis for Noncanonical Substrate Recognition of Cofilin/ADF Proteins by LIM Kinases. *Mol. Cell* **62**, 397–408 (2016).
81. Homan, K. T. & Tesmer, J. J. G. Structural insights into G protein-coupled receptor kinase function. *Curr. Opin. Cell Biol.* **27**, 25–31 (2014).
82. Doble, B. W. & Woodgett, J. R. GSK-3: tricks of the trade for a multi-tasking kinase. *J. Cell Sci.* **116**, 1175–1186 (2003).
83. Valk, E. *et al.* Multistep phosphorylation systems: tunable components of biological signaling circuits. *Mol. Biol. Cell* **25**, 3456 (2014).
84. Kõivomägi, M. *et al.* Multisite phosphorylation networks as signal processors for Cdk1. *Nat. Struct. Mol. Biol.* **20**, 1415–1424 (2013).
85. Holt, L. J. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325**, 1682–1686 (2009).
86. Pawson, T., Raina, M. & Nash, P. Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett.* **513**, 2–10 (2002).
87. Felder, S. *et al.* SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. *Mol. Cell. Biol.* **13**, 1449 (1993).
88. Liu, B. A. *et al.* The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. *Sci. Signal.* **4**, (2011).
89. Grucza, R. A. *et al*. Role of electrostatic interactions in SH2 domain recognition: salt-dependence of tyrosyl-phosphorylated peptide binding to the tandem SH2 domain of the Syk kinase and the single SH2 domain of the Src kinase. *Biochemistry* **39**, 10072–10081 (2000).
90. Pawson, T., Gish, G. D. & Nash, P. SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.* **11**, 504–511 (2001).
91. Grebien, F. *et al.* Targeting the SH2-kinase interface in Bcr-Abl inhibits leukemogenesis. *Cell* **147**, 306–319 (2011).
92. Saksela, K. *et al.* SH3 domain ligand binding: What's the consensus and where's the specificity? *FEBS Lett.* **586**, 2609–2614 (2012).
93. Nishida, M. *et al.* Novel recognition mode between Vav and Grb2 SH3 domains. *EMBO J.* **20**, 2995–3007 (2001).
94. Smith, M. J. *et al*. Screening for PTB domain binding partners and ligand specificity using proteome-derived NPXY peptide arrays. *Mol. Cell. Biol.* **26**, 8461–8474 (2006).
95. Durocher, D. *et al*. The FHA domain is a modular phosphopeptide recognition motif. *Mol.*

*Cell* **4**, 387–394 (1999).

96. Wu, G. *et al.* Structural basis of Smad2 recognition by the Smad anchor for receptor activation. *Science* **287**, 92–97 (2000).
97. Buday, L. & Downward, J. Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* **73**, 611–620 (1993).
98. Schaeper, U. *et al.* Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses. *J. Cell Biol.* **149**, 1419–1432 (2000).
99. Ong, S. H. *et al.* Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins.
100. Schlessinger, J. & Lemmon, M. A. SH2 and PTB domains in tyrosine kinase signaling. *Sci. STKE* **2003**, re12–re12 (2003).
101. Papin, J. A. *et al*. Reconstruction of cellular signalling networks and analysis of their properties. *Nat. Rev. Mol. Cell Biol.* **6**, 99–111 (2005).
102. Gilmore, T. D. Introduction to NF-jB: players, pathways, perspectives. *Oncogene* **25**, 6680–6684 (2006).
103. Hayden, M. S. & Ghosh, S. Signaling to NF-κB. *Genes and Development* vol. 18 2195–2224 (2004).
104. Sil, A. K., *et al*. IκB kinase-α acts in the epidermis to control skeletal and craniofacial morphogenesis. *Nature* **428**, 660–664 (2004).
105. Zhang, Q., Lenardo, M. J. & Baltimore, D. 30 Years of NF-κB: A Blossoming of Relevance to Human Pathobiology. *Cell* vol. 168 37–57 (2017).
106. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harbor perspectives in biology* vol. 1 a000034 (2009).
107. Schoonbroodt, S. *et al.* Crucial Role of the Amino-Terminal Tyrosine Residue 42 and the Carboxyl-Terminal PEST Domain of IκBα in NF-κB Activation by an Oxidative Stress. *J. Immunol.* **164**, 4292–4300 (2000).
108. Hoffmann, A., Natoli, G. & Ghosh, G. Transcriptional regulation via the NF-κB signaling module. *Oncogene* vol. 25 6706–6716 (2006).
109. Wu, C. & Ghosh, S. Differential phosphorylation of the signal-responsive domain of IκBα and IκBβ by IκB kinases. *J. Biol. Chem.* **278**, 31980–31987 (2003).
110. Carmody, R. J. *et al*. H. Negative regulation of toll-like receptor signaling by NF-κB p50 ubiquitination blockade. *Science (80-. ).* **317**, 675–678 (2007).
111. Ghosh, G. *et al*. NF-κB Regulation: Lessons from Structures. *Immunol. Rev.* **246**, 36 (2012).

112. Karin, M. Nuclear factor-κB in cancer development and progression. *Nature* vol. 441 431–436 (2006).
113. Chen, G., Cao, P. & Goeddel, D. V. TNF-Induced Recruitment and Activation of the IKK Complex Require Cdc37 and Hsp90. *Mol. Cell* **9**, 401–410 (2002).
114. Chen, Z. J., Bhoj, V. & Seth, R. B. Ubiquitin, TAK1 and IKK: Is there a connection? *Cell Death and Differentiation* vol. 13 687–692 (2006).
115. Maubach, G., Schmädicke, A.-C. & Naumann, M. NEMO Links Nuclear Factor-κB to Human Diseases. *Trends Mol. Med.* **23**, 1138–1155 (2017).
116. Scheidereit, C. IκB kinase complexes: Gateways to NF-κB activation and transcription. *Oncogene* vol. 25 6685–6705 (2006).
117. Tanaka, H., Fujita, N. & Tsuruo, T. 3-Phosphoinositide-dependent protein kinase-1-mediated IκB kinase β(IKKB) phosphorylation activates NF-κB signaling. *J. Biol. Chem.* **280**, 40965–40973 (2005).
118. Viatour, P. *et al*. Phosphorylation of NF-κB and IκB proteins: Implications in cancer and inflammation. *Trends in Biochemical Sciences* vol. 30 43–52 (2005).
119. Krappmann, D. & Scheidereit, C. A pervasive role of ubiquitin conjugation in activation and termination of IκB kinase pathways. *EMBO Reports* vol. 6 321–326 (2005).
120. Visekruna, A. *et al.* Proteasome-mediated degradation of IκBα and processing of p105 in Crohn disease and ulcerative colitis. *J. Clin. Invest.* **116**, 3195–3203 (2006).
121. Zhong, H. *et al*. The phosphorylation status of nuclear NF-κB determines its association with CBP/p300 or HDAC-1. *Mol. Cell* **9**, 625–636 (2002).
122. Guan, H., Hou, S. & Ricciardi, R. P. DNA binding of repressor nuclear factor-κB p50/p50 depends on phosphorylation of Ser337 by the protein kinase A catalytic subunit. *J. Biol. Chem.* **280**, 9957–9962 (2005).
123. Yu, S. H. *et al*. Stimulation of c-Rel transcriptional activity by PKA catalytic subunit β. *J. Mol. Med.* **82**, 621–628 (2004).
124. Mattioli, I. *et al.* Transient and Selective NF-κB p65 Serine 536 Phosphorylation Induced by T Cell Costimulation Is Mediated by IκB Kinase β and Controls the Kinetics of p65 Nuclear Import. *J. Immunol.* **172**, 6336–6344 (2004).
125. Cohen, S., Achbert-Weiner, H. & Ciechanover, A. Dual Effects of I B Kinase -Mediated Phosphorylation on p105 Fate: SCF -TrCP-Dependent Degradation and SCF -TrCP-Independent Processing. *Mol. Cell. Biol.* **24**, 475–486 (2004).
126. Moorthy, A. K. *et al.* The 20S proteasome processes NF-κB1 p105 into p50 in a translation-independent manner. *EMBO J.* **25**, 1945–1956 (2006).

127. Eliopoulos, A. G. *et al.* Epstein-Barr virus-encoded latent infection membrane protein 1 regulates the processing of p100 NF-κB2 to p52 via an IKKγ/NEMO-independent signalling pathway. *Oncogene* **22**, 7557–7569 (2003).
128. Sun, S. C. Non-canonical NF-κB signaling pathway. *Cell Research* vol. 21 71–85 (2011).
129. Xiao, G., Harhaj, E. W. & Sun, S. C. NF-κB-inducing kinase regulates the processing of NF-κB2 p100. *Mol. Cell* **7**, 401–409 (2001).
130. Liang, C., Zhang, M. & Sun, S. C. β-TrCP binding and processing of NF-κB2/p100 involve its phosphorylation at serines 866 and 870. *Cell. Signal.* **18**, 1309–1317 (2006).
131. Sun, S. C. The non-canonical NF-κB pathway in immunity and inflammation. *Nature Reviews Immunology* vol. 17 545–558 (2017).
132. Xiao, G., Fong, A. & Sun, S. C. Induction of p100 processing by NF-κB-inducing kinase involves docking IκB kinase α (IKKα) to p100 and IKKα-mediated phosphorylation. *J. Biol. Chem.* **279**, 30099–30105 (2004).
133. Ishimaru, N. *et al*. Regulation of naive T cell function by the NF-κB2 pathway. *Nat. Immunol.* **7**, 763–772 (2006).
134. Kato, T. *et al*. CK2 is a C-terminal IκB kinase responsible for NF-κB activation during the UV response. *Mol. Cell* **12**, 829–839 (2003).
135. Imbert, V. *et al.* Tyrosine phosphorylation of IκB-α activates NF-κB without proteolytic degradation of IκB-α. *Cell* **86**, 787–798 (1996).
136. Ryo, A. *et al.* Regulation of NF-κB Signaling by Pin1-Dependent Prolyl Isomerization and Ubiquitin-Mediated Proteolysis of p65/RelA. *Mol. Cell* **12**, 1413–1426 (2003).
137. Bohuslav, J. *et al*. p53 induces NF-κB activation by an IκB kinase-independent mechanism involving phosphorylation of p65 by ribosomal S6 kinase 1. *J. Biol. Chem.* **279**, 26115–26125 (2004).
138. Fischle, W. *et al.* Enzymatic Activity Associated with Class II HDACs Is Dependent on a Multiprotein Complex Containing HDAC3 and SMRT/N-CoR. *Mol. Cell* **9**, 45–57 (2002).
139. Hoberg, J. E. *et al*. IκB Kinase α-Mediated Derepression of SMRT Potentiates Acetylation of RelA/p65 by p300. *Mol. Cell. Biol.* **26**, 457–471 (2006).
140. Yamamoto, Y. *et al*. Histone H3 phosphorylation by IKK-a is critical for cytokine-induced gene expression. (2003).
141. Huang, W. C. *et al*. Phosphorylation of CBP by IKKα Promotes Cell Growth by Switching the Binding Preference of CBP from p53 to NF-κB. *Mol. Cell* **26**, 75 (2007).
142. Lawrence, T. *et al*. IKKα limits macrophage NF-κB activation and contributes to the resolution of inflammation. *Nature* **434**, 1138–1143 (2005).

143. Wegener, E. *et al.* Essential role for IkappaB kinase beta in remodeling Carma1-Bcl10-Malt1 complexes upon T cell activation. *Mol. Cell* **23**, 13–23 (2006).
144. Carayol, N & Wang, C.-Y. IKKalpha stabilizes cytosolic beta-catenin by inhibiting both canonical and non-canonical degradation pathways. *Cell. Signal.* **18**, 1941–1946 (2006).
145. Anest, V. *et al.* A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. *Nature* **423**, 659–663 (2003).
146. Liu, B. *et al.* IKKalpha is required to maintain skin homeostasis and prevent skin cancer. *Cancer Cell* **14**, 212–225 (2008).
147. Waterfield, M. *et al*. IkappaB kinase is an essential component of the Tpl2 signaling pathway. *Mol. Cell. Biol.* **24**, 6040–6048 (2004).
148. Lee, S. *et al.* IkappaB kinase beta phosphorylates Dok1 serines in response to TNF, IL-1, or gamma radiation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17416–17421 (2004).
149. Finnberg, N. & El-Deiry, W. S. Activating FOXO3a, NF-κB and p53 by targeting IKKs: An effective multi-faceted targeting of the tumor-cell phenotype? *Cancer Biol. Ther.* **3**, 614–616 (2004).
150. Lee, D. F. *et al.* IKK beta suppression of TSC1 links inflammation and tumor angiogenesis via the mTOR pathway. *Cell* **130**, 440–455 (2007).
151. Xia, Y. *et al.* Phosphorylation of p53 by IkappaB kinase 2 promotes its degradation by beta-TrCP. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2629–2634 (2009).
152. Lamberti, C. *et al.* Regulation of β-Catenin Function by the IκB Kinases*. *undefined* **276**, 42276–42286 (2001).
153. Prescott, J. & Cook, S. Targeting IKKβ in Cancer: Challenges and Opportunities for the Therapeutic Utilisation of IKKβ Inhibitors. *Cells* **7**, 115 (2018).
154. Liu, S. *et al.* Crystal structure of a human IκB Kinase β asymmetric dimer. *J. Biol. Chem.* **288**, 22758–22767 (2013).
155. Yamaoka, S. *et al.* Complementation cloning of NEMO, a component of the IkappaB kinase complex essential for NF-kappaB activation. *Cell* **93**, 1231–1240 (1998).
156. Wu, C.-J. *et al*. Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF-kappaB activation [corrected]. *Nat. Cell Biol.* **8**, 398–406 (2006).
157. Kensche, T. *et al.* Analysis of nuclear factor-κB (NF-κB) essential modulator (NEMO) binding to linear and lysine-linked ubiquitin chains and its role in the activation of NF-κB. *J. Biol. Chem.* **287**, 23626–23634 (2012).
158. Ea, C.-E. *et al*. Activation of IKK by TNFalpha requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. *Mol. Cell* **22**, 245–257 (2006).

159. Ori, D. *et al.* Essential roles of K63-linked polyubiquitin-binding proteins TAB2 and TAB3 in B cell activation via MAPKs. *J. Immunol.* **190**, 4037–4045 (2013).
160. Israël, A. NF-kappaB activation: Nondegradative ubiquitination implicates NEMO. *Trends Immunol.* **27**, 395–397 (2006).
161. Clark, K., Nanda, S. & Cohen, P. Molecular control of the NEMO family of ubiquitin-binding proteins. *Nat. Rev. Mol. Cell Biol. 2013 1410* **14**, 673–685 (2013).
162. Brenner, D., Blaser, H. & Mak, T. W. Regulation of tumour necrosis factor signalling: live or let die. *Nat. Rev. Immunol.* **15**, 362–374 (2015).
163. Wu, Z.-H. & Miyamoto, S. Induction of a pro-apoptotic ATM-NF-kappaB pathway and its repression by ATR in response to replication stress. *EMBO J.* **27**, 1963–1973 (2008).
164. Fenner, B. J., Scannell, M. S. & Prenh, J. H. M. Expanding the substantial interactome of NEMO using protein microarrays. *PLoS One* **5**, (2010).
165. Sadek, J. *et al.* Modulation of virus-induced NF-κB signaling by NEMO coiled coil mimics. *Nat. Commun.* **11**, (2020).
166. Fennell, L. M., Rahighi, S. & Ikeda, F. Linear ubiquitin chain-binding domains. *FEBS J.* **285**, 2746–2761 (2018).
167. Rushe, M. *et al.* Structure of a NEMO/IKK-associating domain reveals architecture of the interaction site. *Structure* **16**, 798–808 (2008).
168. Hemmi, H. *et al.* The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection. *J. Exp. Med.* **199**, 1641–1650 (2004).
169. Häcker, H. & Karin, M. Regulation and function of IKK and IKK-related kinases. *Sci. STKE* **2006**, (2006).
170. Kawai, T. *et al.* Interferon-alpha induction through Toll-like receptors involves a direct interaction of IRF7 with MyD88 and TRAF6. *Nat. Immunol.* **5**, 1061–1068 (2004).
171. Vercammen, E., Staal, J. & Beyaert, R. Sensing of viral infection and activation of innate immunity by toll-like receptor 3. *Clin. Microbiol. Rev.* **21**, 13–25 (2008).
172. Seth, R. B. *et al*. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* **122**, 669–682 (2005).
173. Ma, X. *et al.* Molecular basis of Tank-binding kinase 1 activation by transautophosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9378–9383 (2012).
174. Clark, K. *et al.* Novel cross-talk within the IKK family controls innate immunity. *Biochem. J.* **434**, 93–104 (2011).
175. Liu, T. *et al*. NF-κB signaling in inflammation. *Signal Transduction and Targeted Therapy* vol. 2 (2017).

176. Platanitis, E. & Decker, T. Regulatory networks involving STATs, IRFs, and NFκB in inflammation. *Front. Immunol.* **9**, (2018).
177. O'Brien, W. T. *et al.* The NLRP3 inflammasome in traumatic brain injury: potential as a biomarker and therapeutic target. *J. Neuroinflammation 2020 171* **17**, 1–12 (2020).
178. Kovacs, S. B. & Miao, E. A. Gasdermins: Effectors of Pyroptosis. *Trends Cell Biol.* **27**, 673–684 (2017).
179. Hayden, M. S. & Ghosh, S. NF-κB in immunobiology. *Cell Research* vol. 21 223–244 (2011).
180. Mosser, D. M. The many faces of macrophage activation. *J. Leukoc. Biol.* **73**, 209–212 (2003).
181. Liu, X. *et al.* Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. *Nature* **535**, 153–158 (2016).
182. Fink, S. L. & Cookson, B. T. Caspase-1-dependent pore formation during pyroptosis leads to osmotic lysis of infected host macrophages. *J. Immunol.* **202**, 1913–1926 (2006).
183. Rathinam, V. A. K. *et al.* The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. *Nat. Immunol.* **11**, 395–402 (2010).
184. Schroder, K. & Tschopp, J. The Inflammasomes. *Cell* vol. 140 821–832 (2010).
185. Greten, F. R. *et al.* NF-κB Is a Negative Regulator of IL-1β Secretion as Revealed by Genetic and Pharmacological Inhibition of IKKβ. *Cell* **130**, 918–931 (2007).
186. Nakahira, K. *et al.* Autophagy proteins regulate innate immune responses by inhibiting the release of mitochondrial DNA mediated by the NALP3 inflammasome. *Nat. Immunol.* **12**, 222–230 (2011).
187. Sasaki, Y. & Iwai, K. Roles of the NF-κB pathway in B-lymphocyte biology. *Curr. Top. Microbiol. Immunol.* **393**, 177–209 (2015).
188. Kayagaki, N. *et al.* BAFF/BLyS receptor 3 binds the B cell survival factor BAFF ligand through a discrete surface loop and promotes processing of NF-κB2. *Immunity* **17**, 515–524 (2002).
189. Xue, L. *et al.* Defective developmental and function of Bcl10-deficient follicular, marginal zone and B1 B cells. *Nat. Immunol.* **4**, 857–865 (2003).
190. MacKay, F. & Schneider, P. Cracking the BAFF code. *Nature Reviews Immunology* vol. 9 491–502 (2009).
191. Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).
192. Wang, L., Wuerffel, R. & Kenter, A. L. NF-κB binds to the immunoglobulin Sγ3 region in

vivo during class switch recombination. *Eur. J. Immunol.* **36**, 3315–3323 (2006).

193. McHeyzer-Williams, M. G. B cells as effectors. *Current Opinion in Immunology* vol. 15 354–361 (2003).
194. Gantke, T. *et al*. IκB kinase regulation of the TPL-2/ERK MAPK pathway. *Immunological Reviews* vol. 246 168–182 (2012).
195. Zhu, J., Yamane, H. & Paul, W. E. Differentiation of Effector CD4 T Cell Populations. *Annu. Rev. Immunol.* **28**, 445–489 (2010).
196. Yu, J. *et al.* Noncanonical NF-κB Activation Mediates STAT3-Stimulated IDO Upregulation in Myeloid-Derived Suppressor Cells in Breast Cancer. *J. Immunol.* **193**, 2574–2586 (2014).
197. Ruan, Q. *et al.* Development of Foxp3+ Regulatory T Cells Is Driven by the c-Rel Enhanceosome. *Immunity* **31**, 932–940 (2009).
198. Chang, J. H. *et al.* Ubc13 maintains the suppressive function of regulatory T cells and prevents their conversion into effector-like T cells. *Nat. Immunol.* **13**, 481–490 (2012).
199. Inohara, N. *et al.* Host recognition of bacterial muramyl dipeptide mediated through NOD2: Implications for Crohn's disease. *J. Biol. Chem.* **278**, 5509–5512 (2003).
200. Inohara, N. & Nuñez, G. NODS: Intracellular proteins involved in inflammation and apoptosis. *Nature Reviews Immunology* vol. 3 371–382 (2003).
201. Sabroe, I. *et al.* Selective Roles for Toll-Like Receptor (TLR)2 and TLR4 in the Regulation of Neutrophil Activation and Life Span. *J. Immunol.* **170**, 5268–5275 (2003).
202. Tang, G. *et al*. Blocking caspase-3-mediated proteolysis of IKKβ suppresses TNF-α-induced apoptosis. *Mol. Cell* **8**, 1005–1016 (2001).
203. Liu, N. *et al.* NF-κB protects from the lysosomal pathway of cell death. *EMBO J.* **22**, 5313–5322 (2003).
204. Li, Z.-W. *et al*. IKKβ Is Required for Peripheral B Cell Survival and Proliferation. *J. Immunol.* **170**, 4630–4637 (2003).
205. Grossmann, M. *et al.* The anti-apoptotic activities of Rel and RelA required during B-cell maturation involve the regulation of Bcl-2 expression. *EMBO J.* **19**, 6351–6360 (2000).
206. Zheng, Y. *et al*. Combined deficiency of p50 and cRel in CD4+ T cells reveals an essential requirement for nuclear factor κB in regulating mature T cell survival and in vivo function. *J. Exp. Med.* **197**, 861–874 (2003).
207. Mitchell, T. C. *et al.* A short domain within Bcl-3 is responsible for its lymphocyte survival activity. in *Annals of the New York Academy of Sciences* vol. 975 132–147 (New York Academy of Sciences, 2002).
208. Wan, Y. Y. & DeGregori, J. The survival of antigen-stimulated T cells requires NFκB-

mediated inhibition of p73 expression. *Immunity* **18**, 331–342 (2003).

209. Ryan, K. M. *et al*. Role of NF-κB in p53-mediated programmed cell death. *Nature* **404**, 892–897 (2000).
210. Rocha, S. *et al*. p53 Represses Cyclin D1 Transcription through Down Regulation of Bcl-3 and Inducing Increased Association of the p52 NF-κB Subunit with Histone Deacetylase 1. *Mol. Cell. Biol.* **23**, 4713–4727 (2003).
211. Shou, Y. *et al*. NF-κB-mediated up-regulation of Bcl-Xs and Bax contributes to cytochrome c release in cyanide-induced apoptosis. *J. Neurochem.* **81**, 842–852 (2002).
212. Kucharczak, J. *et al*. To be, or not to be: NF-κB is the answer - Role of Rel/NF-κB in the regulation of apoptosis. *Oncogene* vol. 22 8961–8982 (2003).
213. Han, Z. *et al*. AP-1 and NF-kappaB regulation in rheumatoid arthritis and murine collagen-induced arthritis. *Autoimmunity* **28**, 197–208 (1998).
214. Davignon, J. L. *et al.* Targeting monocytes/macrophages in the treatment of rheumatoid arthritis. *Rheumatology (Oxford)*. **52**, 590–598 (2013).
215. Novack, D. V. Role of NF-κB in the skeleton. *Cell Res.* **21**, 169–182 (2011).
216. Teng, M. W. L. *et al.* IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases. *Nat. Med.* **21**, 719–729 (2015).
217. Wei, F., Chang, Y. & Wei, W. The role of BAFF in the progression of rheumatoid arthritis. *Cytokine* **76**, 537–544 (2015).
218. Greten, F. R. *et al.* IKKbeta links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell* **118**, 285–296 (2004).
219. Zaph, C. *et al.* Epithelial-cell-intrinsic IKK-beta expression regulates intestinal immune homeostasis. *Nature* **446**, 552–556 (2007).
220. Goverman, J. Autoimmune T cell responses in the central nervous system. *Nat. Rev. Immunol.* **9**, 393–407 (2009).
221. Brüstle, A. *et al.* The NF-κB regulator MALT1 determines the encephalitogenic potential of Th17 cells. *J. Clin. Invest.* **122**, 4698–4709 (2012).
222. van Loo, G. *et al.* Inhibition of transcription factor NF-kappaB in the central nervous system ameliorates autoimmune encephalomyelitis in mice. *Nat. Immunol.* **7**, 954–961 (2006).
223. Yu, X. H., Zheng, X. L. & Tang, C. K. Nuclear Factor-κB Activation as a Pathological Mechanism of Lipid Metabolism and Atherosclerosis. *Adv. Clin. Chem.* **70**, 1–30 (2015).
224. Libby, P., Ridker, P. M. & Hansson, G. K. Progress and challenges in translating the biology of atherosclerosis. *Nature* **473**, 317–325 (2011).
225. Picard, C., Casanova, J. L. & Puel, A. Infectious diseases in patients with IRAK-4, MyD88,

NEMO, or IκBα deficiency. *Clin. Microbiol. Rev.* **24**, 490–497 (2011).

