UAB
Universitat Autònoma de Barcelona

Developing a Pipeline for the Discovery of b-HLH-LZ-Based Therapeutics: Insights from Omomyc Engineering

Hugo Thabussot

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. The access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

Developing a Pipeline for the Discovery of b-HLH-LZ-Based Therapeutics: Insights from Omomyc Engineering

Doctoral Thesis

Hugo Thabussot

Directors: Dr. Laura Soucek and Dr. Marie-Eve Beaulieu

Peptomyc s.l.

Academic tutor: Laura Soucek

PhD program in Biochemistry, Molecular Biology and Biomedicine

Department of Biochemistry and Molecular Biology

Universitat Autònoma de Barcelona

Barcelona, 2024

The tutor and co-director **Dr. Laura Soucek** and the co-director **Dr. Marie-Eve Beaulieu**, Certify:

That the experimental work and the writing of the memory of this doctoral thesis entitled “Developing a Pipeline for the Discovery of b-HLH-LZ-Based Therapeutics: Insights from Omomyc Engineering” have been performed by Hugo Thabussot under their direction and consider that it is suitable to be presented for the degree of PhD in Biochemistry, Molecular Biology and Biomedicine at Universitat Autònoma de Barcelona.

Barcelona, 13th of September 2024.

Tutor and co-director:	**Co-director:**
Dr. Laura Soucek	**Dr. Marie-Eve Beaulieu**

Ph.D. Student: Hugo Thabussot

Table of Contents

Abbreviations 1

Summary 5

I. **Introduction** 9

1. The *c-myc* proto-oncogene 9

1.1. The structural characteristics of the MYC protein and its main interactors 9

1.2. The physiological activity of MYC 14

1.3. Mechanisms of MYC regulation and loss of them in cancer 17

1.4. MYC's role in cancer 21

1.5. MYC inhibition and its challenges 23

2. Omomyc, a first-in-class MYC inhibitor 26

2.1. The use of the Omomyc transgene demonstrates the viability of anti-MYC therapies to treat cancer 26

2.2. The MAX b-HLH-LZ displays intrinsic cell penetration properties 28

2.3. The Omomyc mini-protein is a viable anti-cancer drug candidate 29

3. The emergence of the b-HLH format as a new class of therapeutic protein 35

3.1. Omomyc, a benchmark to surpass? 35

3.2. The lessons learned from decades of characterization and discoveries 39

3.3. Paths and tools to better understand and optimize b-HLH-LZ drug candidates 41

II. **Hypotheses and Objectives** 47

III. **Material and Methods** 51

1. Bacterial transformation and protein expression 51

2. Omomyc and Max* mini-proteins purification 52

3. HisTag purification 52

4. StrepTag purification 53

5. Mini-protein solubility assay 53

6. Circular dichroism 54

6.1. Far-UV spectrum measurement 54

Table of Contents

6.2. Thermal denaturation and protein melting temperature determination 55

7. Cell lines and cell culture 56

8. Endocytosis assay 56

9. Electro-mobility shift assay 58

9.1. Native gel composition 58

9.2. Samples preparation and assay conditions 59

10. Mini-proteins DNA-binding characterization by Enzyme-Linked Immunosorbent Assay (ELISA) 59

11. Co-immunoprecipitation (co-IP) 60

11.1. Cell lysate preparation 60

11.2. Co-IP protocol 60

11.3. Western Blot analysis 61

12. Cell viability and proliferation assay 61

13. Immunofluorescence (IF) 62

14. Competition assay for StrepTag concentration determination in blood-derived matrices 62

14.1. Blood-derived samples 62

14.2. Competition assay 63

IV. **Results** 65

1. Modulation of the cell penetration activity and nuclear localization of b-HLH-LZ proteins using short-stretch mutations and cell penetration tags 65

1.1. Design of the constructs 66

1.2. A simple, HisTag affinity chromatography-based process enabled the purification of V84 to V89 70

1.3. CPP-Omomyc fusion proteins have comparable secondary structures and retain the ability to bind Max*, but show signs of impaired DNA binding by CD 74

1.4. Both the HisTag and CPPs tag impair the variants DNA binding ability 79

1.5. The R9 tag and Omomyc LZ truncation affect solubility and stability 86

1.6. Short stretch mutations may strongly affect Omomyc structure and its ability to bind DNA and MAX 90

Table of Contents

1.7. All variants display the capacity to penetrate cells through ATP-dependent routes and the CPP might affect their cell penetration kinetics 95