226. Li, Q. *et al.* Enhanced NF-kappaB activation and cellular function in macrophages lacking IkappaB kinase 1 (IKK1). *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12425–12430 (2005).
227. Notarangelo, L. D. Functional T cell immunodeficiencies (with T cells present). *Annu. Rev. Immunol.* **31**, 195–225 (2013).
228. Pannicke, U. *et al.* Deficiency of Innate and Acquired Immunity Caused by an IKBKB Mutation . *N. Engl. J. Med.* **369**, 2504–2514 (2013).
229. Smahi, A. *et al.* Genomic rearrangement in NEMO impairs NF-kappaB activation and is a cause of incontinentia pigmenti. The International Incontinentia Pigmenti (IP) Consortium. *Nature* **405**, 466–472 (2000).
230. Williams, B. R. Signal integration via PKR. *Sci. STKE* **2001**, (2001).
231. LaMonica, R. *et al.* VP4 differentially regulates TRAF2 signaling, disengaging JNK activation while directing NF-kappa B to effect rotavirus-specific cellular responses. *J. Biol. Chem.* **276**, 19889–19896 (2001).
232. Flory, E. *et al.* Influenza virus-induced NF-kappaB-dependent gene expression is mediated by overexpression of viral proteins and involves oxidative radicals and activation of IkappaB kinase. *J. Biol. Chem.* **275**, 8307–8314 (2000).
233. Flory, E. *et al.* Plasma membrane-targeted Raf kinase activates NF-kappaB and human immunodeficiency virus type 1 replication in T lymphocytes. *J. Virol.* **72**, 2788–2794 (1998).
234. Bour, S. *et al*. The human immunodeficiency virus type 1 Vpu protein inhibits NF-kappa B activation by interfering with beta TrCP-mediated degradation of Ikappa B. *J. Biol. Chem.* **276**, 15920–15928 (2001).
235. Kim, H. *et al*. Through induction of juxtaposition and tyrosine kinase activity of Jak1, X-gene product of hepatitis B virus stimulates Ras and the transcriptional activation through AP-1, NF-kappaB, and SRE enhancers. *Biochem. Biophys. Res. Commun.* **286**, 886–894 (2001).
236. Dreyfus, D. H. *et al.* Analysis of an ankyrin-like region in Epstein Barr Virus encoded (EBV) BZLF-1 (ZEBRA) protein: Implications for interactions with NF-κB and p53. *Virol. J.* **8**, (2011).
237. Yoshida, H. *et al.* Hepatitis C virus core protein activates nuclear factor kappa B-dependent signaling through tumor necrosis factor receptor-associated factor. *J. Biol. Chem.* **276**, 16399–16405 (2001).
238. Gong, G. *et al*. Human hepatitis C virus NS5A protein alters intracellular calcium levels, induces oxidative stress, and activates STAT-3 and NF-kappa B. *Proc. Natl. Acad. Sci. U. S.*

*A*. **98**, 9599–9604 (2001).

239. Chen, Y. *et al.* The hepatitis C virus protein NS3 suppresses TNF-α-stimulated activation of NF-κB by targeting LUBAC. *Sci. Signal.* **8**, (2015).
240. Tolani, B. *et al*. NEMO is essential for Kaposi's sarcoma-associated herpesvirus-encoded vFLIP K13-induced gene expression and protection against death receptor-induced cell death, and its N-terminal 251 residues are sufficient for this process. *J. Virol.* **88**, 6345–6354 (2014).
241. Xiao, G. *et al.* Retroviral oncoprotein Tax induces processing of NF-kappaB2/p100 in T cells: evidence for the involvement of IKKalpha. *EMBO J.* **20**, 6805–6815 (2001).
242. Fu, D. X. *et al*. Human T-lymphotropic virus type I tax activates I-kappa B kinase by inhibiting I-kappa B kinase-associated serine/threonine protein phosphatase 2A. *J. Biol. Chem.* **278**, 1487–1493 (2003).
243. Bowie, A. *et al.* A46R and A52R from vaccinia virus are antagonists of host IL-1 and toll-like receptor signaling. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10162–10167 (2000).
244. Dixon, L. K. *et al.* African swine fever virus proteins involved in evading host defence systems. *Vet. Immunol. Immunopathol.* **100**, 117–134 (2004).
245. Oie, K. L. & Pickup, D. J. Cowpox virus and other members of the orthopoxvirus genus interfere with the regulation of NF-kappaB activation. *Virology* **288**, 175–187 (2001).
246. Puelles, V. G. *et al.* Multiorgan and Renal Tropism of SARS-CoV-2. *N. Engl. J. Med.* **383**, 590–592 (2020).
247. Thorne, L. G. *et al.* SARS-CoV-2 sensing by RIG-I and MDA5 links epithelial infection to macrophage inflammation. *EMBO J.* **40**, (2021).
248. Fung, S. Y. *et al*. SARS-CoV-2 main protease suppresses type I interferon production by preventing nuclear translocation of phosphorylated IRF3. *Int. J. Biol. Sci.* **17**, 1547 (2021).
249. Paterson, R. W. *et al.* The emerging spectrum of COVID-19 neurology: clinical, radiological and laboratory findings. *Brain* **143**, 3104–3120 (2020).
250. Wenzel, J. *et al.* The SARS-CoV-2 main protease Mpro causes microvascular brain pathology by cleaving NEMO in brain endothelial cells. *Nat. Neurosci.* **24**, 1522 (2021).
251. Erickson, M. A. *et al*. Interactions of SARS-CoV-2 with the Blood–Brain Barrier. *Int. J. Mol. Sci.* **22**, 1–28 (2021).
252. Mantovani, A. *et al*. Cancer-related inflammation. *Nat. 2008 4547203* **454**, 436–444 (2008).
253. Staudt, L. M. Oncogenic activation of NF-kappaB. *Cold Spring Harbor perspectives in biology* vol. 2 (2010).
254. Wertz, I. E. & Dixit, V. M. Signaling to NF-kappaB: regulation by ubiquitination. *Cold*

*Spring Harbor perspectives in biology* vol. 2 (2010).

255. Chapman, M. A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471**, 467–472 (2011).
256. Cabannes, E. *et al*. Mutations in the IkBa gene in Hodgkin's disease suggest a tumour suppressor role for IκBα. *Oncogene 1999 1820* **18**, 3063–3070 (1999).
257. Huang, B. *et al*. Posttranslational modifications of NF-κB: Another layer of regulation for NF-κB signaling pathway. *Cellular Signalling* vol. 22 1282–1290 (2010).
258. Nowicka, A. M. *et al.* A novel pVHL-independent but NEMO-driven pathway in renal cancer promotes HIF stabilization. *Oncogene 2016 3524* **35**, 3125–3138 (2015).
259. Kondylis, V. *et al.* NEMO Prevents Steatohepatitis and Hepatocellular Carcinoma by Inhibiting RIPK1 Kinase Activity-Mediated Hepatocyte Apoptosis. *Cancer Cell* **28**, 582–598 (2015).
260. Schneider, G. & Krämer, O. H. NFKB/p53 crosstalk-a promising new therapeutic target. *Biochimica et Biophysica Acta - Reviews on Cancer* vol. 1815 90–103 (2011).
261. Frank, A. K. *et al.* The Codon 72 Polymorphism of p53 Regulates Interaction with NF- B and Transactivation of Genes Involved in Immunity and Inflammation. *Mol. Cell. Biol.* **31**, 1201–1213 (2011).
262. Kawauchi, K. *et al*. p53 regulates glucose metabolism through an IKK-NF-κB pathway and inhibits cell transformation. *Nat. Cell Biol.* **10**, 611–618 (2008).
263. Schneider, G. *et al.* Cross talk between stimulated NF-κB and the tumor suppressor p53. *Oncogene* **29**, 2795–2806 (2010).
264. Lu, H. *et al.* TNF-α promotes c-REL/ΔNp63α interaction and TAp73 dissociation from key genes that mediate growth arrest and apoptosis in head and neck cancer. *Cancer Res.* **71**, 6867–6877 (2011).
265. Kawahara, T. L. A. *et al.* SIRT6 Links Histone H3 Lysine 9 Deacetylation to NF-κB-Dependent Gene Expression and Organismal Life Span. *Cell* **136**, 62–74 (2009).
266. Porta, C. *et al.* Tolerance and M2 (alternative) macrophage polarization are related processes orchestrated by p50 nuclear factor κB. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14978–14983 (2009).
267. Yu, H., Pardoll, D. & Jove, R. STATs in cancer inflammation and immunity: A leading role for STAT3. *Nature Reviews Cancer* vol. 9 798–809 (2009).
268. Inta, I. *et al.* Bim and Noxa are candidates to mediate the deleterious effect of the NF-κB subunit RelA in cerebral ischemia. *J. Neurosci.* **26**, 12896–12903 (2006).
269. Wing-Man Ho, J. *et al.* UCP4 is a target effector of the NF-κB c-Rel prosurvival pathway

against oxidative stress. *Free Radic. Biol. Med.* **53**, 383–394 (2012).

270. Stephenson, J. *et al*. Inflammation in CNS neurodegenerative diseases. *Immunology* **154**, 204–219 (2018).
271. Tilstra, J. S. *et al.* NF-κB inhibition delays DNA damage - Induced senescence and aging in mice. *J. Clin. Invest.* **122**, 2601–2612 (2012).
272. Lanzillotta, A. *et al.* The acetylation of RelA in Lys310 dictates the NF-κB-dependent response in post-ischemic injury. *Cell Death Dis.* **1**, (2010).
273. Lanzillotta, A. *et al.* NF-κB in innate neuroprotection and age-related neurodegenerative diseases. *Front. Neurol.* **6**, 98 (2015).
274. Caviedes, A. *et al*. BDNF/NF-κB Signaling in the Neurobiology of Depression. *Curr. Pharm. Des.* **23**, (2017).
275. Cui, J. *et al.* Chronic Fluoxetine Treatment Upregulates the Activity of the ERK1/2-NF-κB Signaling Pathway in the Hippocampus and Prefrontal Cortex of Rats Exposed to Forced-Swimming Stress. *Med. Princ. Pract.* **25**, 539–547 (2016).
276. Saha, R. N., Liu, X. & Pahan, K. Up-regulation of BDNF in astrocytes by TNF-α: A case for the neuroprotective role of cytokine. in *Journal of Neuroimmune Pharmacology* vol. 1 212–222 (J Neuroimmune Pharmacol, 2006).
277. Ramirez, S. H. *et al.* Neurotrophins prevent HIV Tat-induced neuronal apoptosis via a nuclear factor-κB (NF-κB)-dependent mechanism. *J. Neurochem.* **78**, 874–889 (2001).
278. Pizzi, M. & Spano, P. F. Distinct roles of diverse nuclear factor-κB complexes in neuropathological mechanisms. *European Journal of Pharmacology* vol. 545 22–28 (2006).
279. Cuccurazzu, B. *et al.* Upregulation of mGlu2 receptors via NF-κB p65 acetylation is involved in the proneurogenic and antidepressant effects of acetyl-L-carnitine. *Neuropsychopharmacology* **38**, 2220–2230 (2013).
280. Durand, J. K. & Baldwin, A. S. Targeting IKK and NF-κB for Therapy. *undefined* **107**, 77–115 (2017).
281. Llona-Minguez, S., Baiget, J. & Mackay, S. P. Small-molecule inhibitors of IκB kinase (IKK) and IKK-related kinases. *http://dx.doi.org/10.4155/ppa.13.31* **2**, 481–498 (2013).
282. Anthony, N. G. *et al.* Inhibitory Kappa B Kinase α (IKKα) Inhibitors That Recapitulate Their Selectivity in Cells against Isoform-Related Biomarkers. *J. Med. Chem.* **60**, 7043–7066 (2017).
283. Smyth, L. A. & Collins, I. Measuring and interpreting the selectivity of protein kinase inhibitors. *J. Chem. Biol.* **2**, 131–151 (2009).
284. Strickson, S. *et al.* The anti-inflammatory drug BAY 11-7082 suppresses the MyD88-

dependent signalling network by targeting the ubiquitin system. *Biochem. J.* **451**, 427–437 (2013).

285. Wang, J. & Gray, N. S. SnapShot: Kinase Inhibitors II. *Mol. Cell* **58**, 710-710.e1 (2015).
286. MacMaster, J. F. *et al.* An inhibitor of IkappaB kinase, BMS-345541, blocks endothelial cell adhesion molecule expression and reduces the severity of dextran sulfate sodium-induced colitis in mice. *Inflamm. Res.* **52**, 508–511 (2003).
287. Byun, M. S., Choi, J. & Jue, D. M. Cysteine-179 of IkappaB kinase beta plays a critical role in enzyme activation by promoting phosphorylation of activation loop serines. *Exp. Mol. Med.* **38**, 546–552 (2006).
288. Bernier, M. *et al.* Binding of manumycin A inhibits IkappaB kinase beta activity. *J. Biol. Chem.* **281**, 2551–2561 (2006).
289. Dong, T. *et al.* Ainsliadimer A selectively inhibits IKKα/β by covalently binding a conserved cysteine. *Nat. Commun.* **6**, (2015).
290. Kopp, E. & Ghosh, S. Inhibition of NF-kappa B by sodium salicylate and aspirin. *Science* **265**, 956–959 (1994).
291. Mladenova, D. *et al.* Sulindac activates NF-κB signaling in colon cancer cells. *Cell Commun. Signal.* **11**, (2013).
292. May, M. J. *et al.* Selective inhibition of NF-kappaB activation by a peptide that blocks the interaction of NEMO with the IkappaB kinase complex. *Science* **289**, 1550–1554 (2000).
293. Grothe, K. et al. IκB kinase inhibition as a potential treatment of osteoarthritis - results of a clinical proof-of-concept study. Osteoarthr. Cartil. **25**, 46–52 (2017).
294. Zhao, Y. et al. Simultaneous targeting therapy for lung metastasis and breast tumor by blocking the NF-κB signaling pathway using Celastrol-loaded micelles. Drug Deliv. **25**, 341–352 (2018).
295. Verstrepen, L. & Beyaert, R. Receptor proximal kinases in NF-κB signaling as potential therapeutic targets in cancer and inflammation. Biochem. Pharmacol. **92**, 519–529 (2014).
296. Pasparakis, M. Regulation of tissue homeostasis by NF-kappaB signalling: implications for inflammatory diseases. Nat. Rev. Immunol. **9**, 778–788 (2009).
297. Yang, J. et al. Conditional ablation of Ikkb inhibits melanoma tumor development in mice. J. Clin. Invest. **120**, 2563–2574 (2010).
298. Zhong, Z. et al. NF-κB Restricts Inflammasome Activation via Elimination of Damaged Mitochondria. Cell **164**, 896–910 (2016).
299. Hsu, L. C. et a*l.* IL-1β-driven neutrophilia preserves antibacterial defense in the absence of the kinase IKKβ. *Nat. Immunol.* **12**, 144–150 (2011).

300. McLoed, A. G. *et al.* Neutrophil-Derived IL-1β Impairs the Efficacy of NF-κB Inhibitors against Lung Cancer. *Cell Rep.* **16**, 120–132 (2016).
301. Pallangyo, C. K., Ziegler, P. K. & Greten, F. R. IKKβ acts as a tumor suppressor in cancer-associated fibroblasts during intestinal tumorigenesis. *J. Exp. Med.* **212**, 2253–2266 (2015).
302. Carl Roth. Autoinduction Media. https://www.carlroth.com/
303. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
304. van Oers, M. M. Opportunities and challenges for the baculovirus expression system. *J. Invertebr. Pathol.* **107**, S3–S15 (2011).
305. Siligardi, G. & Hussain, R. CD spectroscopy: an essential tool for quality control of protein folding. *Methods Mol. Biol.* **1261**, 255–276 (2015).
306. Muller, M. & Demeret, C. CCHCR1 Interacts Specifically with the E2 Protein of Human Papillomavirus Type 16 on a Surface Overlapping BRD4 Binding. *PLoS One* **9**, e92581 (2014).
307. Remy, I. & Michnick, S. W. A highly sensitive protein-protein interaction assay based on Gaussia luciferase. *Nat. Methods 2006 312* **3**, 977–979 (2006).
308. Tannous, B. A. *et al*. Codon-Optimized Gaussia Luciferase cDNA for Mammalian Gene Expression in Culture and in Vivo. *Mol. Ther.* **11**, 435–443 (2005).
309. Krauss, I. R. *et al.* An overview of biological macromolecule crystallization. *Int. J. Mol. Sci.* **14**, 11643–11691 (2013).
310. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
311. McCoy, A. J. *et al.* Phaser crystallographic software. *urn:issn:0021-8898* **40**, 658–674 (2007).
312. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D, Struct. Biol.* **75**, 861–877 (2019).
313. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
314. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *urn:issn:0907-4449* **66**, 486–501 (2010).
315. Scheuermann, T. H. & Brautigam, C. A. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *undefined* **76**, 87–98 (2015).
316. Brautigam, C. A., Zhao, H., Vargas, C., Keller, S. & Schuck, P. Integration and global

analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat. Protoc.* **11**, 882–894 (2016).

317. O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* **25**, 1000–1008 (2018).
318. Hoshino, K. *et al.* IkappaB kinase-alpha is critical for interferon-alpha production induced by Toll-like receptors 7 and 9. *Nature* **440**, 949–953 (2006).
319. Christian, F., Smith, E. L. & Carmody, R. J. The Regulation of NF-κB Subunits by Phosphorylation. *Cells 2016, Vol. 5, Page 12* **5**, 12 (2016).
320. Gates-Tanzer, L. T. & Shisler, J. L. Cellular FLIP long isoform (cFLIPL)–IKKα interactions inhibit IRF7 activation, representing a new cellular strategy to inhibit IFNα expression. *J. Biol. Chem.* **293**, 1745–1755 (2018).
321. Marié, I. *et al*. Phosphorylation-induced dimerization of interferon regulatory factor 7 unmasks DNA binding and a bipartite transactivation domain. *Mol. Cell. Biol.* **20**, 8803–8814 (2000).
322. Caillaud, A. *et al*. Regulatory serine residues mediate phosphorylation-dependent and phosphorylation-independent activation of interferon regulatory factor 7. *J. Biol. Chem.* **280**, 17671–17677 (2005).
323. tenOever, B. R. *et al.* Activation of TBK1 and IKKepsilon kinases by vesicular stomatitis virus infection and the role of viral ribonucleoprotein in the development of interferon antiviral immunity. *J. Virol.* **78**, 10636–10649 (2004).
324. Civas, A. *et al*. Promoter organization of the interferon-A genes differentially affects virus-induced expression and responsiveness to TBK1 and IKKepsilon. *J. Biol. Chem.* **281**, 4856–4866 (2006).
325. Mercurio, F. *et al.* IKK-1 and IKK-2: Cytokine-activated IκB kinases essential for NF-κB activation. *Science (80-. ).* **278**, 860–866 (1997).
326. Müller, F., Graziadei, A. & Rappsilber, J. Quantitative Photo-crosslinking Mass Spectrometry Revealing Protein Structure Response to Environmental Changes. *Anal. Chem.* **91**, 9041–9048 (2019).
327. Mädler, S. et al. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* **44**, 694–706 (2009).
328. Ledsgaard, L. et al. H. Basics of Antibody Phage Display Technology. *Toxins (Basel).* **10**, (2018).

# The E2F4/p130 repressor complex cooperates with oncogenic ΔNp73α to promote cell survival in human papillomavirus 38-transformed keratinocytes and in cancer cells

Valerio Taverniti [1#], Hanna Krynska [1,2#], Assunta Venuti [1], Marie-Laure Straub [2], Cécilia Sirand [1], Eugenie Lohmann [1], Maria Carmen Romero-Medina [1], Stefano Moro [2], Alexis Robitaille [1], Luc Negroni [3], Denise Martinez-Zapien [2], Murielle Masson [2], Massimo Tommasino [1 *], Katia Zanier [2 *]

[1] International Agency for Research on Cancer (IARC), World Health Organization, 150 Cours Albert Thomas, 69008 Lyon, France

[2] Biotechnology and Cell Signaling (CNRS/Université de Strasbourg, UMR 7242), Ecole Superieure de Biotechnologie de Strasbourg, Boulevard Sébastien Brant, BP 10413, F-67412 Illkirch, France.

[3] Proteomics platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) / INSERM U964 / CNRS UMR 7104/ Université de Strasbourg, 1 rue Laurent Fries, BP 10142, F-67404 Illkirch, France.

[#] These authors contributed equally to the work

[*] Correspondence to: zanier@unistra.fr, tommasinom@iarc.fr

KEYWORDS : p53, ΔNp73α, E2F4, protein-protein interactions, gene expression, cellular transformation

## ABSTRACT

Tumor suppressor p53 and its related proteins, p63 and p73, can be synthesized as multiple isoforms lacking part of the N- or C-terminal regions. Specifically, high expression of the ΔNp73α isoform is notoriously associated with various human malignancies characterized by poor prognosis. This isoform is also accumulated by oncogenic viruses such as Epstein–Barr virus (EBV), as well as genus beta human papillomaviruses (HPV) that appear to be involved in carcinogenesis. To gain additional insight into ΔNp73α mechanisms, we have performed proteomics analyses using beta-HPV type 38-transformed keratinocytes (38HK) as an experimental model. We find that ΔNp73α associates with the E2F4/p130 repressor complex through a direct interaction with E2F4. Remarkably, this interaction is favored by the N-terminal truncation of p73 characteristic of ΔNp73 isoforms and is independent of the C-terminal splicing status. We show that the ΔNp73α-E2F4/p130 complex inhibits the expression of specific genes, including genes encoding for negative regulators of proliferation, both in 38HK and in HPV-negative cancer-derived cell lines. Consistently, silencing of E2F4 in 38HK and in cancer cells resulted in induction of senescence. In conclusion, we have identified and characterized a novel transcriptional regulatory complex that exerts pro-proliferative functions in transformed cells.

## AUTHOR SUMMARY

ΔNp63 and ΔNp73 isoforms function as antagonists of tumor suppressor p53 signaling by promoting cellular proliferation. Accumulation of these isoforms, which is associated with carcinogenicity and chemoresistance, can result from infection by well-known oncogenic viruses such as EBV and HPV. Our study focuses on the highly carcinogenic ΔNp73α isoform and uses HPV 38-transformed keratinocytes as a cellular model system. We find that ΔNp73α

physically interacts with the E2F4/p130 repressor transcriptional complex. This, in turn, results in a functional cooperation between the two transcription factors, leading to the inhibition of the expression of specific genes, including genes encoding for negative regulators of proliferation.

## INTRODUCTION

The p53 family of proteins plays a key role in cancer prevention. It consists of three homologous transcription factors (TFs), namely p53, p63, and p73, which bind to common DNA promoter sites (p53RE). The TP53, TP63, and TP73 genes can be expressed as multiple isoforms lacking part of the C- or the N-terminal regions. In the case of TP73, the full-length TAp73 protein is transcribed from the P1 promoter upstream of exon 1, whereas ΔNp73 isoforms are generated from the P2 promoter within intron 3. These isoforms lack the N-terminal transactivation domain (TAD) and consequently act as dominant-negative inhibitors of full-length TAp73 and of its TAp53 and TAp63 homologues [1–3]. Interestingly, the ΔN P2 promoter contains a p53 response element (RE) that enables TAp53 and TAp73 to induce expression of ΔN isoforms [4], thus creating a negative feedback loop that fine-tunes p53 and p73 functions. Additional isoforms of TAp73 and ΔNp73 proteins (i.e. α, β, γ, δ, ε, ζ, θ, η, and η1 isoforms) are generated by alternative splicing within the 3′ region of the gene (exons 10–14), giving rise to differences in the C-terminal regions of the proteins [5,6].

Expression of the TA and ΔN proteins is highly cell- and tissue-specific. Whereas the ΔN isoforms support proliferation, the TA proteins promote cell-cycle arrest, senescence, and apoptosis, suggesting that the ratio between TA and ΔN proteins determines cell fate and oncogenesis [7]. In healthy cells, where ΔNp73 levels are low, c-Abl phosphorylation coupled to Pin1 isomerase binding results in the stabilization of TAp73 [8–10], whereas ΔNp73

isoforms are rapidly degraded by other ubiquitin ligases that can discriminate between TA and ΔN isoforms [11] or *via* the calpain [12] and 20S proteasome [13] pathways.

Mutation of the TP53 gene is the most frequent genetic alteration, present in about 50% of human cancers. In contrast, the TP63 and TP73 genes are rarely mutated but are rather expressed as ΔNp63 and ΔNp73 isoforms in a wide range of human malignancies, which display unfavorable prognosis due to increased drug resistance [7,14,15]. In particular, the ΔNp73α isoform, which comprises an intact (unspliced) C-terminal region, is upregulated in several cancers harboring wild-type TP53 and TP73 genes (including breast, prostate, liver, lung, and thyroid cancer), where it inhibits drug-induced apoptosis [15]. Although the association of ΔNp73α with elevated carcinogenicity and chemoresistance can be explained by alterations of p53-regulated gene expression, it is not yet clear whether ΔNp73α could use additional mechanisms in promoting cellular transformation. Previous independent studies have shown that ΔNp73α is accumulated upon infection with the well-established oncogenic viruses, Epstein–Barr virus (EBV) [16] and other herpesviruses, such as the human cytomegalovirus [17]. Accumulation of the ΔNp73α isoform is also induced in *in vitro* and *in vivo* experimental models by cutaneous beta (β) human papillomavirus (HPV) types [18–20], which appear to be involved in the development of skin squamous cell carcinoma [21].

In this study, we searched for protein factors that modulate the functions of ΔNp73α in transformed cells. We identified and characterized a novel interaction of ΔNp73α with the E2F4-5/p130 transcriptional repressor complex that modulates the cellular gene expression in *in vitro* HPV38-transformed human keratinocytes (38HK) as well as in cancer-derived cells.

**RESULTS**

**ΔNp73α associates with the E2F4/p130 transcriptional repressor complex in 38HK.** 38HK are human keratinocytes immortalized by the E6 and E7 oncoproteins of cutaneous β-HPV38 [18]. Here we used 38HK as a cellular model to perform proteomics analyses of ΔNp73α. The entire ΔNp73α sequence was fused to the N-terminus of the tandem affinity purification (TAP) tag, cloned in a lentiviral vector under the control of the weak promoter of Moloney murine leukaemia virus [22], and stably retrotransduced in 38HK (Fig. 1A). The nuclear fraction of 38HK expressing the ΔNp73α-TAP construct was recovered, and native ΔNp73α complexes were purified using the TAP approach [23] (Fig. 1A and B). Finally, ΔNp73α binding partners were identified by mass spectrometry.

The nuclear ΔNp73α binders identified by detection of more than two unique peptides are listed in S1 Table. Bioinformatics analysis of these binders showed an enrichment of proteins involved in transcriptional and cell-cycle control (Fig. 1C). In contrast, neither the E6 nor the E7 HPV38 oncoproteins were found to be associated with ΔNp73α in these analyses.

Among the nuclear partners that displayed high specificity towards ΔNp73α-TAP compared with the TAP control, we identified the transcription factor E2F4 and the co-repressor retinoblastoma-like 2 (RBL2/p130) (S1 Table). To corroborate this finding, we performed co-immunoprecipitation (co-IP) experiments using nuclear extracts of 38HK stably expressing ΔNp73α that was N-terminally fused to an HA tag (HA-ΔNp73α). Extracts were first applied to a 50–10% sucrose gradient to allow for complex fractionation based on their molecular weight (Fig. 1D, left panel). Incubation with anti-HA antibody clearly co-precipitated endogenous E2F4 and p130 proteins together with HA-ΔNp73α in at least two sucrose fractions (Fig. 1D, lower panel, fractions 11 and 12). Consistently, incubation with an anti-E2F4 antibody led to co-IP of HA-ΔNp73α and p130 in similar sucrose gradient fractions (S1 Fig.). As an alternative approach, we applied 38HK nuclear extracts on a size exclusion chromatography column, which provides more stringent fractionation conditions as compared to sucrose

gradients. We observed that HA-ΔNp73α, E2F4 and p130 proteins co-migrate in some of the fractions eluted from the size exclusion column. Incubation of these fractions with anti-HA antibody clearly co-precipitated endogenous E2F4 and p130 proteins in some of the fractions (S2 Fig.).

Altogether these results indicate that ΔNp73α interacts with several cellular proteins, including transcriptional regulatory complexes.

**ΔNp73α establishes direct protein–protein interactions with E2F4.** E2F4 forms stable heterodimers with DP1/2 [24], whereas p73 hetero-tetramerizes with p63 [25,26]. Consistently, two unique DP1 peptides and several p63 peptides were detected in our proteomics analyses of ΔNp73α binders. To identify the direct binding partner of ΔNp73α, we performed interaction analyses using the *Gaussia princeps* protein complementation assay (GPCA) [27]. This assay, which does not use washing steps, enables detection of protein–protein interactions (PPIs) *in cellulo* with a broad range of affinities, including transient interactions. In this way, we evaluated binary interactions within a protein set that comprises ΔNp73α and ΔNp73β isoforms, TAp63α, E2F4, DP1, and p130. All proteins were cloned as N-terminal fusions to either the Gluc1 or the Gluc2 inactive fragment of the *Gaussia princeps* luciferase. Pairwise Gluc1/Gluc2 combinations were co-expressed in HEK293T cells, and the interactions were monitored by measurement of the luminescence signal. As expected, E2F4 interacts with DP1 and p130, showing the functionality of the recombinant fusion proteins (Fig. 2A). In addition, ΔNp73α interacts with E2F4 but not with p130 or DP1. Such interaction with E2F4 is observed also for the highly carcinogenic ΔNp73β isoform. On the other hand, TAp63α does not interact with any of the proteins tested, thus ruling out indirect contributions by p63 in the assembly of this complex.

Next, we repeated sucrose gradient/co-IP experiments on extracts of 38HK under conditions of E2F4 knockdown. In these experiments, we observed that knockdown of E2F4 by small interfering RNAs (siRNAs) led to an increase in the functional homologue E2F5, and conversely, knockdown of E2F5 raised E2F4 levels (Fig. 2B). Immunoprecipitation of HA-ΔNp73α from sucrose gradient fractions of 38HK treated with scramble siRNA or E2F4 siRNA led to similar recoveries of p130 and DP1 (compare Fig. 2C upper panel with S3 Fig.). In contrast, simultaneous targeting of E2F4 and E2F5 by siRNAs strongly reduced the levels of both E2F4 and E2F5 proteins without affecting the levels of HA-ΔNp73α (Fig. 2B). In this condition, the ability of HA-ΔNp73α to co-immunoprecipitate DP1 and p130 clearly decreased (Fig. 2C and 2D). These results suggest that the ΔNp73α isoform establishes direct PPIs with E2F4 and that, in conditions of E2F4 knockdown, the E2F5 homologue is upregulated and is able to interact with ΔNp73α.

We also compared E2F4 and E2F5 for binding to ΔNp73α by GPCA in HEK293T cells (S4 Fig.). Considering the lower expression levels of the Gluc2-E2F5 fusion protein as compared to Gluc2-E2F4, E2F4 and E2F5 appear to induce similar ΔNp73α binding responses. However, since we did not detect specific E2F5 peptides in the proteomics analyses of ΔNp73α, it is likely that additional factors such as expression levels or post-translational modifications (PTMs) might account for the preference towards E2F4 in 38HK transformed cells.

Overall, our binding analyses indicate that E2F4 is sufficient to interact with ΔNp73α, although additional contributions by other partners (DP1 and p130) to complex stability cannot be excluded.

**E2F4 discriminates between TAp73 and ΔNp73 isoforms.** We reproducibly observed that the E2F4 binding response obtained with full-length TAp73α in GPCA experiments was lower (approximately 50%) than that obtained with the ΔNp73α isoform (Fig. 3A and 3B). This

intriguing observation was confirmed by *in vitro* pulldown analyses that used recombinantly expressed full-length TAp73α and ΔNp73α proteins fused to the maltose binding protein (MBP) and clarified extracts of HEK293T cells overexpressing full-length E2F4 fused to a 3xFlag tag (Fig. 3B, right panel).

Next, we further characterized the interaction between E2F4 and ΔNp73α by exploring the binding properties of a panel of deletion constructs of ΔNp73α and E2F4 by GPCA (Fig. 3A). To search for the region of ΔNp73α that interacts with E2F4, we tested full-length E2F4 against progressive C-terminal deletion constructs of ΔNp73α. We observed that the C-terminal half of ΔNp73α (residues 349–587), comprising the sterile alpha motif (SAM) domain and the disordered linker and C-terminal regions, is dispensable for E2F4 binding, thus indicating that the interaction is independent of the C-terminal splicing status (α, β, γ, δ, ε, ζ, etc.) of ΔNp73 isoforms (Fig. 3C left panel, compare ΔNp73α FL with ΔN(1–348)). In contrast, deletion of the oligomerization domain (OD) strongly decreases binding, probably due to disruption of the tetrameric state and consequent loss of its avidity contributions (Fig. 3C left panel, compare ΔNp73α FL with ΔN(1–302)). Hence, the minimal ΔN(1–348) construct, which fully retains the binding properties of ΔNp73α FL, was further reduced by N-terminal deletions. Deletion of the first 13 residues specific to ΔNp73 isoforms did not affect the interaction. In contrast, deletion of the entire N-terminal disordered region (residues 1–73) reduced binding to 60%, that is to levels comparable to those of full-length TAp73α, whereas further deletion of the DNA-binding domain (DBD) did not have any effect (Fig. 3C left panel, compare TAp73, ΔNp73α FL with ΔN(14–348), ΔN(74–348), and ΔN(262–348)). These results suggest that the N-terminal disordered region of ΔNp73 isoforms contributes to E2F4 binding. However, an additional interaction site is likely to exist, which, on the basis of our binding data, appears to be located within the 262–348 region of ΔNp73 isoforms.

A similar strategy was used to search for the region of E2F4 that interacts with ΔNp73α. Deletion of residues 197–413 from the disordered C-terminus of E2F4, comprising the TAD and the pocket protein binding domain (PPBD), reduced binding activity to 60% (Fig. 3C right panel, compare E2F4 FL with E2F(1–336)), suggesting the existence of a binding site for ΔNp73 within this region. Further deletion of the C-terminus or of the DBD has only moderate effects on the interaction (Fig. 3C right panel, compare E2F4 FL with constructs E2F(1–196) and E2F(89–196), which are probably related to the lower expression levels of these constructs (Fig. 3A right panel). In contrast, deletion of the DP-binding and dimerization (DP-dim) domain decreased binding to 20% (Fig. 3C right panel, compare E2F4 FL with E2F(1–90)). This domain of E2F4 provides avidity contributions but, unlike the OD of ΔNp73α, is also a well-known region for PPIs. Therefore, a contribution of this domain towards ΔNp73 binding cannot be excluded.

To further corroborate these findings, we produced a minimal E2F4/DP1 heterodimer by co-expression in bacteria, which comprised the DP-dim domain and C-terminal region of E2F4 and the DP-dim domain of DP1. Pulldown analyses reproducibly show higher binding activity of the E2F4/DP1 heterodimer towards recombinant MBP-fused ΔN(1-433) (equivalent to the ΔNp73β isoform) and ΔN(1-348) constructs compared to the MBP control condition (Fig. 3D). Taken together, our analyses suggest that the interaction between ΔNp73 and E2F4 is mediated by at least two binding interfaces. One of these interfaces involves an amino acid sequence within the N-terminal disordered region of ΔNp73. The lower binding activity of TAp73 compared with ΔNp73 suggests that, in the context of TAp73, such a sequence is inaccessible to E2F4, probably due to intramolecular interactions (Fig. 3E).

**Expression and regulation of ΔNp73α-E2F4/p130 complex components in transformed *versus* untransformed cells.** We previously showed that expression of HPV38 E6 and E7

oncoproteins induces the accumulation of ΔNp73α in human primary keratinocytes (HPKs) (18). Here, we also evaluated the expression of E2F4 and p130 in primary keratinocytes (HPK) and in HPV38 E6/E7-transformed keratinocyte cell lines (38HK) at both early and late passage. As expected, the levels of oncogenic ΔNp73α were higher in all immortalized 38HK cell lines compared to HPKs (Fig. 4A). In line with previous data [18], detection of ΔNp73α in the early passage 38HK(D2) cell line indicates that accumulation of this isoform occurs after acute infection by the HPV38 virus. The E2F4 and p130 proteins are instead expressed in both primary HPK and transformed 38HK cell lines. For p130, in addition to the characteristic band at 130 kDa, we can observe a band migrating at about 150 kDa in the HPK and early passage 38HK(D2) cell lines, possibly corresponding to ubiquitinated or SUMOylated forms of this protein. Such band at 150 kDa is not present for the late passage 38HK(D3) cell line, which was used throughout this study. In line with these results, analysis of the cancer database revealed that E2F4, and, albeit at a lower extent, the E2F5 homologue, as well as DP1 and p130 are similarly expressed in both normal and cancer tissues from different anatomical sites (S5 Fig.).

It is well-established that the HPV E7 oncoprotein binds to RB family members, thereby inhibiting their interactions with E2F factors. In addition to acting as a steric inhibitor, HPV E7 also promotes RB protein degradation *via* the proteasome. Here, we observe that the E7 oncoprotein from HPV38 (38.E7) binds to p130, as well as to pRb and p107 (Fig. 4B). However, in spite of its high expression levels in HEK293T cells, 38.E7 induces lower binding responses to all three RB proteins compared to the 16.E7 oncoprotein from the "high-risk" HPV16 virus (Fig. 4B), which are indicative of weaker binding affinities. Treatment of 38HK with the proteasome inhibitor MG132 increases both ΔNp73α and p130 levels by about 70% as compared to the DMSO control, showing that these proteins are only partially degraded in 38HK (Fig. 4C). By contrast, E2F4 levels do not change upon treatment with the proteasome

inhibitor. Furthermore, results from cellular fractionation experiments indicate that ΔNp73α and p130 mostly localize to the nucleus of 38HK, with about half of these proteins in the nuclear soluble fraction (i.e. nuclear(S)) and half bound to chromatin (Fig. 4D). Together, these observations strongly suggest that in 38HK p130 is partially protected from the interaction with 38E7 and/or degradation by binding to ΔNp73α, which enhances its association to chromatin. To conclude, all core components of the ΔNp73α-E2F4/p130 complex are expressed in the transformed 38HKs and in cancer cells lines over-expressing ΔNp73α isoform.

**E2F4-5 are required for proliferation of 38HK.** We reasoned that the E2F4/p130 complex might acquire oncogenic functions by partnership with ΔNp73α, which is essential for proliferation of 38HK [18]. Indeed, we observed that E2F4-5 knockdown by siRNAs in 38HK induced a different phenotype compared with the negative control (scramble siRNA), with cells appearing flattened and larger than normal 38HK. This phenotype is a feature of premature senescence. Consistently, assessment of senescence-associated β-galactosidase activity in 38HK showed that simultaneous E2F4-5 knockdown induced an increase in the number of senescent cells by a factor of two at 48 hours after transfection (Fig. 5A). Subsequently, we analyzed 38HK in conditions of single E2F4 or E2F5 knockdown and in conditions of double E2F4-5 knockdown for the presence of the heterochromatin-associated histone mark H3K9Me3 (tri-methylation of Lys9 on histone H3). This analysis enables detection of senescence-associated heterochromatin foci (SAHF), which correspond to heterochromatic regions that are thought to inhibit the expression of proliferation-promoting genes at E2F *loci* [28]. Single knockdown of either E2F4 or E2F5 by siRNAs induces a two-fold increase in the percentage of cells positive for H3K9me3 marks as compared to the scramble treated control (i.e. from 10% to about 20%) (Fig. 5B). Similar results were obtained using shRNAs targeting the 3' UTR of either the E2F4 or of the E2F5 gene, thereby ruling out the possibility that the senescence

phenotype could be an off-target effect of the knockdown approach (S6 Fig.). Remarkably, more efficient depletion by the double E2F4-5 knockdown led to a four-fold increase in the percentage of H3K9me3 positive cells as compared to the scramble treated control (i.e. from 10% to about 40% - Fig. 5B).

These results show that E2F4-5 depletion induces premature senescence in 38HK and are indicative of pro-proliferative functions for these proteins in transformed cells.

**E2F4-5/p130 cooperate with ΔNp73α to inhibit gene expression.** To evaluate the interplay of E2F4-5 and ΔNp73α in gene expression, we first identified genes regulated by E2F4-5, by performing mRNA-seq analyses on 38HK treated with scramble siRNA or with E2F4-5 siRNA. E2F4-5 knockdown reduced the levels of E2F4 and E2F5 proteins by 80% and 45%, respectively (S7A Fig.). This, in turn, led to the upregulation of 2802 genes and to the downregulation of 2110 genes (adjusted *P*-value $\leq 0.05$) (Fig. 6A and S2 and S3 Tables). Gene set enrichment analysis of upregulated genes identified biological processes related to cell-cycle, tumor suppressor signaling, cancer, and viral infection pathways (Fig. 6B left panel). In contrast, downregulated genes were enriched in pathways linked to metabolism and DNA replication (Fig. 6B right panel). In agreement with the senescence phenotype, among the most significantly downregulated genes we can observe cyclin A (CCNA1) and MCM3, which are typically repressed upon formation of SAHF *foci* [28].

Hence, we focused on the genes that are upregulated in conditions of E2F4-5 knockdown, which could be direct targets of the ΔNp73α-E2F4-5/p130 transcriptional complex. We thus selected 23 upregulated genes for validation. In this set, 16 genes are predicted to host REs recognized by p53/p73 TFs within their long promoter regions (defined as 2500 nucleotides upstream of the transcription start site [TSS]) (S4 and S5 Tables). Classical RT-qPCR experiments in 38HK showed that 16 of the 23 genes had significantly higher mRNA levels

under the E2F4-5 knockdown condition than under the scramble condition (Fig. 6C right panel, E2F4-5 siRNA/scramble > 1 and S5 Table), thereby confirming the RNA-seq findings on these genes. In this analysis, we also included 2 downregulated genes (ADAM9 and ANXA10), whose expression was inhibited in condition of E2F4-5 depletion (Fig. 6C left panel, E2F4-5 siRNA/scramble < 1).

Next, we used antisense (AS) oligonucleotides to knock down ΔNp73α expression in 38HK. Notably, knockdown of ΔNp73α induces apoptosis and for this reason a compromise between cell viability and protein depletion has to be searched. Here, we succeeded to decrease ΔNp73α protein levels by about 50%, while leaving E2F4 levels unchanged (Fig. 6D and S7B Fig.). We then checked the effects of ΔNp73α knockdown on the expression of the validated genes. Whereas ADAM9 and ANXA10 did not display significant variations (Fig. 6C left panel), 8 genes from the upregulated and validated set had significantly higher mRNA levels upon ΔNp73α knockdown (Fig. 6C right panel, AS/S ΔNp73α > 1, and S5 Table). These genes are STC1, PRR15L, RIC3, GAD1, MAFB, CDKN1B/p27$^{Kip1}$, CDKN2B/p15$^{INK4b}$ and CDKN2D/p19$^{INK4d}$. Interestingly, 3 of these genes had no REs for p53/p73 within their long promoter regions (i.e. MAFB, GAD1, and CDKN2D/p19$^{INK4d}$). In contrast, only one of the upregulated/validated genes (ANKRD22) had reduced mRNA levels upon ΔNp73α knockdown (Fig. 6C right panel), an effect possibly linked to indirect contributions of ΔNp73α in the regulation of this gene.

Together, these results provide evidence that E2F4/p130 and ΔNp73α cooperate to inhibit the expression of specific genes.

**The interaction of ΔNp73α with E2F4 enhances recruitment to DNA regulatory regions.**

We proceeded to evaluate the recruitment of the ΔNp73α-E2F4/p130 complex to the promoter regions of co-regulated genes by chromatin immunoprecipitation (ChIP). For this, we selected

STC1, whose promoter hosts both E2F and p53/p73 REs, and MAFB, which hosts E2F REs only (S5 Table). 38HK stably expressing HA-ΔNp73α were treated with scramble or E2F4-5 siRNAs, and nuclear extracts were processed for ChIP using anti-HA, anti-p73 and anti-E2F4 antibodies. The eluted DNA was analyzed by qPCR with primers flanking regions 1, 2, and 3 of the STC1 promoter (Fig. 7A), regions 1 and 2 of the MAFB promoter (Fig. 7B), or the negative control region reported by Lee et al. [29] (Fig. 7C).

Results show that HA-ΔNp73 and E2F4 are recruited to regions 1 and 2 of both genes, whereas recruitment of both proteins to region 3 of the STC1 gene is not significant when compared to the negative control (Fig. 7, compare scramble conditions in panels A and B with scramble condition of panel C). In addition, E2F4-5 knockdown significantly reduces the recruitment not only of E2F4 but also of HA-ΔNp73α at region 1 of the STC1 promoter and regions 1 and 2 of the MAFB promoter (Fig. 7).

These data indicate that both ΔNp73α and E2F4-5 bind to the promoter regions of co-regulated genes. Notably, ΔNp73α is recruited to the MAFB promoter even in the absence of a p53/p73 RE. This recruitment depends on the interaction of ΔNp73α with E2F4 and is lost upon knockdown of E2F4-5.

**Contributions of the ΔNp73α-E2F4/p130 complex in cancer cells.** Because several studies have highlighted the importance of ΔNp73α in different types of cancer, we evaluated whether the interaction between ΔNp73α and E2F4 occurs in cancer-derived cell lines. We previously identified an HPV-negative head and neck cancer cell line (HNC-136) and a breast cancer cell line (CAL-51), both with high levels of ΔNp73α and wild-type TP53 [20]. Nuclear extracts from HNC-136 cells were analyzed by sucrose density gradient. Immunoprecipitation of endogenous ΔNp73α using an anti-p73 antibody led to recovery of endogenous E2F4 and p130 proteins (Fig. 8A lower panel, fractions 15–17). Interestingly, ΔNp73α from HNC-136 extracts

was detected in lower-molecular-weight fractions compared with ΔNp73α from 38HK extracts (Fig. 8A upper panel), indicating that these complexes may include additional proteins that depend on the cellular context. In addition, although HNC-136 extracts exhibited expression of E2F5, no E2F5 was recovered upon ΔNp73α immunoprecipitation, suggesting that E2F4 is the preferred partner of ΔNp73α also in this cell line.

As in the case of 38HK, knockdown of E2F4-5 in both HCN-136 and CAL-51 cells (S8 Fig.) led to premature senescence (Fig. 8B), suggesting pro-proliferative functions for E2F4-5 also in cancer cells. We therefore evaluated the contributions of E2F4-5 and ΔNp73α towards gene expression in these cell lines, by an approach similar to that described above but focusing on the genes that were found to be co-regulated by both TFs in 38HK. RT-qPCR analysis showed that the expression of most of these genes was rescued upon knockdown of E2F4-5 in both HNC-136 (S9A Fig.) and CAL-51 cells (Fig. 8C, white histograms). In contrast, ΔNp73α depletion induced massive apoptosis in both cancer cell lines. However, by carefully screening a range of AS oligonucleotide concentrations, we could identify one condition for the CAL-51 cell line that resulted in a knockdown of ΔNp73α to approximately 60% at 24 hours after transfection (S9B Fig.). Remarkably, even with such a partial knockdown we could observe an increase in the expression of most genes, with 3 genes (STC1, MAFB and CDKN2D/p19$^{INK4d}$) showing statistically significantly higher mRNA levels under conditions of either E2F4-5 knockdown or ΔNp73α knockdown (Fig. 8C, grey histograms).

Together, our data show that the ΔNp73α-E2F4/p130 complex regulates the expression of specific genes in cancer cells.

**p19$^{INK4d}$ induces senescence in 38HK.** The CDKN2D/p19$^{INK4d}$ gene is repressed by both ΔNp73α and E2F4-5 in both 38HK and CAL51cell lines. A previous study showed that the p19$^{INK4d}$ protein is involved in senescence mechanisms by contributing to heterochromatin

formation [30]. To test whether similar mechanisms occur in our cellular model system, we retro-transduced $p19^{INK4d}$ or negative control pLXSN-GFP plasmid in 38HK cells, which were subsequently analyzed for senescence-associated β-galactosidase activity and SAHF *foci*. Results show that expression of $p19^{INK4d}$ induces a two-fold increase in the number of senescent cells in the β-galactosidase assay and a 3-fold increase in the number of cells positive for the H3K9me3 marker (Fig. 9). These results suggest that repression of the CDKN2D/$p19^{INK4d}$ gene by the ΔNp73α-E2F4-5/p130 prevents senescence, thus favoring cellular proliferation.

## DISCUSSION

In this study, we present results from what is, to the best of our knowledge, the first reported proteomics analysis on a ΔNp73 isoform, which has enabled us to identify the E2F4/p130 repressor complex as a nuclear partner of ΔNp73α. E2F4 belongs to the E2F family of proteins, whose members act as transcriptional activators (E2F1/2/3/6) and repressors (E2F4-5). E2F TFs associate with DP1–4 proteins through the dimerization domain and interact with the retinoblastoma (RB) family proteins (pRb, p130, p107) *via* the TAD. This interaction with RB proteins prevents recruitment of the transcriptional machinery, thereby inhibiting transcriptional activity. In the canonical model of cell-cycle progression, E2F4 and E2F5 mediate the repression of cell-cycle genes and cell-cycle arrest in the G0 phase and are part, together with p107/p130, DP1/2, and MuvB, of the DREAM complex [31,32]. Here, besides E2F4, DP1 and p130, none of the other DREAM complex subunits (i.e. LIN proteins, PAF, RBBP4) could be detected in association with ΔNp73α, suggesting that ΔNp73α-E2F4/p130 and DREAM are distinct complexes.

E2F4-5 and associated proteins (DP1 and p130) are expressed in both healthy and cancer tissue. However, our findings suggest that the interaction with ΔNp73α alters the E2F4-5 properties

in transformed cells. Although the E2F5 homologue can replace E2F4 in conditions of E2F4 depletion, E2F4 seems to be the preferred binding partner of ΔNp73α in transformed cells. Such preference might be related to the lower expression levels of E2F5 compared to E2F4, or, alternatively, to PTMs specific for E2F4 that might enhance binding affinity. In line with these findings, many previous studies have unveiled oncogenic functions for E2F4 but not for E2F5, despite the strong similarity between these two homologues [33].