2. Impact of the net charge on the Omomyc-derived mini-proteins mode of action 99

2.1. Design of the constructs 100

2.2. The StrepTag-II is a viable replacement for the HisTag 102

2.3. Net charge modifications may impact the variants solubility profile 106

2.4. The charge variants retained the ability to bind MYC and MAX, despite a destabilized structure 109

2.5. The charge variants retained the abilities to bind DNA and to discriminate specific from non-specific DNA sequences, as homodimer or heterodimer with Max* 113

2.6. The charge variant ability to inhibit cell growth does not correlate with their ability to penetrate cells 119

2.7. A newly developed assay highlights the charge variants different behavior in plasma 122

V. **Discussion** 131

Section 1 132

Section 2 143

VI. **Conclusions** 156

Section 1 156

Section 2 156

VII. **References** 157

Acknowledgement 184

Abbreviations

2YT: 2x Yeast extract Tryptone
ACF: Amicon® Centrifugal Filter (tube)
AF488: Alexa Fluor 488
ATCC: American Type Culture Collection
ATP: Adenosine Triphosphate
BBB: Blood Brain Barrier
b-HLH-LZ: basic-helix-loop-helix-leucine zipper
BR: Basic Region
BRD4: Bromodomain-containing Protein 4
BSA: Bovine Serum Albumin
CAK: Cdk-activating kinase
CD: Circular Dichroism
CDK: Cyclin-Dependent Kinase
CDKI: Cyclin-Dependent Kinase Inhibitor
CDR: Complementarity-Determining Region
CDX: Cell line-Derived Xenograft
CEX: Cation Exchange Chromatography
ChIP-qPCR: Chromatin Immunoprecipitation-Quantitative Polymerase Chain Reaction
CPP: Cell Penetrating Peptide
CRD: Coding Region instability Determinant
DAPI: 4',6-diamidino-2-phenylindole
DMEM: Dulbecco's Modified Eagle Medium
DMSO: Dimethyl Sulfoxide
DNA: deoxyribonucleic acids
DTT: Dithiothreitol
E-Box: enhancer box
EDTA: Ethylenediaminetetraacetic Acid
EGFR: Epidermal Growth Factor Receptor
EGTA: ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
ELISA: Enzyme-Linked Immunosorbent Assay
EM-1: Endomorphin-1 peptide
EMSA: Electro-Mobility Shift Assay
FBS: Fetal Bovine Serum
FDA: United States of America Food and Drug Administration
FP: Fluorescence Polarization
FSC-A: Forward Scatter Area
GFP: Green Fluorescent Protein
GMP: Good Manufacturing Practices

GSEA: Gene Set Enrichment Analysis
HCP: Host Cell Protein
IF: Immunofluorescence
IL: Interleukin
IRES: internal ribosome entry sites
LB: Luria-Bertani (broth)
LLOD: Lower Limit of Detection
mAb: Monoclonal Antibody
MALDI-TOF: Matrix-Assisted Laser Desorption/Ionization (mass spectrometry)
MAX: Myc-associated factor X
MB: MYC box
MGA: MAX Gene-Associated protein
MIZ-1: MYC-interacting Zinc finger protein 1
MNT: Max's Next Tango
MRE: Mean Residual Ellipticity
MS: Mass Spectrometry
MSD: Merck Sharp & Dohme
MWCO: Molecular Weight Cut-Off
MXD: MAX Dimerization Protein
Ncl: Nucleolin
NHDF: Normal Human Dermal Fibroblast
NMR: Nuclear Magnetic Resonance
NLS: Nuclear Localization Signal
Npm1: Nucleophosmin 1
NSCLC: Non-Small Cell Lung Carcinoma
ORC: Origin Recognition Complex
pAb: Polyclonal Antibody
PBS: Phosphate-buffered Saline (solution)
PDAC: Pancreatic Ductal Adenocarcinoma
PDB: Protein Data Bank
PES: Polyethersulfone
PFA: Perfluoroalkoxy Alkane
pI: Isoelectric point
PK: Pharmacokinetics
PLA: Proximity Ligation Assay
PMN: Proximal MYC Network
QBP1: Polyglutamine-Binding Peptide 1
Rb: retinoblastoma protein
RBP: RNA-binding Protein
RECIST: Response Evaluation Criteria In Solid Tumors