Many of our analyses were performed in the 38HK model of viral transformation, in which ΔNp73α is the main expressed isoform, whereas TAp73 levels are under the detection limit [18]. Our observations indicate that in 38HK p130 is, at least in part, protected from the hijacking by the HPV38 E7 oncoprotein, most likely through the association with ΔNp73α, which enhances its association to chromatin. Consistently, neither the E6 nor the E7 viral oncoproteins are detected as partners of ΔNp73α in our proteomics analyses. The detection of a ΔNp73α-E2F4/p130 complex in extracts of the HPV-negative HNC-136 cancer cell line confirms that the ΔNp73α-E2F4 interaction is independent of the viral oncoproteins.

Results from binding analyses suggest that truncation of the TAD region of p73, giving rise to ΔN isoforms, leads to the exposure of a binding site for E2F4 and that this interaction mechanism is independent of the C-terminal splicing status of ΔNp73 isoforms. This finding strongly suggests that oncogenic ΔNp73 isoforms acquire novel PPIs with respect to TAp73, which are reminiscent of the gain-of-function PPIs established by hotspot oncogenic mutants of p53 with other TFs [34,35]. However, unlike mutant p53 interactions, which lead to activation of transcription, the interaction of ΔNp73α with the E2F4/p130 complex results in repression of gene expression.

It is well established that direct PPIs between TFs belonging to different families enhance cooperative binding to DNA and that this results in TF redirection to different sites [36–38]. Consistently, ChIP experiments investigating TF recruitment to the promoters of the STC1 and

MAFB genes showed that E2F4 enhances binding of ΔNp73α to DNA. Remarkably, the MAFB gene analyzed here does not host any p53/p73 REs in its long promoter region, thus providing evidence that the interaction between ΔNp73α and E2F4 is sufficient to redirect ΔNp73α to E2F4-regulated genes.

Redirection arising from PPIs between TFs can lead to different physiological outcomes. This is well documented in the case of p53 gain-of-function interactions, which lead to the activation of transcriptional programs related to cellular proliferation. Here, we observed that in 38HK, as well as in HNC-136 and CAL-51 cancer cell lines, E2F4-5 knockdown triggers premature senescence, which is associated to global chromatin changes and formation of SAHF *foci*. These findings are in agreement with previous studies showing that E2F4 acts as an oncogene with pro-proliferative and anti-chemotherapeutic properties (reviewed in [33]) and highly expressed in prostate [39] and breast [40] cancers. We propose that the oncogenic properties of E2F4-5 observed in this work are, at least in part, related to its partnership with ΔNp73α. Transcriptomics analyses combined with classical RT-qPCR experiments enabled us to identify eight genes that are co-regulated by ΔNp73α and E2F4-5 in 38HK. Among these genes we found three members of the cyclin-dependent kinase inhibitor (*Cdkn*) family, namely CDKN1B/$p27^{Kip1}$, CDKN2B/$p15^{INK4b}$ and CDKN2D/$p19^{INK4d}$. The products of these genes are negative regulators of cell-cycle progression by binding and inhibiting cyclin-dependent kinases CDK1/2 ($p27^{Kip1}$) and CDK4/6 ($p27^{Kip1}$, $p15^{INK4b}$ and $p19^{INK4d}$ ), and represent well-established or candidate tumor suppressors [41–43]. Multiple studies have shown that *Cdkn* genes have both redundant and compensatory roles [44,45] and that co-deletion of *Cdkn* genes in some human cancers are responsible for a worsen phenotype [46–48]. $p27^{Kip1}$ is regarded as an effector of senescence acting through the PTEN-$p27^{Kip1}$ pathway [49,50]. Of note, the CDKN1A gene coding for $p21^{cip1/warf1}$, homologue of $27^{Kip1}$ and implicated in initiation of senescence, is repressed by ΔNp73α but not by E2F4-5 in 38HK (*data not shown*). $p15^{INK4b}$

and p19$^{INK4d}$ belong instead to the INK4 family of CDK inhibitors. p15$^{INK4b}$ is transcribed together with the functional homologue p16$^{INK4a}$ from the INK4a/ARF/INK4b locus, which is deleted in a wide spectrum of tumors [51]. Several studies have shown that p15$^{INK4b}$ plays a role in the maintenance of senescence [52–54] and is considered a marker of the senescence state [55]. p19$^{INK4d}$ is co-repressed by ΔNp73α and E2F4-5 in both 38HK and CAL51 cell lines. Previous studies have highlighted the antiproliferative properties of this protein and its inactivation in human cancers [56–58]. Here, we find that expression of p19$^{INK4d}$ induces cellular senescence associated to heterochromatin formation in 38HK. This observation is in agreement with a previous report showing that p19$^{INK4d}$ associates to SAHF *foci* in another cell line [30]. Interestingly, here we find that E2F4-5 depletion also upregulates the expression of the high-mobility group HMGA2 protein (S2 Table), which was previously shown to act in a mutually reinforcing manner with *Cdkn* inhibitors to promote SAHF formation [59].

In addition to *Cdkn* genes, the STC1 and MAFB genes were also found to be co-regulated by ΔNp73α and E2F4-5 in both 38HK and in the CAL-51 cancer cell line. Stanniocalcin 1, the product of the STC1 gene, was previously reported as one of the most highly secreted SASP (senescence-associated secretory phenotype) factors in senescent cells [60]. The MAFB gene instead encodes for a TF belonging to the AP1 family, which acts both as activator and as repressor of transcription [61]. Besides its well-established role as an oncogene, MAFB has also been reported to act as a tumor suppressor in certain cellular contexts, by antagonizing oncoproteins such as HRAS and BRAF [62]. Altogether our results on the repression of specific genes in transformed cells provide evidence for oncogenic contributions of the ΔNp73α-E2F4/p130 complex in different cellular contexts. Yet, our data do not exclude that ΔNp73α or E2F4/p130 act alone or as part of other complexes exerting oncogenic functions.

Our characterization of the ΔNp73α-E2F4 interaction suggests that a main site for E2F4 binding resides in the N-ter unstructured arm of ΔNp73 isoforms and that this site is locked in

TAp73. TAp73 is known to encompass a C-terminal transcription inhibitory domain (TID), which associates with the TAD domain [63]. Deletion of the TID domain did not increase TAp73 binding to E2F4 (*data not shown*), suggesting that TID is not involved in the inhibition of the E2F4 interaction site. By contrast, two recent studies have reported an intramolecular interaction between the N-terminal TAD region and the core domain of p53, which results in the attenuation of DNA-binding activity [64,65]. Thus, a similar interaction may exist in TAp73, with the ΔN truncation resulting in the exposure of a PPI site that can engage in interactions with other TFs. Future biochemical and structural studies will provide more precise insights into the mechanisms underlying the ΔNp73α-E2F4 interaction.

In conclusion, our study provides evidence that ΔNp73 isoforms acquire novel PPIs with other TFs. The ΔNp73α partner investigated here is E2F4, which acquires oncogenic properties by supporting cell proliferation and survival. Our work suggests that the cooperation with factors such as E2F4 enhances and redirects the DNA-binding activity of ΔNp73 isoforms in cancer.

## MATERIALS AND METHODS

### DNA constructs

The ORFs encoding *human* E2F4, *human* E2F5 and *human* DP1 were purchased from Addgene, whereas the ORF encoding human p130 from Origene.

*Retroviral constructs.* The ΔNp73α-TAP construct used for retrotransduction was generated in two steps. First, the open reading frame (ORF) coding for ΔNp73α was amplified by PCR and cloned in the pCDNA-TAP vector [23], thus enabling expression of the TAP tag fused to the C-terminus of ΔNp73α. Second, the pCDNA ΔNp73α-TAP plasmid was used as a template for PCR amplification of the ΔNp73α-TAP fusion, which was cloned into pBABE-puro retroviral vector [22] (pBABE ΔNp73α-TAP). For the HA-ΔNp73α construct, the ΔNp73α ORF was

amplified with an oligo complementary to its N-terminus and containing the HA tag and directly cloned into pBABE-puro (pBABE HA-ΔNp73α). pMSCV-p19Ink4d-IRES-GFP plasmid expressing *mouse* p19INK4d (sharing 87% sequence identity with human p19INK4d ) [66] was obtained from Addgene.

*Constructs for expression in mammalian cells*. ORFs encoding for ΔNp73α, TAp73α, TAp63α, E2F4, E2F5, DP1, and p130 were amplified by PCR and cloned into the pDONR207 vector by recombination cloning (Gateway system, Invitrogen). The resulting pEntry clones were then transferred into the GPCA destination vectors pSPICA-N1 and pSPICA-N2 (both derived from the pCiNeo mammalian expression vector) and into a pcDNA3 vector enabling expression of an N-terminal 3xFlag tag. The GPCA vectors enable expression of the Gluc1 (pSPICA-N1) and Gluc2 (pSPICA-N2) complementary fragments of the *Gaussia princeps* luciferase, which are linked to the N-terminal ends of the tested proteins by a flexible 20 amino acid linker [27].

*Constructs for E.coli expression*. ORFs encoding for ΔNp73α, TAp73α and constructs of these proteins were cloned in the NcoI and KpnI sites of a modified pETM-41 vector containing an N-terminal MBP tag followed by a TEV cleavage site. The E2F4(84-413) and DP1(199-350) DNA constructs for expression of the minimal heterodimer were cloned in the NdeI and BamHI sites of the pmCS and pnEA vectors [67]. These vectors allow for fusion to N-terminal 6xHis (pmCS) and GST (pnEA) tags, thus giving rise to the 6xHis-E2F4(84-413) and GST-DP1(199-350) fusions.

*shRNA constructs for gene knockdown.* Oligonucleotides coding for shRNAs, which target the 3'-UTR of E2F4 and E2F5 genes, were chemically synthesized by Eurofins Biolab. The oligonucleotide sequences used for gene knockdown are described in S6 Table. Complementary oligonucleotide-annealed products were cloned into the NdeI and EcoRI sites of the lentiviral

expression vector pLKO.1. As a negative control a plasmid expressing the scramble sequence (MISSION pLKO.1-puro shRNA Control Plasmid DNA) was purchased from Sigma-Aldrich. All constructs were verified by DNA sequencing.

**Cell lines and cell culture**

Phoenix, NIH 3T3, HEK293T, HNC-136 and CAL-51 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% calf serum (NIH 3T3) at 37°C with 5% $CO_2$.

38HK from different donors [68] were grown together with NIH 3T3 feeder layers in FAD medium containing 3 parts Ham's F12, 1 part DMEM, 2,5% fetal calf serum, insulin (5 μg/ml), epidermal growth factor (10 ng/ml), cholera toxin (8.4 ng/ml), adenine (24 μg/ml), and hydrocortisone (0.4 μg/ml). Feeder layers were prepared by treating NIH 3T3 with mitomycin C for 2 hours. Around $3x10^5$ of treated NIH 3T3 were co-cultured with 38HK cells in T75 cell culture flasks. Feeder layers were removed by incubating the cell co-cultures with 5 ml of PBS 1X supplemented with 2mM EDTA. In this way, more than 95% of the feeder cells were removed. After feeder removal, 38HK cells were collected by scraping for analysis.

HPK cells were freshly isolated from neonatal foreskin and cultured in Keratinocyte Growth Medium 2 (PromoCell, Heidelberg, Germany).

**Cell line generation**

Retrovirus transduction system was used to generate 38HK cells stably expressing ΔNp73α-TAP, HA-ΔNp73α and p19INK4d constructs. High-titer retroviral supernatants were generated by transient transfection of Phoenix cells with the retroviral constructs described above and used to infect 38HK as described previously [69]. Briefly, 500 μl of DNA mix (10 μg plasmid DNA, 248 mM $CaCl_2$) were gently mixed to 500 μl of 2X HBS-buffered saline (1.5 mM

Na2HPO4, 50 mM HEPES, 280 mM NaCl, 10 mM KCl, 12 mM Dextrose, pH 7.05). The transfection mix was then used to transfect Phoenix cells cultured in 5 ml fresh medium supplemented with 25 μM Chloroquine for 6/8 hours. After 48 hours, the culture medium containing the retrovirus was filtered (0.2 μm filter), mixed with 5 μl of polybrene (Sigma) and used to infect 38HK cell cultures for 3 hours. 24 hours after infection, 38HK were selected in 0.2 μg/ml of puromycin for 3-5 days.

**Transfection conditions**

38HK, HNC-136, and CAL-51 cell lines were transiently transfected with siRNAs and sense/antisense oligonucleotides using Lipofectamine 2000 (Invitrogen). After 4 hours, the transfection mix was removed and the cells cultured in FAD medium (without antibiotic and cholera toxin). The sequences of oligos used for gene knockdown are described in S6 Table.

38HK were transiently transfected with lentiviral shRNA plasmids using TransIT-Keratinocytes Transfection Reagent (Mirus) according to the manufacturer's protocol. HEK293T cells were transfected using JetPEI® (Polyplus transfection).

**Proteomics**

38HK (about $2x10^8$ total cells) stably expressing ΔNp73α-TAP or TAP tag alone (negative control) were resuspended in 8 ml of cold buffer A (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5mM DTT, 0.2 mM EDTA, protease inhibitor cocktail mix 1x, 10 mM NaF) and incubated for 30 minutes on ice. Then, cells were lysed by passing the mix through a 25 gauge needle 15 times and centrifuged for 10 minutes at 13400 rpm at 4°C. The supernatant (cytoplasmic soluble fraction) was flash-frozen and stored at -80°C, while the pellet (the nuclear fraction) was resuspended in 7 ml of cold buffer B (20 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 250 mM NaCl, 20% glycerol, 0.5 mM DTT, 0.2 mM EDTA, protease inhibitor cocktail

mix 1x, 10 mM NaF) and incubated on ice for 1 hour. The mix was centrifuged for 10 minutes at 13,400 rpm at 4°C and the supernatant (nuclear soluble fraction) was collected. Each lysis step was checked by Western blot (Fig. 1b, upper panel).

Buffer A was added to the nuclear soluble fraction to reach a final concentration of 200 mM NaCl and 16% glycerol. The resulting nuclear protein extract was transferred to an ultra-clear polycarbonate tube and centrifuged at 40,000 rpm for 1 hour at 4 °C using a SW41 rotor (Beckmann). After centrifugation, the supernatant was carefully collected and incubated overnight with 100 µl (dry bead volume) of prewashed IgG Sepharose beads. Beads were then washed 3 times with 10 ml of IPP150 buffer (10 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.1% NP40) and once with 10 ml of TEV buffer (10 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.1% NP40, 0.5 mM EDTA, 1 mM DTT). Subsequently, beads were resuspended in 1 ml of TEV cleavage buffer containing 15 µl (10 U/µl) of acTEV protease (Invitrogen) and incubated with gentle agitation for 4 hours at 16 °C. Protein complexes were eluted by gravity flow. Then, 3 ml of calmodulin binding buffer (10 mM β-mercaptoethanol, 10 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 2 mM $CaCl_2$, 0.1% NP40) and 3 µl of 1 M $CaCl_2$ were added to the 1 ml eluate containing the protein complexes to chelate the EDTA present in the TEV cleavage buffer. The resulting mix was incubated with 100 µl of prewashed (dry bead volume) calmodulin-Sepharose beads (Agilent Technologies) for 2 hours with gentle agitation at 4 °C. Beads were then washed 3 times with 10 ml of calmodulin binding buffer. The protein complexes were recovered with 5 consecutive elutions (200 µl each) with calmodulin elution buffer (10 mM β-mercaptoethanol, 10 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 mM magnesium acetate, 1 mM imidazole, 2 mM EGTA, 0.1% NP40). An additional elution with 1% SDS was performed to recover all the remaining proteins (Fig. 1b, lower panel).

Elution fractions 2 from the ΔNp73α-TAP and TAP purifications were partially digested with trypsin and analyzed by LC-MS using an Orbitrap ELITE instrument equipped with a C18

Accucore 50 cm column. The generated data were analyzed using the Proteome Discoverer 2.4 tool. Proteins enriched more than 10-fold in ΔNp73α-TAP compared with the control (TAP-only) experiment are listed in S1 Table. Enrichment is calculated from the ratio of the sums of peptide peak areas in test and control experiments.

**Sucrose gradient/co-immunoprecipitation (co-IP)**

Sucrose density gradients were performed as previously described [70] with minor modifications. Briefly, step gradients were made by superposing sucrose solutions of different concentration (50%, 40%, 30%, 20%, 10%) in an ultra-clear polycarbonate tube (Beckman), and a linear gradient was allowed to form overnight at 4 °C. Then, 1.5-2 mg of nuclear extracts from 38HK HA-ΔNp73α or HNC-136 cells were carefully transferred to the top of the sucrose gradient and the protein complexes were separated based on their molecular weight by ultracentrifugation at 35,300 rpm for 16 hours at 4°C using the SW41 rotor. After centrifugation, 500 µl fractions were collected from the bottom of the tube by gravity flow.

ΔNp73α complexes were immunoprecipitated from each fraction using 30 µl of slurry HA-agarose beads (Sigma-Aldrich, ref. A2095) or 20 µl of pre-complexed p73-Sepharose beads (dry bead volume). Briefly, each fraction was incubated with the beads for 2 hours (HA-beads) or 4 hours (p73-beads). After incubation, the beads were washed 5 times with 1ml of washing buffer (20mM Tris-HCl pH 7.5, 1.5 mM $MgCl_2$, 150 mM NaCl, 0.2 mM EDTA, 0.1% Igepal). The protein complexes were eluted in 1x loading dye buffer.

**GPCA assay**

HEK293T cells were transfected using the reverse transfection method. Transfection mixes containing 100 ng of pSPICA-N2 and 100 ng of pSPICA-N1 plasmids expressing test proteins plus JetPEI® (Polyplus transfection) were dispensed in white 96-well plates. HEK293T cells

were then seeded on the DNA mixes at a concentration of 4.2 x$10^4$ cells per well. At 48 hours after transfection, cells were washed with 50 µl of PBS and lysed with 40 µl of *Renilla* lysis buffer (Promega, E2820) for 30 minutes with agitation. *Gaussia princeps* luciferase enzymatic activity was measured using a Berthold Centro LB960 luminometer by injecting 50 µl per well of luciferase substrate reagent (Promega, E2820) and counting luminescence for 10 seconds. Results are expressed as a fold change normalized over the sum of controls, specified herein as normalized luminescence ratio (NLR) [27]. For a given protein pair A/B, NLR = (Gluc1-A + Gluc2-B) / [(Gluc1-A + Gluc2) +(Gluc1 + Gluc2-B)].

**Cellular fractionation**

38HK (out 2.5x$10^7$ total cells) stably expressing HA-ΔNp73α were resuspended in 1 ml of cold buffer A (described in the proteomics section) and incubated for 15 minutes on ice. Then, cells were lysed by passing the mix through a 25 gauge needle 15 times and centrifuged for 5 minutes at 12000 rpm at 4°C. The supernatant, corresponding to the cytoplasmic soluble fraction, was recovered, while the pellet (the nuclear fraction) was resuspended in 200 µl of cold buffer B (described in the proteomics section) and incubated on ice for 30 min. The mix was centrifuged for 15 minutes at 12000 rpm at 4°C. The supernatant (nuclear soluble fraction) was recovered, while the pellet was resuspended in the chromatin digestion mix, consisting of 190µl of buffer A supplemented with 6µl of CaCl2 (0.1M), 2µl of BSA (10mg/ml) and 2µl of Micrococcal Nuclease (M0247S, New England). The mix was incubated at 4°C on rotating wheel for 2 hours. Then, the mix was lysed by passing the mix through a 27 gauge needle 15 times. Nuclease activity was stopped by addition of EGTA to a final concentration of 5.6 mM. Proteins were released from the digested chromatin by addition of 23.5µl of buffer C (200 mM HEPES-KOH pH 7.9, 15 mM $MgCl_2$, 2.5 M NaCl, 5 mM DTT, 2 mM EDTA, protease inhibitor cocktail mix 1x, 100 mM NaF) and incubating at 4°C on rotating wheel for 40 min. The mix

was centrifuged for 5 minutes at 12000 rpm at 4°C and the supernatant (i.e. the chromatin fraction) recovered.

**β-galactosidase and SAHF (senescence-associated heterochromatin foci) staining for senescence analyses**

38HK, HNC-136, and CAL-51 cells were transiently transfected with siRNAs or lentiviral shRNA plasmids and cultured for 48 hours. 38HK transduced with p19INK4d were instead cultured for 72 h. The NIH 3T3 feeder layer was removed with PBS/EDTA from 38HK cultures prior to senescence analyses.

Senescence was assessed using the Senescence β-Galactosidase Staining Kit at pH 6 following the manufacturer's instructions (Cell Signaling Technology). For SAHF staining, 38HK cells were layered on slides coated with polylysine and fixed in 4% paraformaldehyde in PBS (pH 7.4) for 15 min at room temperature, and permeabilized with 0.1% Triton X-100 in PBS for 15 min [71]. Cells were incubated with H3K9me3 antibody (abcam; ab1220) for 2 hours at room temperature, followed by incubation with Alexa Fluor 488-conjugated or Alexa Fluor 555-conjugated secondary antibody for 1 hour at room temperature and mounted using Vectashield Antifade Mounting Medium with DAPI. The slides were visualized using a Nikon Eclipse Ti wide-field inverted fluorescence video microscope. The images thus captured were analyzed by NIS-Element software from Nikon.

**Pulldown assays**

The minimal E2F4/DP1 heterodimer was produced by co-expression of 6xHis-E2F4 (84-413) and GST-DP1(199-350) in *E. coli* BL21 DE3 cells overnight at 15 °C. The bacterial pellet (500 ml expression) was resuspended in lysis buffer (20 mM Tris pH 8.0, 400 mM NaCl, 10% glycerol, 5mM DTT, lysozyme, 100 μg/ml DNAse, 100 μg/ml RNAse, cOmplete EDTA-free

(Roche)) and lysed by sonication. Cleared extracts were applied to $Ni^{2+}$-NTA resin previously equilibrated in buffer A (20 mM Tris pH 8.0, 400 mM NaCl, 10% glycerol, 2mM DTT). After extensive washing the E2F4/DP1 heterodimer was eluted by applying buffer A supplemented with 250 mM imidazole competitor. Subsequently, the sample was concentrated and then buffer exchanged using a Nap10 (GE healthcare) column equilibrated in buffer B (20 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 2mM DTT).

Over-expression of ΔNp73α and TAp73α proteins (full-length and truncated constructs) fused to MBP was carried out overnight in *E. coli* BL21 DE3 cells at 15 °C. Cell pellets (50 ml expressions) were resuspended in lysis buffer (20 mM Tris pH 8.0, 250 mM NaCl, 10% glycerol, 0.2% NP-40, 2mM DTT, lysozyme, 100 µg/ml DNAse, 100 µg/ml RNAse, cOmplete EDTA-free (Roche)), lysed by sonication and cleared by centrifugation. Supernatants were then incubated with 100 µl of pre-equilibrated amylose resin beads for 2 hours at 4°C. Subsequently, resin was extensively washed with PD buffer (20 mM Tris pH 8.0, 150 mM NaCl, 2mM DTT, cOmplete EDTA-freeFor the pulldown experiment, 10 µl of amylose resin coupled to MBP-p73/ΔNp73α proteins were incubated with clarified lysates of HEK293T expressing 3xFlag-E2F4 or recombinant purified E2F4/DP1 heterodimer for 2 hours at 4°C. After two quick washing steps with PD buffer, complexes were eluted by incubation with 20 µl of PD buffer supplemented with 20 mM maltose for 15 minutes at 4°C. PD reactions were migrated onto two separate 10% SDS-PAGE gel. One gel was used for Western blot to detect 3xFlag-E2F4, His-E2F4(84-413) or GST-DP1(199-350), the other for Coomassie staining to detect MBP-p73/ΔNp73α proteins.

**Immunoblotting**

Western blot detection of endogenous proteins was performed using the following antibodies: p73 (Abcam, ref: ab215038), E2F4 (Santa Cruz Biotechnology, ref. sc-398543X), E2F5

(Genetex, ref. GTX129491), DP1 (Abcam, ref. ab124678), p130 (Cell Signaling, ref. 13610S) β-actin (clone C4, MP Biomedicals), GAPDH (6C5, ref. sc-32233, Santa Cruz). HA-ΔNp73α was detected using the HA-peroxidase antibody (Roche, ref: 12013819001). ΔNp73α-TAP was detected using the anti-TAP antibody (Thermofisher Scientific, ref. CAB1001). 3xFlag-E2F4 was detected using the anti-Flag antibody (Sigma, ref. F3165) antibody. The protein expression levels of the Gluc1- and Gluc2-fusion proteins were evaluated using the anti-Gluc antibody (New England Biolabs, ref. E8023).

**mRNA-seq**

38HK cells transfected with scramble siRNA or E2F4-5 siRNAs were collected at 48 hours after transfection. Total RNA was extracted from 38HK (about $10^6$ cells per sample) using the RNeasy Mini kit from QIAGEN and quantified by Qubit.

A total of 6 samples (3 for scramble siRNA and 3 for E2F4-5 siRNA) were analyzed by the GenomEast platform of IGBMC (Illkirch, France). RNA-seq libraries were generated from 500 ng of total RNA using the TruSeq Stranded mRNA Library Prep Kit and TruSeq RNA Single Indexes kits A and B (Illumina, San Diego, CA), according to the manufacturer's instructions. The read length was 50 nt. The mean total reads per sample was 59,103,999.

Mapping of the reads was processed with STAR 2.7.3a on the primary assembly of the latest release of the human genome [72] (GRCh38.p13, release 33, PRI version: https://www.gencodegenes.org/human/ ) with corresponding comprehensive gene annotations. No soft clipping was accepted. Of the reads, 79% mapped once on the genome, 14% multiple times and 6.4% were below the minimum length threshold to map. Reads were counted using htseq-count version 0.11.2 [73] with reverse strand matching (option “stranded reverse”). Differential expression analysis was done with DESeq2 1.24.0 [74] with Benjamini-Hochberg correction for multiple tests on R 3.6.2.

**RT-qPCR**

Total RNA was extracted from cultured cells using the NucleoSpin RNA II Kit (Macherey-Nagel). The RNA obtained was reverse-transcribed to cDNA using the RevertAid H minus First Strand cDNA Kit (Life Technologies) according to the manufacturer's protocols. Real-time quantitative PCR (qPCR) was performed using the LightCycler 480 SYBR Green I Master (Roche) or the Mesa Green qPCR MasterMix Plus for SYBR Assay (Eurogentec) with the primers listed in S6 Table. Primers were selected on PrimerBank database [75]. Reactions were run in triplicate and expression was normalized to GAPDH. The expression analysis was performed using the MxPro QPCR software (Agilent).

**Chromatin immunoprecipitation (ChIP)**

ChIP was performed using the Shearing ChIP and OneDay ChIP kits (Diagenode) according to the manufacturer's instructions. Briefly, cells were sonicated to obtain DNA fragments of 200–500 bp. Sheared chromatin was immunoprecipitated with antibodies against the following proteins/tags: HA (Abcam, ref. ab9110), p73 (Abcam, ref. ab215038), E2F4 (Santa Cruz Biotechnology, ref. SC-398543X), p130 (Cell Signaling, ref. 13610S). 10% of the sheared chromatin was kept as the input for the ChIP.

Immunoprecipitated chromatin has been analysed by q-PCR using the LightCycler 480 SYBR Green I Master (Roche) on a LightCycler® 96 Instrument or the Mesa Green qPCR MasterMix Plus for SYBR Assay (Eurogentec) on a Stratagene Mx3005P Multiplex Quantitative Real Time PCR System. The sequences of primers used for qPCR are described in S6 Table. Primers surrounding the target region were checked for specificity using the NCBI Primer designing tool. ChIP qPCR results were analyzed by evaluating signal of enrichment over noise normalized to Input.

### Quantification and statistical analysis

Quantification of protein levels from western blot bands was done using the Evolution-Capt Edge software (Vilber) or ImageLab software (Biorad). The immunofluorescence images were processed using ImageJ software. The data presented are expressed as means ± SD from three independent experiments. *P*-values are calculated using unpaired Student's t-test.

## DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [76] partner repository with the dataset identifier PXD022947. The RNAseq analyses have been deposited to the GEO database [77] with the identifier GSE162816. Original data files for Western blot and senescence analyses are deposited on the public repository Mendeley Data (doi: 10.17632/cd5hsz8z8w.1).

## ACKNOWLEDGEMENTS

The authors would like to thank Bertrand Seraphin (IGBMC, Strasbourg) for the pCDNA-TAP vector, Yves Jacob (Institut Pasteur, France) for the pSPICA vectors and Christophe Romier (IGBMC, Strasbourg) for the pnEA and pmCS vectors. The authors are grateful to Gunter Stier (BZH, University of Heidelberg), Christian Gaiddon (INSERM U1113, Strasbourg), the engineers of the GenomEast platform of IGBMC (Christelle Thibault-Charpentier and Bernard Jost), and members of the IARC and BSC-UMR7242 teams for precious help and advice.

The authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the institutions with which they are affiliated.

## AUTHOR CONTRIBUTIONS

Conceptualization: Valerio Taverniti, Massimo Tommasino and Katia Zanier.

Investigation: Valerio Taverniti, Hanna Krynska, Marie-Laure Straub, Cécilia Sirand, Maria Carmen Romero-Medina, Stefano Moro, Luc Negroni, Denise Martinez-Zapien.

Methodology: Valerio Taverniti, Hanna Krynska, Marie-Laure Straub, Cécilia Sirand, Assunta Venuti, Murielle Masson.

Formal analysis: Valerio Taverniti, Marie-Laure Straub, Eugenie Lohmann, Alexis Robitaille, Luc Negroni.

Visualization: Valerio Taverniti, Katia Zanier.

Writing (original draft): Massimo Tommasino, Katia Zanier.

Project administration: Massimo Tommasino, Katia Zanier.

Funding acquisition: Massimo Tommasino, Katia Zanier.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

## FUNDING

This work received institutional support from IARC/WHO, CNRS, and Université de Strasbourg. The work was supported by grants from ‘Fondation ARC pour la recherche sur le cancer’ (ref. PJA 20151203192) and the ‘Institut National de la Santé et de la Recherche Médicale’ (ref. ENV201610) to MT, and Agence Nationale de la Recherche (ANR JCJC, ref. ANR-13-JSV8-0004-01), ‘Ligue contre le Cancer’ (Région Grand Est - CCIR), ‘Fondation pour La Recherche Medicale’ (Equipes FRM, ref. DEQ20180339231), and ‘Alsace contre le Cancer’ to KZ.

## FIGURE LEGENDS

**Figure 1.** Proteomics analyses in 38HK identify the E2F4/p130 complex as a partner of ΔNp73α. **(A)** Illustration of the proteomics approach used to identify nuclear binding partners of ΔNp73α. (**B**) Expression and purification of ΔNp73α-TAP complexes. (*Upper panel*) Distribution of ΔNp73α-TAP in cytoplasmic and nuclear fractions of 38HK extracts. The two fractions were centrifuged and analyzed by Western blot using an anti-TAP antibody. T: total extract; S: supernatant; P: pellet. ΔNp73α-TAP is present mainly in the nuclear fraction (Cyt(P) and Nuc(S), see also Materials and Methods section). (*Lower panel*) Silver-stained 10% SDS-PAGE analysis of elution fractions 1–5 from the second affinity purification step (calmodulin resin) for ΔNp73α-TAP and control (TAP) purifications. An additional elution with SDS was performed to recover all the remaining proteins. (**C**) Pathway analysis of nuclear ΔNp73α binding partners. The analysis was performed using the Reactome database [78]. Only significant pathways are shown (defined by a false discovery rate (FDR) value ≤ 0.02) and ranked based on the $-\log_{10}$ of the associated *P*-value (*: ≤ 0.05; **: ≤ 0.01; ***: ≤ 0.001; ****: ≤ 0.0001). Histogram bar size shows the number of input proteins involved in the corresponding pathway. See also S1 Table. (**D**) Sucrose gradient/co-IP experiments on endogenous 38HK proteins. (*Left panel*) Sucrose fractions of 38HK nuclear extracts stably expressing HA-ΔNp73α were migrated on a 10% SDS-PAGE gel and analyzed by Western blot using antibodies recognizing the HA tag, E2F4 and p130 proteins. (*Right panel*) Indicated fractions were immunoprecipitated using anti-HA antibody-coupled agarose beads. Unfractionated nuclear extract (NE) was loaded on the same gel as a reference for protein migration. Images for NE and sucrose fractions are derived from different exposures of the same membrane (see original Western blot image on Mendeley data).

**Figure 2.** ΔNp73α establishes a direct interaction with E2F4. (**A**) (*Left panel*) Representative dataset for GPCA analysis of binary interactions between ΔNp73/TAp63 proteins and components of the E2F4/p130 complex. Pairwise combinations of Gluc1-ΔNp73/TAp63 and Gluc2-E2F4/DP1/p130 fusion constructs were co-transfected in HEK293T cells. After 48 hours, the *Gaussia princeps* luciferase enzymatic activity was measured and expressed as normalized luminescence ratio (NLR) value [27]. The interactions of Gluc1-E2F4 with Gluc2-E2F4/DP1/p130 proteins are reported as internal controls. Error bars show standard deviations derived from triplicate measurements. (*Right panel*) Expression levels of Gluc1-fused ΔNp73/TAp63 and E2F4 proteins, and of Gluc2-fused E2F4, DP1 and p130 proteins in HEK293T cells. At 24h post-transfection cells were collected, lysed by incubation on ice and centrifuged. Aliquots of cleared supernatants were loaded onto an 8% SDS-PAGE gel. Proteins were visualized by Western blot using a polyclonal anti-Gluc antibody, which allows to detect both Gluc1 and Gluc2 fragments, albeit with lower efficiency for the Gluc2 fragment. Multiple bands are detected for E2F4, which probably correspond to different phosphorylation states of the protein [79]. (**B**) Expression of ΔNp73α, E2F4-5, DP1 and p130 proteins in 38HK cells transfected with scramble (SC) siRNAs or siRNAs against either E2F4 (siE2F4) or E2F5 (siE2F5) or E2F4 plus E2F5 (siE2F4-5). (*Left panel*) mRNA levels measured by RT-qPCR. (*Right panel*) Protein levels as determined by Western blot analysis. All samples were migrated on the same gel (see original Western blot image on Mendeley data). (**C**) Sucrose gradient/co-IP experiments using nuclear extracts from 38HK in the scramble siRNA condition (*upper panels*) and in the double E2F4-5 knockdown condition (*lower panels*). Indicated fractions were immunoprecipitated using anti-HA antibody-coupled agarose beads. HA-ΔNp73α, E2F4, E2F5, DP1, and p130 were detected by Western blot. IgG(H) chains migrate in close proximity with E2F5 and DP1 proteins. See also S3 Fig. for experiments under single E2F4 or E2F5

knockdown conditions. NE: unfractionated nuclear extract. See also legend of Fig. 1D. (**D**) Quantification of immunoprecipitated HA-ΔNp73α, DP1, and p130 proteins in the scramble and double E2F4-5 knockdown conditions of the co-IP experiments shown in panel C. For each protein, band intensities in the individual immunoprecipitated fractions are normalized by the band intensity in the NE control ($I_{fr}/I_{NE}$). Then, the $I_{fr}/I_{NE}$ values of fractions 12 to 16 are summed to get an estimate of the total protein levels. Since immunoprecipitation of HA-ΔNp73α seems to be more efficient for the E2F4-5 knockdown compared to the scramble condition (i.e. 1.2-fold), the total DP1 and p130 protein vales are further normalized for the HA-ΔNp73α levels in each condition.

**Figure 3.** Identification of domains involved in the interaction between ΔNp73α and E2F4. (**A**) (*Left panel*) Schematic representation of the Gluc1-fused TA/ΔNp73α and Gluc2-fused E2F4 constructs analyzed by GPCA. The ΔN(1–433) construct corresponds to the ΔNp73β isoform. Domains of TAp73α/ΔNp73α: transactivation domain (TAD, yellow); DNA-binding domain (DBD, blue); oligomerization domain (OD, red); sterile alpha motif domain (SAM, green); 13-amino-acid peptide specific for ΔNp73 isoforms (ΔN peptide, purple). Domains of E2F4: DNA-binding domain (DBD, blue); DP-binding and dimerization domain (DP-dim, red); transactivation domain (TAD, yellow); pocket protein binding domain (PPBD, orange). (*Right panel*) Expression levels of Gluc1- and Gluc2-fused TA/ΔNp73α and E2F4 constructs in HEK293T cells. Proteins were resolved on a 25%/10% gradient SDS-gel and visualized by Western blot using an anti-Gluc antibody. (**B**) (*Left panel*) GPCA analysis of the pairwise interactions of Gluc1-TAp73α and Gluc1-ΔNp73α FL with Gluc2-E2F4 FL. (*Right panel*) MBP-pulldown analysis of the interactions of TAp73α and ΔNp73α FL with E2F4 FL. Amylose resin coupled to MBP-TAp73α and MBP-ΔNp73α FL constructs was incubated with clarified extracts from HEK293T cells transiently expressing 3xFlag-E2F4 FL. PD reactions

were migrated on two separate 10% SDS-PAGE gels. One gel was used to detect E2F4 by Western blot using an anti-Flag antibody, and the second gel was used to detect MBP fusions by Coomassie staining. The SAM domain (residues 405–502 of ΔNp73α) was used as a negative control and was migrated on the same gel as the TAp73α and ΔNp73α constructs (see original images on Mendeley data). (**C**) GPCA analyses of pairwise ΔNp73α/E2F4 interactions. (*Left panel*) Gluc1-ΔNp73α deletion constructs versus Gluc2-E2F4 FL. (*Right panel*) Gluc1-ΔNp73α FL versus Gluc2-E2F4 deletion constructs. The NLR values were normalized to the ΔNp73α FL/E2F4 FL interaction. (**B**-**C**) The NLR values are averages from three independent experiments, with each experiment being performed in triplicate. *P*-values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (*: $P < 0.05$; **: $P < 0.01$). (**D**) Pulldown analysis of the interactions of MBP-ΔN(1-433) (i.e. -ΔNp73β) and MBP-ΔN(1-348) constructs with the minimal E2F4(84-413)/DP1(199-350) heterodimer. The heterodimer was produced in *E. coli* by co-expression of the E2F4 and DP1 constructs fused to His and GST tags, respectively. MBP is used as a negative control. PD reactions were migrated on two separate 10% SDS-PAGE gels. One gel was used to detect His-E2F4 and GST-DP1 by Western blot using an anti-His and anti GST antibodies, whereas the second gel was used to detect MBP-ΔNp73 fusions by Coomassie staining. (**E**) Proposed mechanism of ΔNp73α interaction with the E2F4/p130 complex. Red hatched box: E2F4 binding site.

**Figure 4.** Expression and regulation of ΔNp73α-E2F4/p130 complex components in primary HPK and transformed 38HK cells. (**A**) Immunoblotting of ΔNp73α and E2F4 and p130 proteins in human foreskin HPK and 38HK cell lines from 3 different donors (D1-D3). *: HPK(D2) and 38HK(D2) cell lines are from the same donor. 38HK(D2) is an early passage (i.e. passage 16) cell line, whereas the 38HK(D3) is a late passage cell line, which was used throughout this study. (**B**) (*Left panel*) GPCA analysis of the pairwise interactions between Gluc1-38.E7/E6 or

Gluc1-16.E7 proteins and Gluc2-fused RB proteins co-expressed in HEK293T cells. (*Middle and right panels*) Expression levels of Gluc1-38.E7/E6 and Gluc1-16.E7 proteins and Gluc2-RB proteins in HEK293T cells. Gluc1-38.E7/E6 and Gluc1-16.E7 were migrated on a 12% SDS-PAGE gel, whereas Gluc2-RB proteins on a 8% gel. (**B**) Comparison of ΔNp73α, E2F4 and p130 levels in 38(HK) in control *versus* MG132 treatment conditions. Cells were treated with 10 μM MG132 and incubated for 4 hours. (*Left panel*) Western blot analysis. Samples were loaded on a 10% SDS-PAGE gel. (Right panel) Quantification of proteins from the Western blot analysis. Protein levels were normalized to actin. (**D**) Distribution of ΔNp73α and E2F4 and p130 proteins in the nuclear soluble (Nuclear(S)), chromatin and cytoplasm fractions of 38HK. The nuclear soluble and chromatin fractions were obtained as described in the Materials and Methods section.

**Figure 5** Depletion of E2F4-5 in induces senescence in 38HK. (**A**) *In situ* senescence-associated β-galactosidase staining of 38HK. (*Left panel*) Representative photomicrographs of β-galactosidase staining at pH 6 on 38HK treated with scramble or E2F4-5 siRNA at 48 hours after transfection. (*Right panel*) Percentage of senescent 38HK under scramble and double E2F4-5 siRNA conditions. (**B**) Immunofluorescent staining for SAHF in 38HK. (*Left panel*) Representative images obtained using an epifluorescence microscope of 38HK cells transfected with scramble (SC) siRNAs or siRNAs against either E2F4 (siE2F4), E2F5 (siE2F5) or E2F4 and E2F5 (siE2F4-5). At 48 hours post-transfection, cells were layered on coverslips coated with polylysine and probed using an anti-H3K9me3 antibody followed by secondary Alexa Fluor 488-conjugated antibody. Nuclei were stained with DAPI (coloured blue). Images were merged using ImageJ software. Enlarged regions are indicated by white rectangles. (*Right panel*) Quantification of cells positive for $H3K9me^3$ marks in the four conditions was performed using ImageJ software; 150 cells were considered for each sample. (**A**-**B**) The data are averages

of three independent experiments. *P*-values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (**: $P < 0.01$).

**Figure 6.** Genome-wide analysis (mRNA-seq) of E2F4-5 double knockdown in 38HK and co-regulation by ΔNp73α. **(A)** Volcano plot representation of genes deregulated by transfection of siRNA against E2F4-5. *x*-axis: $\log_2$ transformation of the fold change; genes with a positive value are upregulated, whereas genes with a negative value are downregulated. *y*-axis: $-\log_{10}$ transformation of the adjusted *P*-value (*P*adj). Blue circles: significantly deregulated genes associated with *P*adj ≤ 0.05. Grey circles: non-significantly deregulated genes. Red circles: deregulated genes that have been tested in classical RT-qPCR experiments. A total of 30,274 genes were analyzed. **(B)** Biological processes enriched in upregulated (*left panel*) and downregulated (*right panel*) genes. Pathway analysis was performed using the KEGG-2019 database [80]. Only significant pathways are shown (defined by *P*adj ≤ 0.05) and ranked based on the associated *P*-value (*: ≤ 0.05; **: ≤ 0.01; ***: ≤ 0.001; ****: ≤ 0.0001). Histogram bar size shows the number of genes involved in the corresponding pathway. (**C**) RT-qPCR analysis of downregulated (*left panel*) or upregulated (*right panel*) genes from the RNA-seq experiment, under conditions of E2F4-5 knockdown (white histograms) or ΔNp73α knockdown (grey histograms). For all conditions, mRNA levels were first normalized to GAPDH expression. Then, mRNA levels in E2F4-5 or ΔNp73α knockdown 38HK were normalized to the mRNA levels observed under negative control conditions (scramble siRNA or sense ΔNp73α). The red line shows the expression level of each gene under scramble or sense control conditions (set to 1). "E2F4-5 siRNA/scramble" and "AS/S ΔNp73α" values > 1 and < 1 indicate upregulation and downregulation of gene expression, respectively. Underlined genes are significantly rescued upon either E2F4-5 or ΔNp73α knockdown compared with the control conditions. The values reported are averages of three independent experiments. *P*-

values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (*: $P < 0.05$; **: $P < 0.01$). (**D**) ΔNp73α and E2F4 protein levels in 38HK in conditions of scramble and ΔNp73α knockdown using sense (S) and antisense (AS) oligonucleotides (see also S7 Fig.).

**Figure 7.** E2F4-5 knockdown decreases ΔNp73α recruitment to the promoters of the STC1 and MAFB genes. (**A, B**, *upper panels*) Schematic illustration of the long promoter regions (+2500 nucleotides from TSS) of the STC1 and MAFB genes. Numbers are the positions of TP53/TP73 and E2F4 REs associated with high prediction scores (see S4 Table). (**A, B,** *lower panels* and **C**) Results from ChIP experiments. 38HK expressing HA-ΔNp73α were treated with scramble (grey histograms) or E2F4-5 (white histograms) siRNAs and harvested at 48 hours after transfection. Nuclear extracts were processed for ChIP using an anti-HA antibody recognizing HA-ΔNp73α (HA), an anti-p73 antibody recognizing both p73 and ΔNp73α proteins (p73) or anti-E2F4 (E2F4) antibodies. The eluted DNA was analyzed by qPCR with primers flanking regions 1, 2, and 3 of the STC1 promoter (**A**), regions 1 and 2 of the MAFB promoter (**B**), or the negative control region (S6 Table) (**C**). The amount of DNA bound by each protein is expressed as percentage of DNA in the input. (**A**-**C**)The data are averages of three independent experiments. *P*-values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (*: $P < 0.05$).

**Figure 8.** ΔNp73α cooperates with E2F4/p130 in HPV-negative cancer cells. (**A**) Sucrose gradient/co-IP experiments using nuclear extracts of HNC-136 cells. Sucrose fractions were immunoprecipitated using anti-p73 antibody-coupled beads and analyzed for endogenous ΔNp73α, E2F4, E2F5, and p130 proteins. Note that the p130 signal (highlighted by a red asterix) is partially masked by a closely migrating non-specific band. See also legend of Fig. 1D. (**B**) *In situ* senescence-associated β-galactosidase staining of HNC-136 and CAL-51 cells

at 48 hours after transfection. (*Left panel*) Representative photomicrographs of β-galactosidase staining at pH 6 in cells treated with scramble or E2F4-5 siRNA at 48 h after transfection. (*Right panel*) Percentage of senescent cells under scramble and E2F4-5 siRNA conditions. (**C**) Expression profiles of selected genes in CAL-51 cells. mRNA levels determined by RT-qPCR under conditions of E2F4-5 knockdown (white histograms) or ΔNp73α knockdown (grey histograms). The red line shows the expression level of each gene under scramble or sense control conditions (set to 1). "E2F4-5 siRNA/scramble" and "AS/S ΔNp73α" values > 1 and < 1 indicate upregulation and downregulation of gene expression, respectively. (**B**-**C**) The data are averages from three independent experiments. *P*-values are obtained from unpaired *t*-test, *n* = 3 biological triplicates (*: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$). See also Fig. S8 for E2F4-5 and ΔNp73α levels in the scramble and knockdown conditions.

**Figure 9**. Expression of $p19^{INK4d}$ induces senescence in 38HK. (**A**) $p19^{INK4d}$ mRNA levels determined by RT-qPCR in 38HK retro-transduced with pMSCV-$p19^{INK4d}$-IRES-GFP or with pLXSN-GFP (negative control). (**B**) *In situ* senescence-associated β-galactosidase staining. (*Left panel*) Representative photomicrographs of β-galactosidase staining at pH 6. (*Right panel*) Percentage of senescent 38HK in the negative control and $p19^{INK4d}$ expression conditions. The data are averages and standard deviations from three independent experiments. (**C**) Immunofluorescent staining for SAHF. (*Left panel*) Representative images obtained using an epifluorescence microscope. Cells were layered on coverslips coated with polylysine and probed using an anti-H3K9me3 antibody followed by secondary Alexa Fluor 555-conjugated antibody. Nuclei were stained with DAPI (coloured blue). (*Right panel*) Quantification of cells positive for $H3K9me^3$ marks in the in the negative control and $p19^{INK4d}$ expression conditions was performed using ImageJ software; 150 cells were considered for each sample. (**B**-**C**) The

data are averages from three independent experiments. *P*-values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (**: $P < 0.01$, ***: $P < 0.001$)

## LEGENDS FOR SUPPLEMENTARY FIGURES

**S1 Figure. Complex immunoprecipitation using an anti-E2F4 antibody.** (**A**) IP experiment on total 38HK nuclear extracts using either IgG or an anti-E2F4 antibody. Detection of endogenous ΔNp73α, E2F4 and p130 proteins was done by Western blotting using anti-p73, anti-E2F4 and anti-p130 antibodies, respectively. (**B**) Sucrose gradient/co-IP experiments performed on 38HK nuclear extracts stably expressing HA-ΔNp73α. Indicated fractions were immunoprecipitated using anti-E2F4 antibody and analyzed by Western blot using antibodies recognizing the HA tag, E2F4, and p130 proteins, respectively. Unfractionated nuclear extract (NE) was loaded on the same gel as a reference for protein migration. Images for NE and sucrose fractions are derived from different exposures of the same membrane (see original Western blot image on Mendeley data). Note that the p130 (highlighted by a red Asterix) is partially masked by a closely migrating non-specific band.

**S2 Figure. Size exclusion chromatography/co-IP experiments.** (*Top panel*) 38HK nuclear extracts stably expressing HA-ΔNp73α were applied to a Superose 6 10/300 increase column. (*Middle panel*) Fractions 20 to 30 were found to contain ΔNp73α, E2F4 and p130 proteins. Gel filtration fractions were migrated on a 10% SDS-PAGE gel and analyzed by Western blot using antibodies recognizing the HA tag, E2F4, and p130 proteins. (*Bottom panel*) Fractions 20 to 30 were immunoprecipitated using anti-HA antibody-coupled agarose beads and analyzed for HA-ΔNp73α, E2F4, and p130 proteins.

**S3 Figure. Sucrose gradient/co-IP experiments in conditions of single E2F4 or E2F5 knockdown.** 38HK were transfected with siRNAs against E2F4 (**A**) or against E2F5 (**B**). (*Left panels)* Sucrose fractions. (*Right panels*) Immunoprecipitations of sucrose fractions using an anti-HA antibody. See also legend of Fig. 1D of the main text.

**S4 Figure. Comparison of ΔNp73α interactions with E2F4 and E2F5.** (*Left panel*) GPCA analysis of Gluc1-ΔNp73α *versus* Gluc2-E2F4 and Gluc2-E2F5. (*Right panel)* Expression levels of Gluc1-ΔNp73α, Gluc2-E2F4 and Gluc2-E2F5 in HEK293T cells. Note that the differences in binding responses between E2F5 and E2F4 are related to the different expression levels of the two fusion proteins.