RNA: Ribonucleic Acid
RPMI: Roswell Park Memorial Institute (medium)
RT: Room Temperature
SCLC: small cell lung cancer
SDS-PAGE: sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SPPS: Solid-Phase Peptide Synthesis
SPR: Surface Plasmon Resonance
SSC-A: Side Scatter Area
TAD: Transactivation Domain
TAM: Tumor-Associated Macrophage
TCEP: Tris(2-carboxyethyl)phosphine
TEMED: N,N,N',N'-tetramethylethylenediamine
Tm: Melting Temperature
TNBC: Triple-Negative Breast Cancer
Treg: regulatory T cell
UTR: Untranslated Region
UV: Ultraviolet
VEGF: Vascular Endothelial Growth Factor
VHIR: Vall d'Hebron Institute of Research

Summary

MYC is a transcription factor playing a central role in human biology through its regulation of processes such as cellular proliferation and differentiation or metabolism. MYC was shown to be deregulated in the majority of human cancers through overexpression and loss of normal regulatory controls and to strongly contribute to tumor growth and maintenance. For this reason, MYC inhibition represents a particularly appealing therapeutic target in oncology. However, MYC's intrinsic disordered nature and nuclear localization made its direct inhibition challenging. In 2021, Peptomyc brought to the clinic the first direct mini-protein MYC inhibitor, named Omomyc, which acts as a MYC dominant negative, with intrinsic cell penetration properties. This major milestone in oncology has paved the way to the development of a new class of therapeutic proteins, in the shape of b-HLH-LZ domains, and was an inspiration for the design of similar molecules.

The work presented in this thesis aimed at establishing the foundations for the efficient development of new Omomyc-based drug candidates, through the development of relevant tools and characterization methods. We did so while exploring two strategies to enhance the anti-tumoral Omomyc activity:

First, we investigated the modulation of the cell penetration efficiency and nuclear localization of Omomyc-derived variants through the introduction of short-stretch mutations and fusion to cell-penetrating peptides (CPPs). Through the study of a panel of 23 Omomyc- or Max*-derived variants, we determined that the fusion of CPPs to Omomyc variants could improve their cell penetration kinetics and that their overall cell penetration efficiency could be improved through their fusion to an Epidermal Growth Factor Receptor (EGFR)-targeting sequence. However, this improvement came at the cost of reduced structural stability, less specific DNA-binding ability and decreased solubility. Furthermore, truncations to the leucine zipper domain affected the protein's stability, with larger truncations leading to poor recombinant expression and reduced solubility. Finally, we determined that the use of HisTag could be valuable to streamline the purification of large variants panel, but impacted their DNA-binding ability, critically limiting these variants evaluation.

The second part of this thesis explored how changes in the net charge of Omomyc-derived proteins influence their solubility, their binding to MYC,

MAX and DNA, as well as their cellular activity. First, we could demonstrate that variants with a lower isoelectric point (pI) could exhibit more advantageous solubility profiles both at neutral and basic pH conditions. Despite having an overall reduced structural stability, these charge-modified variants retained the ability to bind to both MYC and MAX, suggesting that their core binding functions were mainly preserved despite structural changes. However, the variants with the most profound pI change presented significantly reduced affinity for DNA, which was seemingly rescued through heterodimerization with Max*. Interestingly, these variants capacity for cell penetration did not correlate with their effectiveness in inhibiting cell proliferation, suggesting that additional factors, such as intracellular stability, nuclear localization, or subtle changes in MAX and/or MYC affinity could play pivotal roles in determining their therapeutic efficacy. Furthermore, a newly developed assay highlighted the charge variants different behavior in plasma and whole blood and suggested that Omomyc variants net charge modulation could impact their pharmacokinetic profile.

Overall, the results of this thesis provide important insights into the challenges and opportunities of developing Omomyc-derived mini-proteins as therapeutic agents. While the fusion of CPPs enhances cell penetration, trade-offs in structural stability and functional activity must be carefully balanced. Charge mutations show promise for improving solubility and stability, which could enhance in vivo pharmacokinetics. Overall, this work advances the understanding of how MYC inhibition can be achieved through rational design of mini-proteins and lays the groundwork for further optimization of b-HLH-LZ proteins as viable cancer therapeutics.

“People who think they know everything are a great annoyance, especially to those of us who do.”

Isaac Asimov

I. Introduction

1. The *c-myc* proto-oncogene

Originally described as a cellular analog of the avian viral *v-myc* in 1979[1,2], *c-myc* is without a doubt, one of the most studied proto-oncogenes, with an average of nearly 2,500 articles mentioning the term "myc" published every year over the last decade. This gene encodes the MYC transcription factor and is part of the *myc* family together with *l-myc* and *n-myc*. Several isoforms of the MYC protein are encoded by the *c-myc* gene[3], but they all display a transactivation domain (TAD) on their N-terminus that can recruit the transcription machinery, a central region and a characteristic basic-helix-loop-helix-leucine zipper (b-HLH-LZ) domain enabling the MYC protein to dimerize and bind deoxyribonucleic acids (DNA). MYC exerts most of its functions through DNA binding and must associate with the MYC-Associated factor X (MAX) to do so. Upon heterodimerization with the MAX protein via their respective b-HLH-LZ domains, MYC gains the ability to bind specific DNA sequences called enhancer box (or E-Box, CANNTG)[4–6], located upstream of target gene coding sequences. Importantly, MYC transcriptional activity is crucial during development, and a *c-myc* gene knockout in mice was found to be systematically lethal before the day 10.5 of gestation[7]. Through the binding of more than 30,000 E-Box sequences spread throughout the genome, MYC can regulate up to 15% of the identified human genes[8,9], justifying its common designation as "master regulator" or "orchestrator" of genome transcription.

1.1. The structural characteristics of the MYC protein and its main interactors

The MYC protein is characterized by an intrinsically disordered structure, due to a certain lack of defined secondary and tertiary structure. Nevertheless, this flexible protein is able to bind the MAX b-HLH-LZ domain, and as a heterodimer, to recognize E-Box DNA sequences with very high affinities, with dissociation constants in the low nM range[10]. A clear increase in α-helix content is observed upon binding of MYC to MAX, as extensively demonstrated by circular dichroism (CD)[11,12]. During this process, both respective HLH-LZ domains fold, while assembling in a highly stable

complex, although their respective basic region (BR) remain mainly disordered in the absence of DNA[13]. The heterodimerization between MYC and MAX is driven by the complementarity of their LZ domains, which form a coiled-coil structure upon dimerization. While unfavorable electrostatic interactions destabilize MYC homodimers, they are avoided in the heterodimers with MAX, which instead achieve high stability through favorable electrostatic interactions, hydrogen bonds formation and van der Waals forces between specific residues at the interface of their LZ domains (Figure 1). The HLH domains were also shown to interact closely forming a four-helix bundle arrangement at the center of the b-HLH-LZ dimer, which coordinates the orientation and connects both the BR and LZ at each end[14]. Of note, the assembly of MYC and MAX HLHs, stabilized in part by the interaction between MYC's L396 and p22 MAX isoform F43 residues, form the only domain in that dimer that could be assimilated to a hydrophobic core, the rest of the dimer structure being plane and rather featureless[14].

The MYC/MAX heterodimer achieves specificity in DNA binding through mechanisms that combine conformational selection with induced-fit folding[15]. Initially, the basic regions of MYC and MAX are only partially folded, giving them the flexibility to align with the major groove of DNA. This pre-alignment helps positioning basic residues to interact with the DNA backbone. Upon encountering a specific DNA sequence, like the canonical E-box (CACGTG), the MYC/MAX heterodimer BRs undergoes conformational selection. This is followed by an induced-fit process, where the protein structure becomes more stable and adopts an α-helical conformation optimized for specific DNA binding. In contrast, when the heterodimer binds to non-specific DNA sequences, the BRs α-helical conformation is not fully stabilized, resulting in lower binding affinity. This leads to a strong preference for specific sequences, with the full length MYC/MAX heterodimer showing approximately 100-fold higher affinity for E-Box sequences compared to non-specific ones[14,15].