**S5 Figure. Expression level of E2F4, E2F5, TDP1 and p130 in normal and tumor tissues.** The expression matrix plot of E2F4, E2F5, TDP1 and p130 in tumors (T) or normal tissues (N) has been created with the GEPIA2 tool (http://gepia2.cancer-pku.cn)- Multiple Gene Analysis- Multiple Gene Comparison. Gene list: E2F4, E2F5, TDP1 and p130. Dataset: all cancer types. Matched Normal data: TCGA tumor + TCGA normal + GTEx normal. The density of color in each block represents the median expression value of the indicated gene in a given tissue, normalized by the maximum median expression value across all blocks. Different genes in same tumors or normal tissues are compared. GTE: Genotype-Tissue Expression.

**ACC**: Adrenocortical carcinoma, **BLCA**: Bladder Urothelial Carcinoma, **BRCA**: Breast invasive carcinoma, **CESC**: Cervical squamous cell carcinoma and endocervical adenocarcinoma, **CHOL**: Cholangio carcinoma, **COAD**: Colon adenocarcinoma, **DLBC**: Lymphoid Neoplasm Diffuse Large B-cell Lymphoma, **ESCA**: Esophageal carcinoma, **GBM**: Glioblastoma multiforme, **HNSC:** Head and Neck squamous cell carcinoma, **KICH**: Kidney Chromophobe, **KIRC**: Kidney renal clear cell carcinoma, **KIRP**: Kidney renal papillary cell

carcinoma, **LAML**: Acute Myeloid Leukemia, **LGG**: Brain Lower Grade Glioma, **LIHC**: Liver hepatocellular carcinoma, **LUAD**: Lung adenocarcinoma, **LUSC**: Lung squamous cell carcinoma, **OV**: Ovarian serous cystadenocarcinoma, **PAAD**: Pancreatic adenocarcinoma, **PCPG**: Pheochromocytoma and Paraganglioma, **PRAD**: Prostate adenocarcinoma, **READ**: Rectum adenocarcinoma, **SARC** : Sarcoma, **SKCM**: Skin Cutaneous Melanoma, **STAD**: Stomach adenocarcinoma, **TGCT**: Testicular Germ Cell Tumors, **THCA**: Thyroid carcinoma, **THYM**: Thymoma, **UCEC**: Uterine Corpus Endometrial Carcinoma, **UCS**: Uterine Carcinosarcoma.

**S6 Figure. Immunofluorescent staining for SAHF in 38HK transfected with plasmids encoding for shRNAs against E2F4 or E2F5**. (**A**) mRNA levels of E2F4 and E2F5 genes in conditions of knockdown using pLKO.1 shE2F4 or shE2F5 plasmids as measured by RT-qPCR. (**B**) (*Upper panel*) Representative images obtained using an epifluorescence microscope of 38HK cells transfected with pLKO.1 shE2F4 or shE2F5. At 48 hours post-transfection, cells were layered on coverslips coated with polylysine and probed using an anti-H3K9me3 antibody followed by secondary Alexa Fluor 488-conjugated antibody. Nuclei were stained with DAPI (coloured blue). Images were merged using ImageJ software. (*Lower panel*) Quantification of cells positive for $H3K9me^3$ marks in the four conditions was performed using ImageJ software; 150 cells were considered for each sample. (**A**-**B**) The data are averages of three independent experiments. *P*-values are obtained from unpaired *t*-test, $n = 3$ biological triplicates (**: $P < 0.01$; ***: $P < 0.001$).

**S7 Figure. Protein levels in the scramble and knockdown conditions in 38HK. (A)** (*Upper panel*) Western blot analysis of E2F4 and E2F5 protein levels in the 38HK cultures used for mRNA-seq analysis (three independent transfection experiments with scramble siRNA and

three independent transfections with E2F4-5 siRNAs). All samples were run on the same gel (see also Mendeley for original image). (*Lower panel*) Average levels of E2F4 and E2F5 proteins in the scramble (white histograms) and E2F4-5 (black histograms) siRNA conditions. (**B**) (*Upper panel*) Western blot analysis of ΔNp73α, E2F4 and p130 protein levels in 38HK treated with sense (S) or antisense (AS) oligonucleotides against ΔNp73α (three independent transfection experiments with S and three independent transfections with AS oligos). These 38HK cultures were used for the gene expression analysis shown in Fig. 6C of the main text. (*Lower panel*) Average levels of ΔNp73α and E2F4 in the S (white histograms) and AS (black histograms) conditions. (**A-B**) Protein levels in the knockdown conditions have been first normalized to GADPH and then to the respective control conditions. Total proteins used for the Western blot have been extracted using the Nucleospin RNA/Protein kit (Macherey-Nagel) following the manufacturer's protocol, which includes a protein precipitation step.

**S8 Figure. E2F4-5 and ΔNp73α levels in the scramble and knockdown conditions in cancer cells (A)** RT-qPCR analysis of E2F4 and E2F5 expression in HNC-136 and CAL-51 cell cultures treated with scramble (white histograms) or E2F4-5 siRNA (black histograms) and used for the gene expression analysis shown in Fig. 8 of the main text and in Fig. S9. mRNA levels for the E2F4-5 siRNA condition are expressed relative to the scramble (control condition, set to 1). Error bars report on the standard deviations for three independent transfection experiments. **(B)** (*Upper panel*) Western blot analysis of ΔNp73α protein levels in CAL-51 cultures treated with sense (S) or antisense (AS) oligonucleotides against ΔNp73α (three independent transfection experiments with S and three independent transfections with AS oligos). These CAL-51 cultures were used for the gene expression analysis shown in Fig. 8 of the main text. (*Lower panel*) Average ΔNp73α levels in the S and AS conditions. Protein levels

in the AS conditions were first normalized to actin and then to the respective control (S) conditions (control, set to 1).

**S9 Figure. Expression profiles of selected genes in HNC-136 cells.** mRNA levels determined by RT-qPCR in conditions of E2F4-5 knockdown. The red bar corresponds to the expression level of each gene in the scramble or sense control conditions (set to 1). "E2F4-5 siRNA/scramble" values > 1 and < 1 indicate upregulation and downregulation of gene expression, respectively. Error bars report on the standard deviations from three independent experiments. *: $P < 0.05$; **: $P < 0.01$. See also legend of Fig. 6 of the main text.

## LEGENDS FOR SUPPLEMENTARY TABLES

**S1 Table. Nuclear protein binding partners of ΔNp73α.** List of proteins displaying an enrichment of at least 10-fold in the ΔNp73α-TAP proteomics experiment compared with control (TAP-only) conditions. All proteins listed have been detected by more than two unique peptides, with the exception of DP1 that was detected by two unique peptides.

**S2 Table. mRNA-seq analysis: genes significantly upregulated upon E2F4-5 knockdown in 38HK.** Significantly upregulated genes are associated with adjusted *P*-values ≤ 0.05. Genes are ranked based on adjusted *P*-values. The presence or absence of E2F4 and TP53/TP73 RE within the long promoter region (+2500 nt from TSS) is reported for each gene.

**S3 Table. mRNA-seq analysis: genes significantly downregulated upon E2F4-5 knockdown in 38HK.** Significantly downregulated genes are associated with adjusted *P*-values ≤ 0.05. See also legend of S2 Table.

**S4 Table. Results of E2F4 and p53/p73 RE predictions for genes upregulated upon E2F4-5 knockdown in 38HK.** Sequences of upregulated genes were recovered from the Ensembl database version 86 dataset and biomaRt 2.42.0. The 2500 nt region upstream of TSS (defined as the long promoter region) was searched for E2F4 and TP53/TP73 REs using the JASPAR2020 [81] (https://github.com/da-bar/JASPAR2020) 0.99.8 and TFBSTools 1.25.2 packages. The score threshold was set to 0.85.

**S5 Table. Effects of E2F4-5 and ΔNp73α knockdown on the expression of selected genes in 38HK.** Genes which are significantly upregulated in the RNA-seq analysis and that have been tested in classical RT-qPCR experiments. +: rescue associated with significant *p*-value; -/+: rescue associated with higher average mRNA levels but not significant *p*-values; -: no rescue.

**S6 Table. Oligonucleotide sequences for gene knockdown, RT-qPCR, and ChIP experiments.**

## REFERENCES

1. Lee C-W, Thangue NBL. Promoter specificity and stability control of the p53-related protein p73. Oncogene. 1999;18: 4171–4181. doi:10.1038/sj.onc.1202793

2. Fontemaggi G, Kela I, Amariglio N, Rechavi G, Krishnamurthy J, Strano S, et al. Identification of Direct p73 Target Genes Combining DNA Microarray and Chromatin Immunoprecipitation Analyses. J Biol Chem. 2002;277: 43359–43368. doi:10.1074/jbc.m205573200

3. Murray-Zmijewski F, Lane DP, Bourdon J-C. p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. Cell Death Differ. 2006;13: 962–972. doi:10.1038/sj.cdd.4401914

4. Grob TJ, Novak U, Maisse C, Barcaroli D, Lüthi AU, Pirnia F, et al. Human ΔNp73 regulates a dominant negative feedback loop for TAp73 and p53. Cell Death Differ. 2001;8: 1213–1223. doi:10.1038/sj.cdd.4400962

5. Ueda Y, Hijikata M, Takagi S, Chiba T, Shimotohno K. New p73 variants with altered C-terminal structures have varied transcriptional activities. Oncogene. 1999;18: 4993–4998. doi:10.1038/sj.onc.1202817

6. Vikhreva P, Melino G, Amelio I. p73 Alternative Splicing: Exploring a Biological Role for the C-Terminal Isoforms. J Mol Biol. 2018;430: 1829–1838. doi:10.1016/j.jmb.2018.04.034

7. Melino G, Laurenzi VD, Vousden KH. p73: Friend or foe in tumorigenesis. Nat Rev Cancer. 2002;2: 605–615. doi:10.1038/nrc861

8. Gong J, Costanzo A, Yang H-Q, Melino G, Kaelin WG, Levrero M, et al. The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage. Nature. 1999;399: 806–809. doi:10.1038/21690

9. Agami R, Blandino G, Oren M, Shaul Y. Interaction of c-Abl and p73α and their collaboration to induce apoptosis. Nature. 1999;399: 809–813. doi:10.1038/21697

10. Mantovani F, Piazza S, Gostissa M, Strano S, Zacchi P, Mantovani R, et al. Pin1 Links the Activities of c-Abl and p300 in Regulating p73 Function. Mol Cell. 2004;14: 625–636. doi:10.1016/j.molcel.2004.05.007

11. Sayan BS, Yang AL, Conforti F, Tucci P, Piro MC, Browne GJ, et al. Differential control of TAp73 and ΔNp73 protein stability by the ring finger ubiquitin ligase PIR2. Proc National Acad Sci. 2010;107: 12877–12882. doi:10.1073/pnas.0911828107

12. Munarriz E, Bano D, Sayan AE, Rossi M, Melino G, Nicotera P. Calpain cleavage regulates the protein stability of p73. Biochem Bioph Res Co. 2005;333: 954–960. doi:10.1016/j.bbrc.2005.05.188

13. Asher G, Tsvetkov P, Kahana C, Shaul Y. A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. Gene Dev. 2005;19: 316–321. doi:10.1101/gad.319905

14. Candi E, Agostini M, Melino G, Bernassola F. How the TP53 Family Proteins TP63 and TP73 Contribute to Tumorigenesis: Regulators and Effectors. Hum Mutat. 2014;35: 702–714. doi:10.1002/humu.22523

15. Di C, Yang L, Zhang H, Ma X, Zhang X, Sun C, et al. Mechanisms, function and clinical applications of DNp73. Cell Cycle. 2013;12: 1861–1867. doi:10.4161/cc.24967

16. Accardi R, Fathallah I, Gruffat H, Mariggiò G, Calvez-Kelm FL, Voegele C, et al. Epstein - Barr Virus Transforming Protein LMP-1 Alters B Cells Gene Expression by Promoting Accumulation of the Oncoprotein ΔNp73α. Plos Pathog. 2013;9: e1003186. doi:10.1371/journal.ppat.1003186

17. Allart S, Martin H, Détraves C, Terrasson J, Caput D, Davrinche C. Human cytomegalovirus induces drug resistance and alteration of programmed cell death by accumulation of δN-p73α}. J Biol Chem. 2002;277: 29063–29068. doi:10.1074/jbc.m201974200

18. Accardi R, Dong W, Smet A, Cui R, Hautefeuille A, Gabet A, et al. Skin human papillomavirus type 38 alters p53 functions by accumulation of ΔNp73. Embo Rep. 2006;7: 334–340. doi:10.1038/sj.embor.7400615

19. Cornet I, Bouvard V, Campo MS, Thomas M, Banks L, Gissmann L, et al. Comparative Analysis of Transforming Properties of E6 and E7 from Different Beta Human Papillomavirus Types. J Virol. 2012;86: 2366–2370. doi:10.1128/jvi.06579-11

20. Accardi R, Scalise M, Gheit T, Hussain I, Yue J, Carreira C, et al. IκB Kinase β Promotes Cell Survival by Antagonizing p53 Functions through ΔNp73α Phosphorylation and Stabilization. Mol Cell Biol. 2011;31: 2210–2226. doi:10.1128/mcb.00964-10

21. Rollison DE, Viarisio D, Amorrortu RP, Gheit T, Tommasino M. An Emerging Issue in Oncogenic Virology: the Role of Beta Human Papillomavirus Types in the Development of Cutaneous Squamous Cell Carcinoma. J Virol. 2019;93. doi:10.1128/jvi.01003-18

22. Morgenstern JP, Land H. A series of mammalian expression vectors and characterisation of their expression of a reporter gene in stably and transiently transfected cells. Nucleic Acids Res. 1990;18: 1068–1068. doi:10.1093/nar/18.4.1068

23. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol. 1999;17: 1030–1032. doi:10.1038/13732

24. Liban TJ, Medina EM, Tripathi S, Sengupta S, Henry RW, Buchler NE, et al. Conservation and divergence of C-terminal domain structure in the retinoblastoma protein family. Proc National Acad Sci. 2017;114: 4942–4947. doi:10.1073/pnas.1619170114

25. Joerger AC, Rajagopalan S, Natan E, Veprintsev DB, Robinson CV, Fersht AR. Structural evolution of p53, p63, and p73: Implicationfor heterotetramer formation. Proc National Acad Sci. 2009;42: 17705–17710.

26. Gebel J, Luh LM, Coutandin D, Osterburg C, Löhr F, Schäfer B, et al. Mechanism of TAp73 inhibition by ΔNp63 and structural basis of p63/p73 hetero-tetramerization. Cell Death Differ. 2016;23: 1930–1940. doi:10.1038/cdd.2016.83

27. Cassonnet P, Rolloy C, Neveu G, Vidalain P-O, Chantier T, Pellet J, et al. Benchmarking a luciferase complementation assay for detecting protein complexes. Nat Methods. 2011;8: 990–992. doi:10.1038/nmeth.1773

28. Narita M, Nuñez S, Heard E, Narita M, Lin AW, Hearn SA, et al. Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence. Cell. 2003;113: 703–716. doi:10.1016/s0092-8674(03)00401-x

29. Lee B-K, Bhinge AA, Iyer VR. Wide-ranging functions of E2F4 in transcriptional activation and repression revealed by genome-wide analysis. Nucleic Acids Res. 2011;39: 3558–3573. doi:10.1093/nar/gkq1313

30. Sonzogni SV, Ogara MF, Belluscio LM, Castillo DS, Scassa ME, Cánepa ET. p19INK4d is involved in the cellular senescence mechanism contributing to heterochromatin formation. Biochimica Et Biophysica Acta Bba - Gen Subj. 2014;1840: 2171–2183. doi:10.1016/j.bbagen.2014.03.015

31. Litovchick L, Sadasivam S, Florens L, Zhu X, Swanson SK, Velmurugan S, et al. Evolutionarily Conserved Multisubunit RBL2/p130 and E2F4 Protein Complex Represses Human Cell Cycle-Dependent Genes in Quiescence. Mol Cell. 2007;26: 539–551. doi:10.1016/j.molcel.2007.04.015

32. Schmit F, Korenjak M, Mannefeld M, Schmitt K, Franke C, Eyss B von, et al. LINC, a Human Complex That is Related to pRB-Containing Complexes in Invertebrates Regulates the Expression of G 2 /M Genes. Cell Cycle. 2007;6: 1903–1913. doi:10.4161/cc.6.15.4512

33. Hsu J, Sage J. Novel functions for the transcription factor E2F4 in development and disease. Cell Cycle. 2016;15: 3183–3190. doi:10.1080/15384101.2016.1234551

34. Freed-Pastor WA, Prives C. Mutant p53: one name, many proteins. Gene Dev. 2012;26: 1268–1286. doi:10.1101/gad.190678.112

35. Kim MP, Lozano G. Mutant p53 partners in crime. Cell Death Differ. 2018;25: 161–168. doi:10.1038/cdd.2017.185

36. Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, et al. DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature. 2015;527: 384–388. doi:10.1038/nature15518

37. Morgunova E, Taipale J. Structural perspective of cooperative transcription factor binding. Curr Opin Struc Biol. 2017;47: 1–8. doi:10.1016/j.sbi.2017.03.006

38. Stampfel G, Kazmar T, Frank O, Wienerroither S, Reiter F, Stark A. Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature. 2015;528: 147–151. doi:10.1038/nature15545

39. Waghray A, Schober M, Feroze F, Yao F, Virgin J, Chen YQ. Identification of differentially expressed genes by serial analysis of gene expression in human prostate cancer. Cancer Res. 2001;61: 4283–6.

40. Rakha EA, Armour JAL, Pinder SE, Paish CE, Ellis IO. High-resolution analysis of 16q22.1 in breast carcinoma using DNA amplifiable probes (multiplex amplifiable probe hybridization technique) and immunohistochemistry. Int J Cancer. 2005;114: 720–729. doi:10.1002/ijc.20738

41. Razavipour SF, Harikumar KB, Slingerland JM. p27 as a Transcriptional Regulator: New Roles in Development and Cancer. Cancer Res. 2020. doi:10.1158/0008-5472.can-19-3663

42. Guan K-L, Jenkins CW, Li Y, O'Keefe CL, Noh S, Wu X, et al. Isolation and Characterization of p19INK4d, a p16- related Inhibitor Specific to CDK6 and CDK4. Molecular Biology of the Cell. 1996;7: 57–70. doi:10.1091/mbc.7.1.57

43. Chan FKAM, Zhang J, Cheng L, Shapiro DN, Winoto A. Identification of Human and Mouse p19, a Novel CDK4 and CDK6 Inhibitor with Homology to p16ink4. Molecular and Cellular Biology. 1995;15: 2682–2688. doi:0.1128/MCB.15.5.2682

44. Cánepa ET, Scassa ME, Ceruti JM, Marazita MC, Carcagno AL, Sirkin PF, et al. INK4 proteins, a family of mammalian CDK inhibitors with novel biological functions. Iubmb Life. 2007;59: 419–426. doi:10.1080/15216540701488358

45. Franklin DS, Godfrey VL, O'Brien DA, Deng C, Xiong Y. Functional Collaboration between Different Cyclin-Dependent Kinase Inhibitors Suppresses Tumor Growth with Distinct Tissue Specificity. Mol Cell Biol. 2000;20: 6147–6158. doi:10.1128/mcb.20.16.6147-6158.2000

46. Solomon DA, Kim J-S, Jean W, Waldman T. Conspirators in a Capital Crime: Co-deletion of p18INK4c and p16INK4a/p14ARF/p15INK4b in Glioblastoma Multiforme. Cancer Res. 2008;68: 8657–8660. doi:10.1158/0008-5472.can-08-2084

47. García-Fernández RA, García-Palencia P, Sánchez MÁ, Gil-Gómez G, Sánchez B, Rollán E, et al. Combined loss of p21waf1/cip1 and p27kip1 enhances tumorigenesis in mice. Lab Invest. 2011;91: 1634–1642. doi:10.1038/labinvest.2011.133

48. Martín-Caballero J, Flores JM, García-Palencia P, Collado M, Serrano M. Different cooperating effect of p21 or p27 deficiency in combination with INK4a/ARF deletion in mice. Oncogene. 2004;23: 8231–8237. doi:10.1038/sj.onc.1207863

49. Majumder PK, Grisanzio C, O'Connell F, Barry M, Brito JM, Xu Q, et al. A Prostatic Intraepithelial Neoplasia-Dependent p27Kip1 Checkpoint Induces Senescence and Inhibits Cell Proliferation and Cancer Progression. Cancer Cell. 2008;14: 146–155. doi:10.1016/j.ccr.2008.06.002

50. Cristofano AD, Acetis MD, Koff A, Cordon-Cardo C, Pandolfi PP. Pten and p27KIP1 cooperate in prostate cancer tumor suppression in the mouse. Nat Genet. 2001;27: 222–224. doi:10.1038/84879

51. Kim WY, Sharpless NE. The Regulation of INK4/ARF in Cancer and Aging. Cell. 2006;127: 265–275. doi:10.1016/j.cell.2006.10.003

52. Krimpenfort P, IJpenberg A, Song J-Y, Valk M van der, Nawijn M, Zevenhoven J, et al. p15Ink4b is a critical tumour suppressor in the absence of p16Ink4a. Nature. 2007;448: 943–946. doi:10.1038/nature06084

53. Fuxe J, Akusjärvi G, Goike HM, Roos G, Collins VP, Pettersson RF. Adenovirus-mediated overexpression of p15INK4B inhibits human glioma cell growth, induces replicative senescence, and inhibits telomerase activity similarly to p16INK4A. Cell Growth Differ Mol Biology J Am Assoc Cancer Res. 2000;11: 373–84.