The crystal structure of the MYC and MAX b-HLH-LZ domains heterodimer bound to a canonical E-Box CACGTG (Figure 1A), solved by Nair and Burkley in 2003[16], demonstrated the importance of the H359, E363 and R367 MYC residues for specific recognition of the sequence, by forming a series of specific hydrogen bond with the sequence bases. Fisher and Goding highlighted in an early study the importance of the glutamic acid residue, when they reported that its mutation to a glutamine destabilized MYC's

ability to bind DNA[17]. On the other hand, B. Amati and colleagues showed that the mutation of the residues R364, R366 and R367 for alanine residues would suffice to abrogate MYC's ability to bind DNA[4]. Those results were further confirmed by his group in a subsequent study showing that a recombinant MYC^{RA} b-HLH-LZ domain did not display any significant affinity for DNA[18]. More recently, Pellanda et al. demonstrated that both the recombinant b-HLH-LZ domain and full length MYC^{HEA} (H359A, E363A) mutant would still retain a high affinity for DNA, although full length MYC^{HEA}' preferred DNA sequence would shift to a non-canonical E-Box (CACGTC)[18].

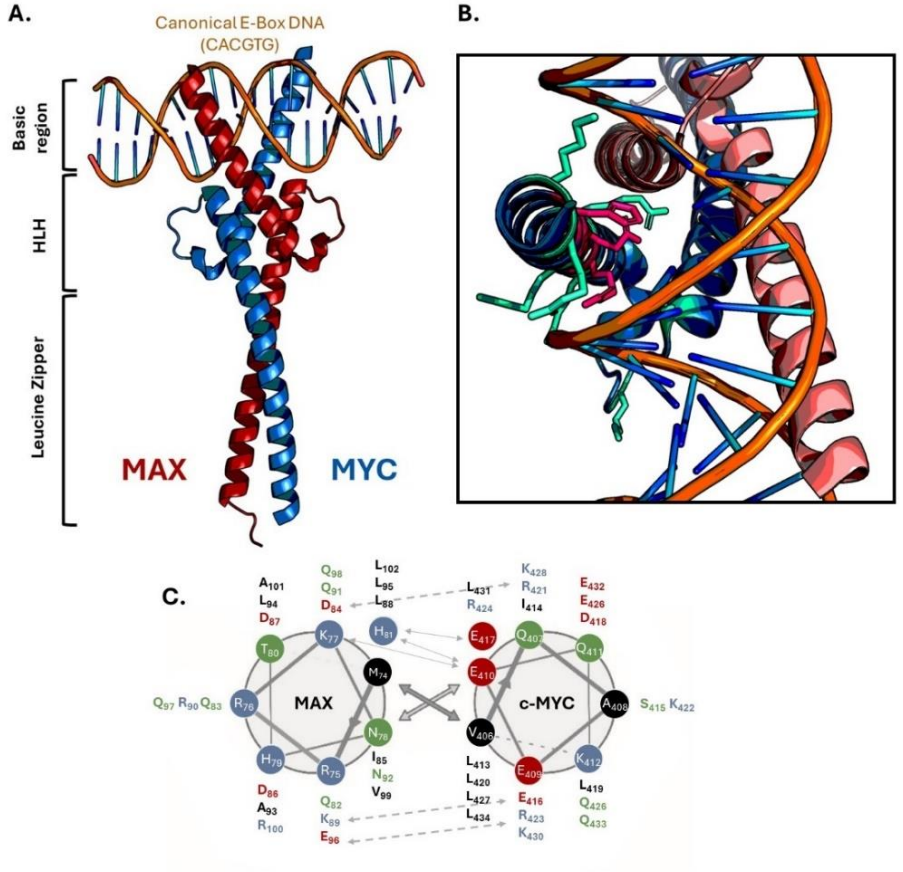


Figure 1, A. *X-ray crystallography structure of the MYC/MAX b-HLH-LZ heterodimer complexed to the canonical E-Box DNA sequence (PDB entry: 1NKP).* ***B.*** *Close-up representation of the MYC basic region and the major groove of DNA. The side chains of the residues interacting with DNA were represented, those forming non-specific electrostatic interactions with the DNA backbone are colored in green and the ones participating to MYC's specificity for the E-Box are colored in pink.* ***C.*** *Helical-wheel representation for the interfacial interactions in the LZ domains during MYC-MAX interaction. Adapted from Beaulieu et al., 2020.*[14]

While sequence recognition is mediated through the formation of H-bonds, overall DNA binding is enabled by electrostatic interactions between the negatively charged DNA phosphodiester backbone and the basic residues located in MYC's basic region (K355, R356, R364, R366, R367), Helix 1 (K371) or in its loop (K389 and K392)[16]. MYC/MAX dimers binding to DNA was seen to not only stabilize their respective basic region folding but their whole b-HLH-LZ domain. In turn, the MAX's LZ homodimerization stability was shown to impact MAX b-HLH-LZ homodimers' affinity for DNA[19], and a similar observation could be expected for MYC/MAX-derived dimers, although that was not formally demonstrated yet, to my knowledge.