54. Erickson S, Sangfelt O, Heyman M, Castro J, Einhorn S, Grandér D. Involvement of the Ink4 proteins p16 and p15 in T-lymphocyte senescence. Oncogene. 1998;17: 595–602. doi:10.1038/sj.onc.1201965

55. Kuilman T, Michaloglou C, Mooi WJ, Peeper DS. The essence of senescence. Gene Dev. 2010;24: 2463–2479. doi:10.1101/gad.1971610

56. Bai F, Chan HL, Smith MD, Kiyokawa H, Pei X-H. p19Ink4d Is a Tumor Suppressor and Controls Pituitary Anterior Lobe Cell Proliferation. Mol Cell Biol. 2014;34: 2121–2134. doi:10.1128/mcb.01363-13

57. Wang Y, Jin W, Jia X, Luo R, Tan Y, Zhu X, et al. Transcriptional repression of CDKN2D by PML/RARα contributes to the altered proliferation and differentiation block of acute promyelocytic leukemia cells. Cell Death Dis. 2014;5: e1431. doi:10.1038/cddis.2014.388

58. Dreidax D, Bannert S, Henrich K-O, Schröder C, Bender S, Oakes CC, et al. p19-INK4d inhibits neuroblastoma cell growth, induces differentiation and is hypermethylated and downregulated in MYCN-amplified neuroblastomas. Hum Mol Genet. 2014;23: 6826–6837. doi:10.1093/hmg/ddu406

59. Narita M, Narita M, Krizhanovsky V, Nuñez S, Chicas A, Hearn SA, et al. A Novel Role for High-Mobility Group A Proteins in Cellular Senescence and Heterochromatin Formation. Cell. 2006;126: 503–514. doi:10.1016/j.cell.2006.05.052

60. Basisty N, Kale A, Jeon OH, Kuehnemann C, Payne T, Rao C, et al. A proteomic atlas of senescence-associated secretomes for aging biomarker development. Plos Biol. 2020;18: e3000599. doi:10.1371/journal.pbio.3000599

61. Eychène A, Rocques N, Pouponnot C. A new MAFia in cancer. Nat Rev Cancer. 2008;8: 683–693. doi:10.1038/nrc2460

62. Pouponnot C, Sii-Felice K, Hmitou I, Rocques N, Lecoin L, Druillennec S, et al. Cell context reveals a dual role for Maf in oncogenesis. Oncogene. 2006;25: 1299–1310. doi:10.1038/sj.onc.1209171

63. DeYoung MP, Ellisen LW. p63 and p73 in human cancer: defining the network. Oncogene. 2007;26: 5169–5183. doi:10.1038/sj.onc.1210337

64. Krois AS, Dyson HJ, Wright PE. Long-range regulation of p53 DNA binding by its intrinsically disordered N-terminal transactivation domain. Proc National Acad Sci. 2018;115: 201814051. doi:10.1073/pnas.1814051115

65. He F, Borcherds W, Song T, Wei X, Das M, Chen L, et al. Interaction between p53 N terminus and core domain regulates specific and nonspecific DNA binding. Proc National Acad Sci. 2019;116: 201903077. doi:10.1073/pnas.1903077116

66. Forget A, Ayrault O, Besten W den, Kuo M-L, Sherr CJ, Roussel MF. Differential post-transcriptional regulation of two Ink4 proteins, p18Ink4c and p19Ink4d. Cell Cycle. 2008;7: 3737–3746. doi:10.4161/cc.7.23.7187

67. Diebold M-L, Fribourg S, Koch M, Metzger T, Romier C. Deciphering correct strategies for multiprotein complex assembly by co-expression: Application to complexes as large as the histone octamer. J Struct Biol. 2011;175: 178–188. doi:10.1016/j.jsb.2011.02.001

68. Gabet A-S, Accardi R, Bellopede A, Popp S, Boukamp P, Sylla BS, et al. Impairment of the telomere/telomerase system and genomic instability are associated with keratinocyte immortalization induced by the skin human papillomavirus type 38. Faseb J. 2008;22: 622–632. doi:10.1096/fj.07-8389com

69. Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. Proc National Acad Sci. 1993;90: 8392–8396. doi:10.1073/pnas.90.18.8392

70. Panda A, Martindale J, Gorospe M. Polysome Fractionation to Analyze mRNA Distribution Profiles. Bio-protocol. 2017;7. doi:10.21769/bioprotoc.2126

71. Venuti A, Pastori C, Pennisi R, Riva A, Sciortino MT, Lopalco L. Class B β-arrestin2-dependent CCR5 signalosome retention with natural antibodies to CCR5. Sci Rep-uk. 2016;6: 39382. doi:10.1038/srep39382

72. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21. doi:10.1093/bioinformatics/bts635

73. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31: 166–169. doi:10.1093/bioinformatics/btu638

74. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15: 550. doi:10.1186/s13059-014-0550-8

75. Wang X, Spandidos A, Wang H, Seed B. PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. Nucleic Acids Res. 2012;40: D1144–D1149. doi:10.1093/nar/gkr1013

76. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2018;47: gky1106-. doi:10.1093/nar/gky1106

77. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 2013;41: D991–D995. doi:10.1093/nar/gks1193

78. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2017;46: gkx1132-. doi:10.1093/nar/gkx1132

79. Paquin M-C, Cagnol S, Carrier JC, Leblanc C, Rivard N. ERK-associated changes in E2F4 phosphorylation, localization and transcriptional activity during mitogenic stimulation in human intestinal epithelial crypt cells. Bmc Cell Biol. 2013;14: 33. doi:10.1186/1471-2121-14-33

80. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 2018;47: gky962-. doi:10.1093/nar/gky962

81. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2019;48: D87–D92. doi:10.1093/nar/gkz1001

Figure 1

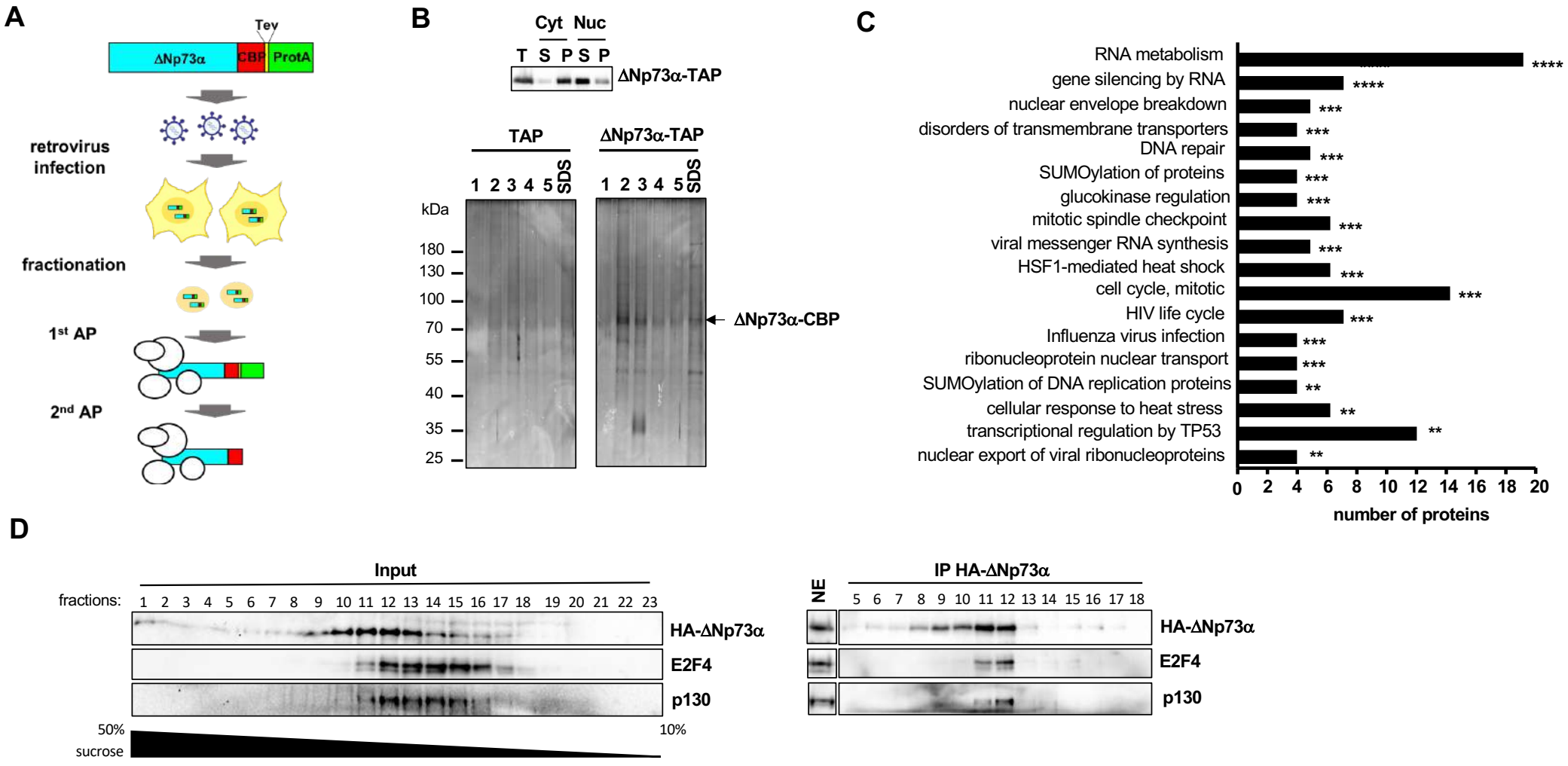

A
ΔNp73α
CBP
ProtA
Tev
retrovirus infection
fractionation
1st AP
2nd AP
B
Cyt Nuc
T S P S P
ΔNp73α-TAP
TAP
ΔNp73α-TAP
1 2 3 4 5 SDS
kDa
180
130
100
70
55
40
35
25
ΔNp73α-CBP
C
RNA metabolism
gene silencing by RNA
nuclear envelope breakdown
disorders of transmembrane transporters
DNA repair
SUMOylation of proteins
glucokinase regulation
mitotic spindle checkpoint
viral messenger RNA synthesis
HSF1-mediated heat shock
cell cycle, mitotic
HIV life cycle
Influenza virus infection
ribonucleoprotein nuclear transport
SUMOylation of DNA replication proteins
cellular response to heat stress
transcriptional regulation by TP53
nuclear export of viral ribonucleoproteins
0 2 4 6 8 10 12 14 16 18 20
number of proteins
D
Input
fractions: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
HA-ΔNp73α
E2F4
p130
50%
sucrose
10%
IP HA-ΔNp73α
NE
5 6 7 8 9 10 11 12 13 14 15 16 17 18
HA-ΔNp73α
E2F4
p130

Figure 2

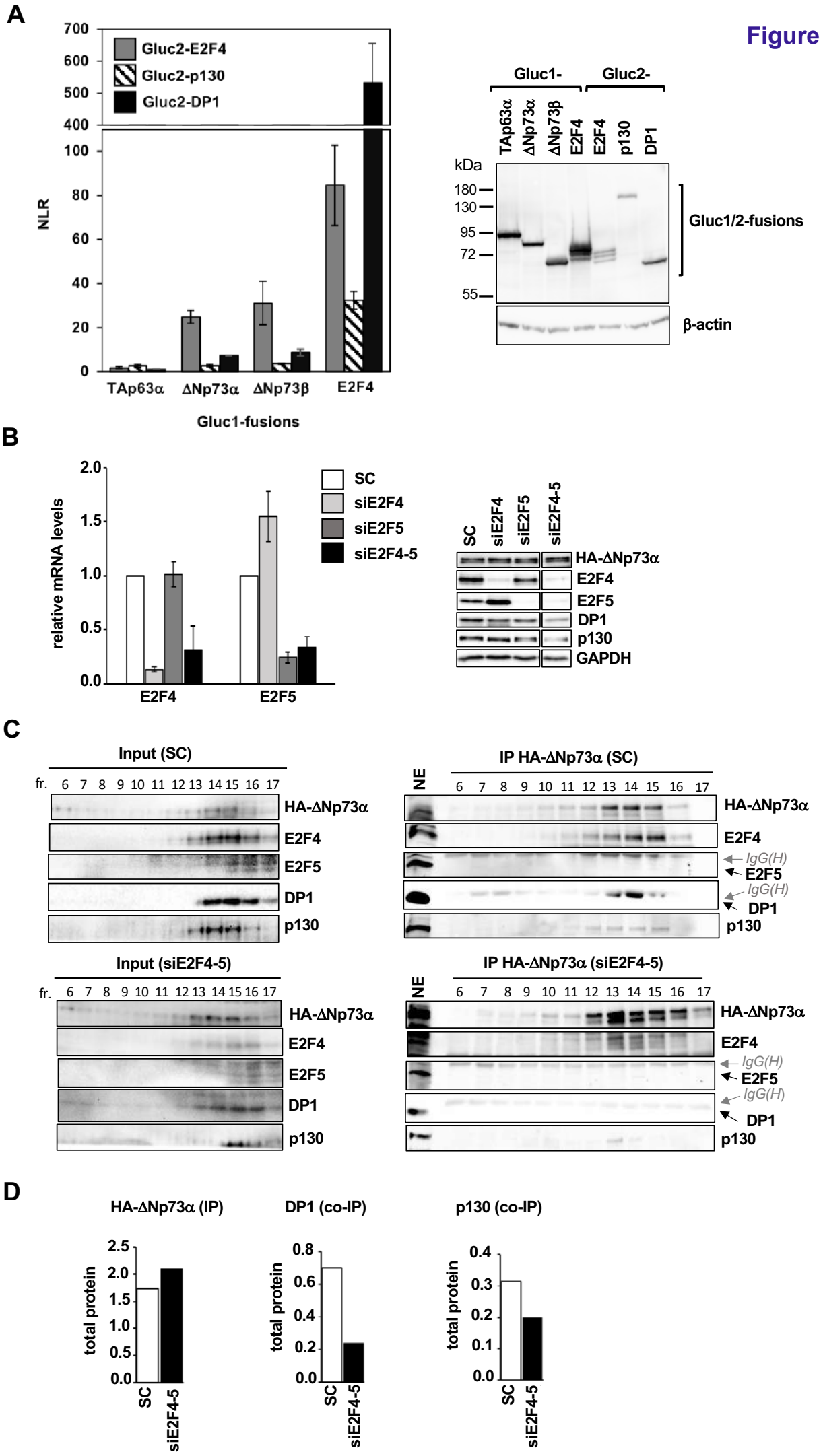
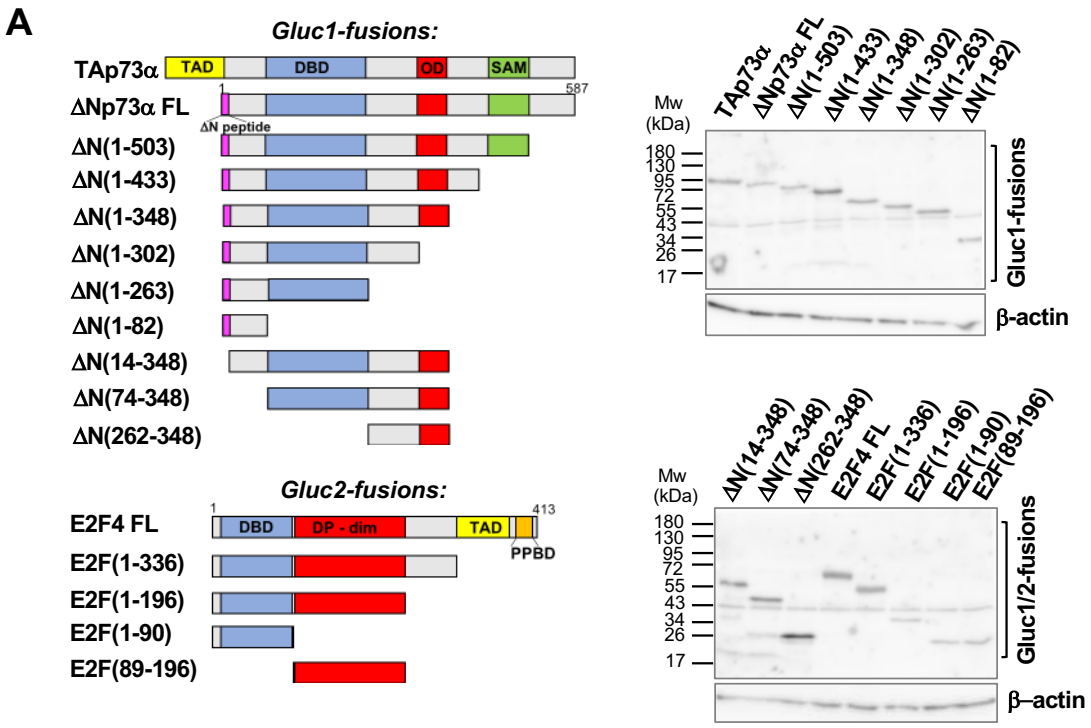
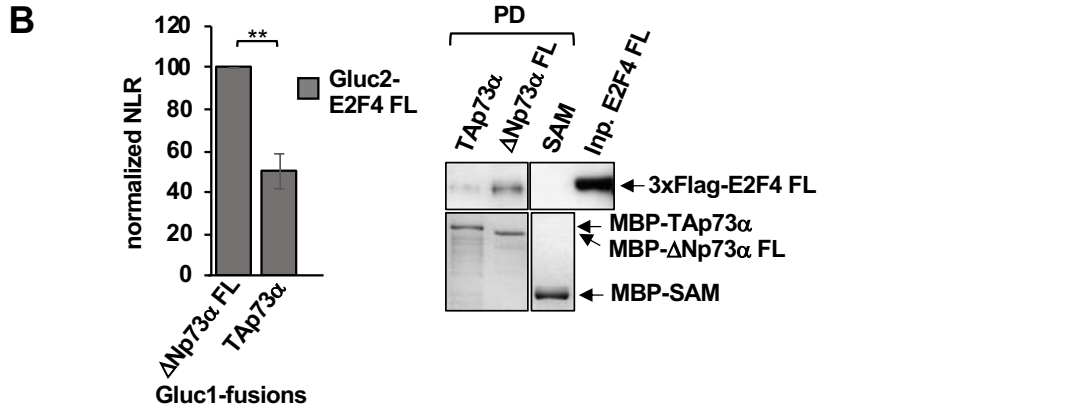
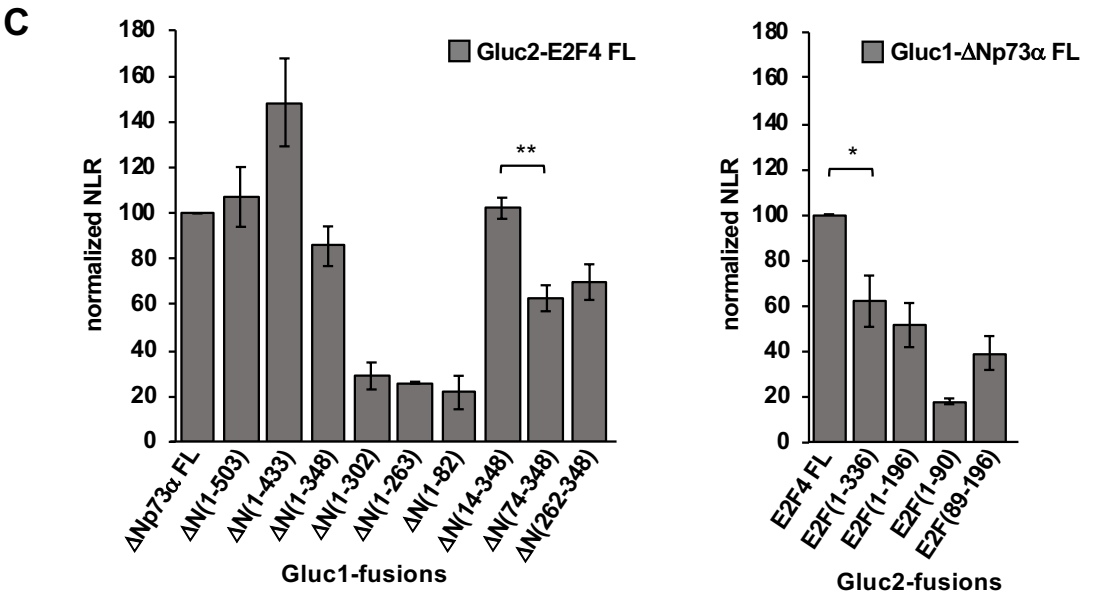
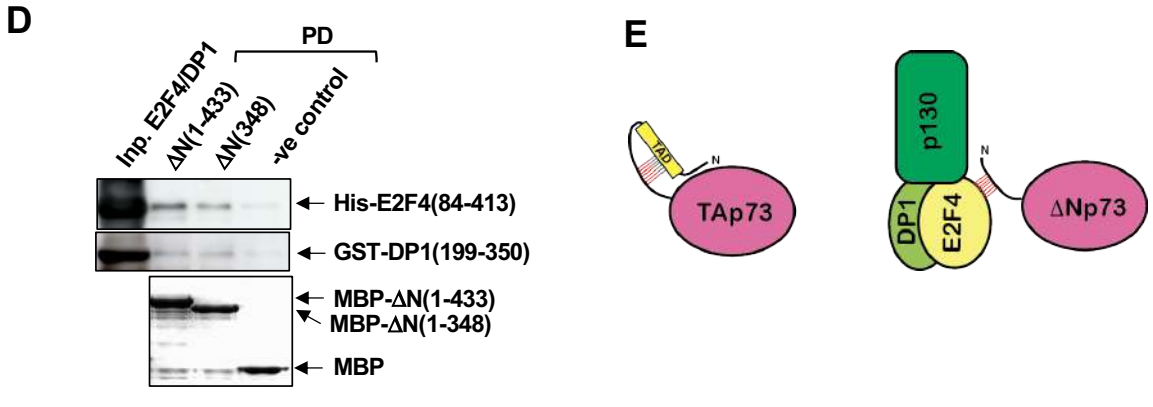

A
Gluc2-E2F4
Gluc2-p130
Gluc2-DP1
NLR
TAp63α
ΔNp73α
ΔNp73β
E2F4
Gluc1-fusions
Gluc1-
Gluc2-
kDa
180
130
95
72
55
Gluc1/2-fusions
β-actin
B
SC
siE2F4
siE2F5
siE2F4-5
relative mRNA levels
HA-ΔNp73α
E2F4
E2F5
DP1
p130
GAPDH
C
Input (SC)
IP HA-ΔNp73α (SC)
NE
fr.
IgG(H)
Input (siE2F4-5)
IP HA-ΔNp73α (siE2F4-5)
D
HA-ΔNp73α (IP)
DP1 (co-IP)
p130 (co-IP)
total protein

Figure 3


A
Gluc1-fusions:
TAp73α
ΔNp73α FL
ΔN(1-503)
ΔN(1-433)
ΔN(1-348)
ΔN(1-302)
ΔN(1-263)
ΔN(1-82)
ΔN(14-348)
ΔN(74-348)
ΔN(262-348)
TAD
DBD
OD
SAM
ΔN peptide
Gluc2-fusions:
E2F4 FL
E2F(1-336)
E2F(1-196)
E2F(1-90)
E2F(89-196)
DP - dim
PPBD
Mw (kDa)
Gluc1-fusions
β-actin
Gluc1/2-fusions


B
normalized NLR
Gluc2-E2F4 FL
ΔNp73α FL
TAp73α
Gluc1-fusions
PD
SAM
Inp. E2F4 FL
3xFlag-E2F4 FL
MBP-TAp73α
MBP-ΔNp73α FL
MBP-SAM


C
Gluc2-E2F4 FL
Gluc1-ΔNp73α FL
normalized NLR
Gluc1-fusions
Gluc2-fusions


D
PD
Inp. E2F4/DP1
ΔN(1-433)
ΔN(348)
-ve control
His-E2F4(84-413)
GST-DP1(199-350)
MBP-ΔN(1-433)
MBP-ΔN(1-348)
MBP
E
TAD
TAp73
p130
DP1
E2F4
ΔNp73

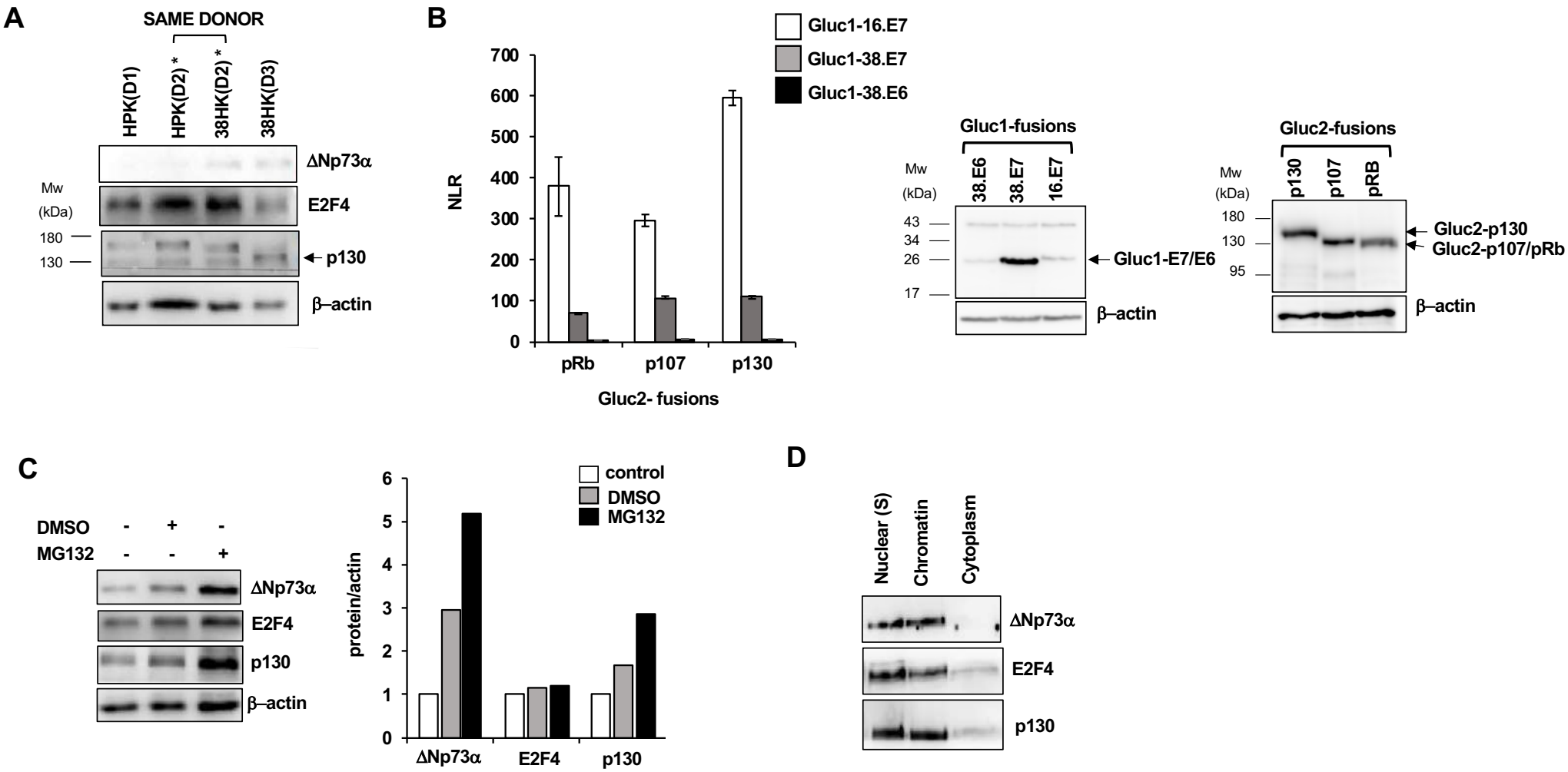

A
SAME DONOR
HPK(D1)
HPK(D2) *
38HK(D2) *
38HK(D3)
Mw
(kDa)
180
130
ΔNp73α
E2F4
p130
β–actin
B
Gluc1-16.E7
Gluc1-38.E7
Gluc1-38.E6
NLR
700
600
500
400
300
200
100
0
pRb
p107
p130
Gluc2- fusions
Gluc1-fusions
Mw
(kDa)
38.E6
38.E7
16.E7
43
34
26
17
Gluc1-E7/E6
β–actin
Gluc2-fusions
p130
p107
pRB
180
130
95
Gluc2-p130
Gluc2-p107/pRb
β–actin
C
DMSO - + -
MG132 - - +
ΔNp73α
E2F4
p130
β–actin
control
DMSO
MG132
protein/actin
6
5
4
3
2
1
0
ΔNp73α
E2F4
p130
D
Nuclear (S)
Chromatin
Cytoplasm
ΔNp73α
E2F4
p130

Figure 5

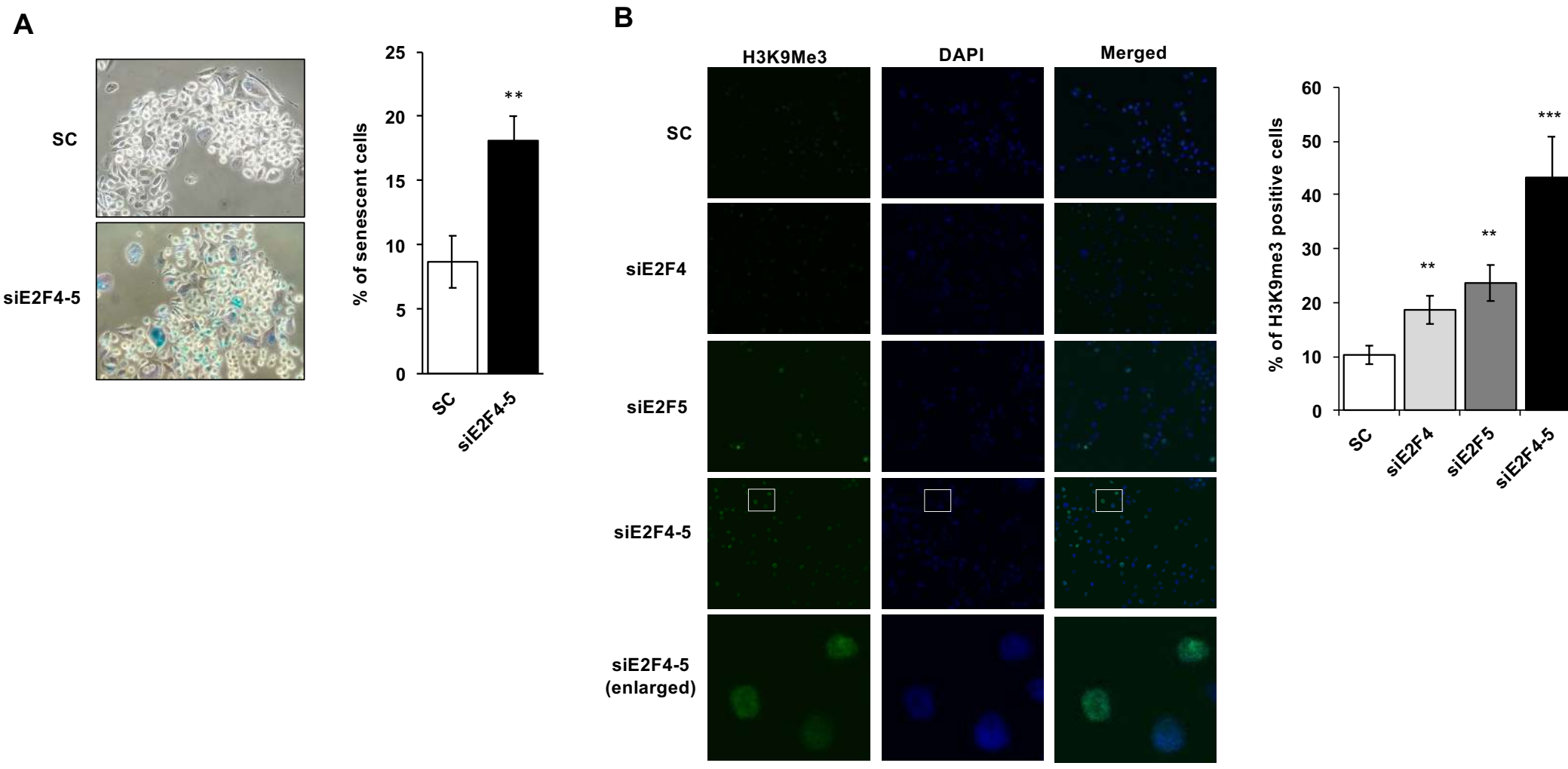
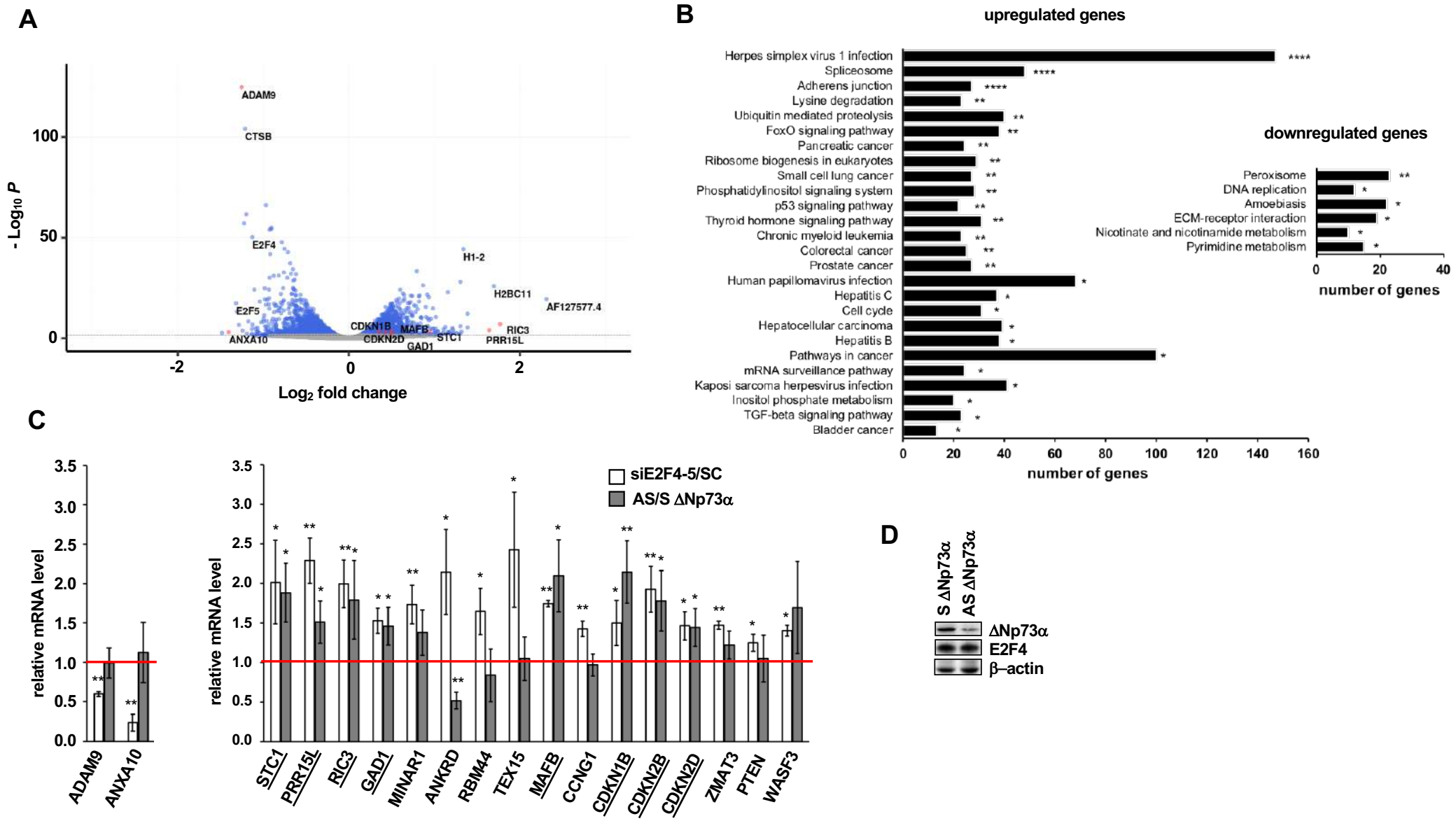
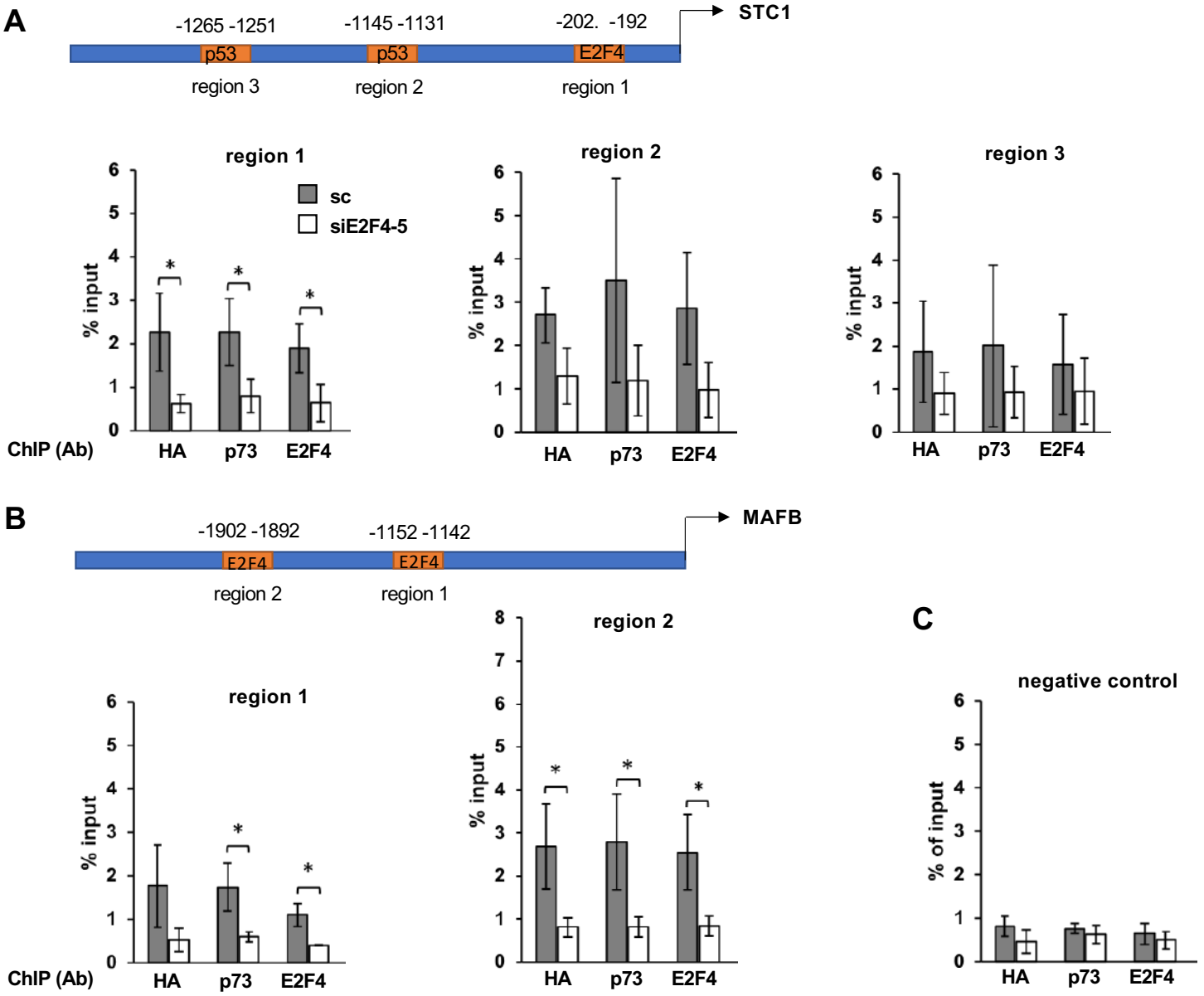
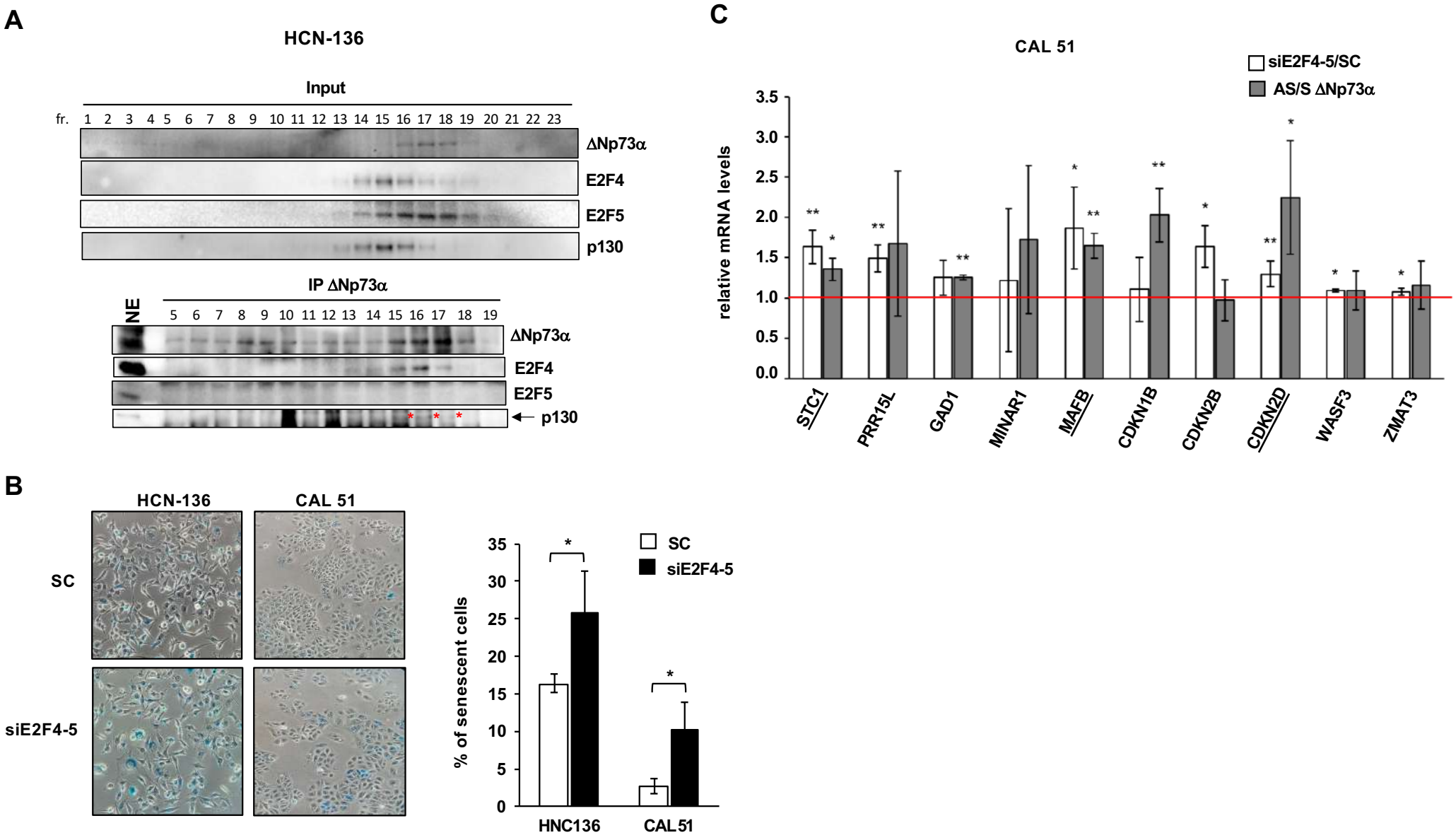
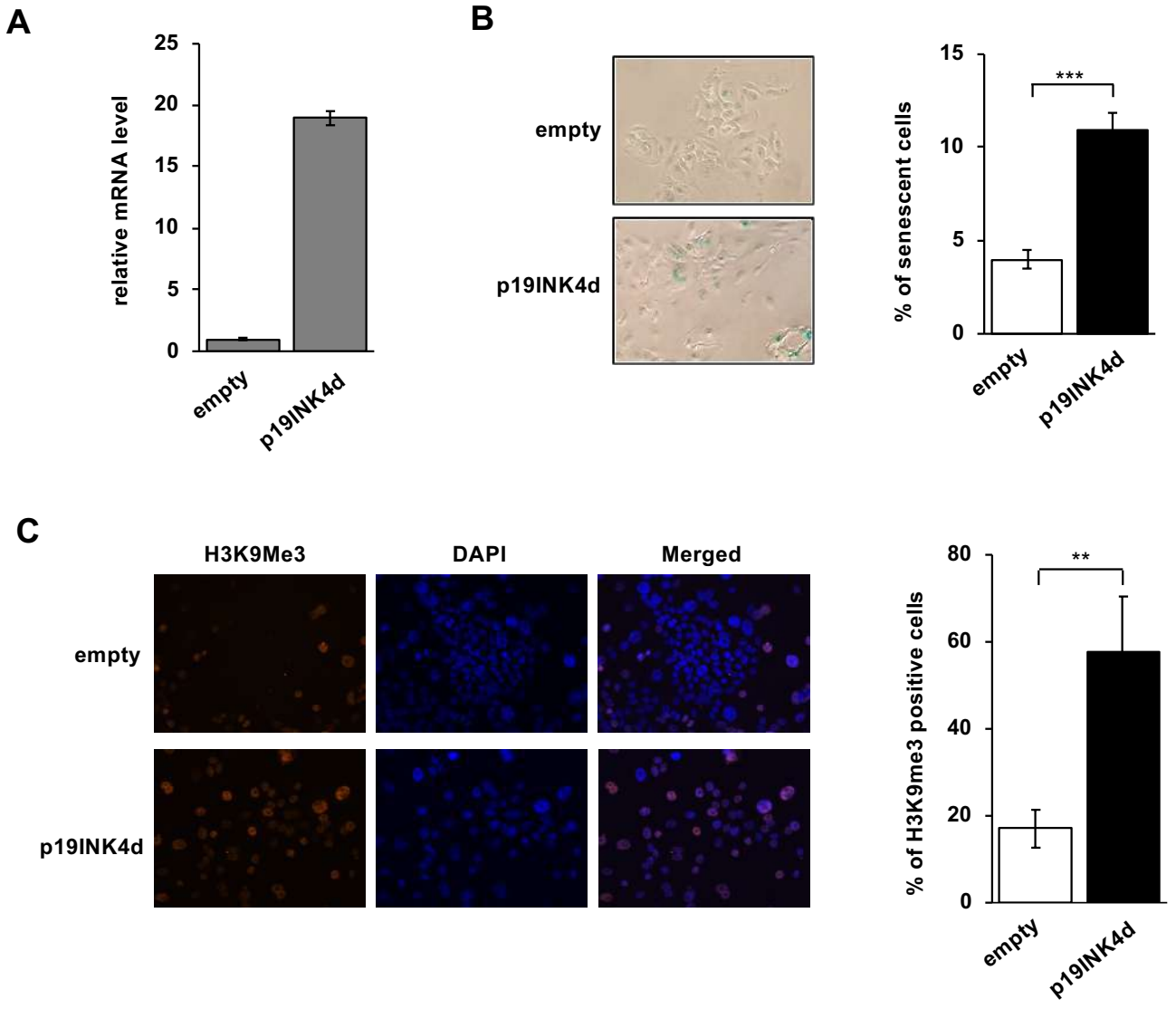

A
SC
siE2F4-5
% of senescent cells
25
20
15
10
5
0
**
SC
siE2F4-5
B
H3K9Me3
DAPI
Merged
SC
siE2F4
siE2F5
siE2F4-5
siE2F4-5 (enlarged)
% of H3K9me3 positive cells
60
50
40
30
20
10
0
**
**
***
SC
siE2F4
siE2F5
siE2F4-5

Figure 6


A
ADAM9
CTSB
E2F4
H1-2
H2BC11
AF127577.4
E2F5
CDKN1B
MAFB
RIC3
ANXA10
CDKN2D
STC1
GAD1
PRR15L
- Log10 P
Log2 fold change
B
upregulated genes
Herpes simplex virus 1 infection
Spliceosome
Adherens junction
Lysine degradation
Ubiquitin mediated proteolysis
FoxO signaling pathway
Pancreatic cancer
Ribosome biogenesis in eukaryotes
Small cell lung cancer
Phosphatidylinositol signaling system
p53 signaling pathway
Thyroid hormone signaling pathway
Chronic myeloid leukemia
Colorectal cancer
Prostate cancer
Human papillomavirus infection
Hepatitis C
Cell cycle
Hepatocellular carcinoma
Hepatitis B
Pathways in cancer
mRNA surveillance pathway
Kaposi sarcoma herpesvirus infection
Inositol phosphate metabolism
TGF-beta signaling pathway
Bladder cancer
number of genes
downregulated genes
Peroxisome
DNA replication
Amoebiasis
ECM-receptor interaction
Nicotinate and nicotinamide metabolism
Pyrimidine metabolism
number of genes
C
relative mRNA level
siE2F4-5/SC
AS/S ΔNp73α
ADAM9
ANXA10
STC1
PRR15L
RIC3
GAD1
MINAR1
ANKRD
RBM44
TEX15
MAFB
CCNG1
CDKN1B
CDKN2B
CDKN2D
ZMAT3
PTEN
WASF3
D
S ΔNp73α
AS ΔNp73α
ΔNp73α
E2F4
β–actin


A
-1265 -1251
-1145 -1131
-202. -192
STC1
p53
p53
E2F4
region 3
region 2
region 1
region 1
region 2
region 3
sc
siE2F4-5
% input
ChIP (Ab)
HA
p73
E2F4
B
-1902 -1892
-1152 -1142
MAFB
E2F4
E2F4
region 2
region 1
C
negative control
% of input


A
HCN-136
Input
fr. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
ΔNp73α
E2F4
E2F5
p130
IP ΔNp73α
NE
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
ΔNp73α
E2F4
E2F5
p130
B
HCN-136
CAL 51
SC
siE2F4-5
% of senescent cells
SC
siE2F4-5
HNC136
CAL51
C
CAL 51
siE2F4-5/SC
AS/S ΔNp73α
relative mRNA levels
STC1
PRR15L
GAD1
MINAR1
MAFB
CDKN1B
CDKN2B
CDKN2D
WASF3
ZMAT3


A
relative mRNA level
25
20
15
10
5
0
empty
p19INK4d
B
empty
p19INK4d
% of senescent cells
15
10
5
0
***
empty
p19INK4d
C
H3K9Me3
DAPI
Merged
empty
p19INK4d
% of H3K9me3 positive cells
80
60
40
20
0
**
empty
p19INK4d

# Résumé en français suivi des mots-clés en français

L'inhibiteur de la kinase κB (IKK) est le régulateur principal de la signalisation du NF-kB, impliqué dans l'activation et la régulation de l'inflammatoires et de la réponse immunitaire innée. Lors de la stimulation des récepteurs, IKK induit la translocation des facteurs NF-κB dans le noyau où ils régulent l'expression des gènes. Les espèces du complexe IKK régulant NF-κB sont le canonique (IKK1/IKK2/NEMO) et le complexe alternatif ($IKK1_2$). Les informations sur la structure soient disponibles pour les sous-unités IKK mais les détails sur les mécanismes d'interactions IKK-substrat fait défaut.

Nous avons identifié un nouveau motif d'amarrage, qui est basé sur le consensus YDDΦXΦ et qui est présent dans l'extrémité C-terminale désordonnée de IκBα. Ce motif est nécessaire et suffisant pour les interactions de IκBα avec les sous-unités IKK1 et 2. Une corrélation directe entre l'affinité de l'interaction du motif et l'activité kinase d'IKK2 est observée. Nous fournissons également un modèle structural 3D de cette interaction basé sur la diffraction aux rayons X et des données de spectrométrie de masse.

Le motif YDDΦXΦ est également présent dans d'autres substrats d'IKK (IκBβ, p100 et IRF7). De plus, tests préliminaires montrent que ce motif est également reconnu par la sous-unité régulatrice NEMO. Enfin, les peptides synthétiques contenant le motif YDDΦXΦ sont capables d'inhiber la phosphorylation du substrat par IKK2, ouvrant la voie au développement futur d'une nouvelle classe d'inhibiteurs sélectifs d'IKK.


**Mots clés : interaction protéique, IKK complexe, NF-κB, IκBα, inflammation, système immunitaire, cancer**


# Résumé en anglais suivi des mots-clés en anglais

The inhibitor of kB kinase (IKK) is the master regulator of NF-kB signalling and plays a fundamental role in the activation and regulation of inflammatory processes and innate immune response. Upon receptor stimulation, IKK induces the translocation of NF-kB factors into the nucleus where they regulate gene expression.

The IKK kinase complex is represented by the canonical complex (IKK1/IKK2/NEMO) and the alternative complex ($IKK1_2$). Although structural information is available for IKK1/2, detailed knowledge of the mechanisms of IKK- substrate interactions is lacking.

In this work, we have identified a novel YDDΦXΦ docking motif present in the disordered C-terminus of IκBα. This motif is necessary for the interactions with IKK1 and IKK2. A direct correlation between the affinity of the YDDΦXΦ docking interaction and IKK2 kinase activity is observed. We provide a 3D structural model of this interaction based on x-ray diffraction and X-linking mass spectrometry data. The YDDΦXΦ motif is found in other IKK substrates (IκBβ, p100 and IRF7).

Moreover, preliminary analyses indicate that this motif is recognized by the NEMO regulatory subunit. Finally, synthetic peptides containing the YDDΦXΦ motif are able to inhibit substrate phosphorylation by IKK2, opening the way to future development of a new class of selective IKK inhibitors.


**Keywords : protein interaction, IKK complex, NF-κB, IκBα, inflammation, immune system, cancer**