Apart from its b-HLH-LZ domain, the MYC sequence comprises 6 short sequences well-conserved across the *myc* family members, called MYC homology boxes or simply MYC boxes (MB), all throughout its transactivation domain (TAD) and central region. MB0 and MB1 are found near the N-terminus of MYC and play an important role in MYC stability[20–23]. Additionally, the negatively charged MB0 was recently proposed to have a modulating activity on MYC DNA binding by interacting with the intramolecular BR[24]. The MYC's TAD encompassing MB0, MBI and MBII mediates MYC interactions with many transcription activators and repressors, such as TRRAP, GCN5, TIP48 or p400, and is therefore critical for MYC's transcriptional activity, as skillfully summarized by Conacci-Sorell and colleagues[25]. MYC's TAD was also shown to recruit and induce the assembly of the ribonucleic acid (RNA) polymerase II through its interaction with the TATA binding protein[26].

In addition, MBIIIa (missing in MYCL) and MBIIIb were shown to mediate MYC's activity on chromatin remodeling through their interactions with proteins like WDR5[27] or histone deacetylase[28]. Moreover, the MBIIIb is localized in a larger MYC central region ($MYC_{226-269}$), commonly referred to as PEST, since its sequence is enriched in proline (P), glutamic acid (E), serine (S) and threonine (T) residues, which were shown to play an important role in MYC's protein stability[29].

Finally, the MBIV was shown to interact with the host-cell factor 1 protein, which contributes to MYC transcriptional activation[30]. Of note, through the study of a ΔMBIV MYC variant, Cowling et al. highlighted the importance of this region in DNA binding, apoptosis, transformation and G2 arrest[31].

Importantly, two nuclear localization signal (NLS) sequences enable MYC to translocate into the nucleus after translation. The first to be identified, and probably the most potent, is located between residues 320 to 328 near the MBIV, while the second one is located in MYC's BR[32–35].

Together with other b-HLH-LZ proteins, MYC and MAX are part of a network commonly called the proximal MYC network (PMN) (Figure 2). Among the proteins that are part of the PMN, the Max dimerization protein 1 (MXD1), its paralogs MXD2-4, the Max's Next Tango protein (MNT) and the MAX gene-associated protein (MGA) were described to antagonize MYC activity, through at least three distinct mechanisms. First, they compete with MYC for MAX heterodimerization. Since MAX is MYC's obligate partner to bind DNA, MYC is kept away from DNA through MAX sequestration when MAX levels are limiting. Second, as heterodimers with MAX, the MXD proteins and MNT also compete for binding to the same DNA sequences as the MYC/MAX dimers. Finally, while MYC binding to E-Box enables the recruitment of transcription activators of a target gene, the MXD proteins recruit transcription co-repressors instead[25]. The activation of these MYC antagonists has been shown to induce cell cycle arrests and differentiation, as one could expect. Interestingly, MNT deficiency in mice results in defective embryonic growth and decreased viability. In contrast, the development of mice with MXD1-MXD4 single gene knock-out was not significantly affected, but the mice showed signs of delayed terminal granulocyte differentiation (MXD1), were subject to ectopic cell proliferation and more prone to develop neoplasia (MXD2), or displayed increased sensitivity to apoptotic signals in some tissues (MXD3)[25].

Another member of the network is MGA, a 336 kDa protein with the particularity to comprise a T-box DNA binding motif in addition to the b-HLH-LZ domain, giving it the ability to specifically bind distinct DNA motives simultaneously. Although it might not be relevant in more physiological contexts, overexpression of MGA was observed to repress MYC activity, similarly to the MXD proteins[36].

Finally, an important part of MYC's transcriptional repressing activity is mediated through its interaction with the MYC-interacting Zinc finger protein 1 (MIZ-1) through its b-HLH-LZ domain. Although competing with MAX for MYC binding, MYC/MIZ-1 dimerization is not accompanied by a dramatic MYC structure folding and the ability of this heterodimer to bind DNA is

unclear[37]. The implication of the MYC/MIZ-1 interaction will be discussed later on in the text.

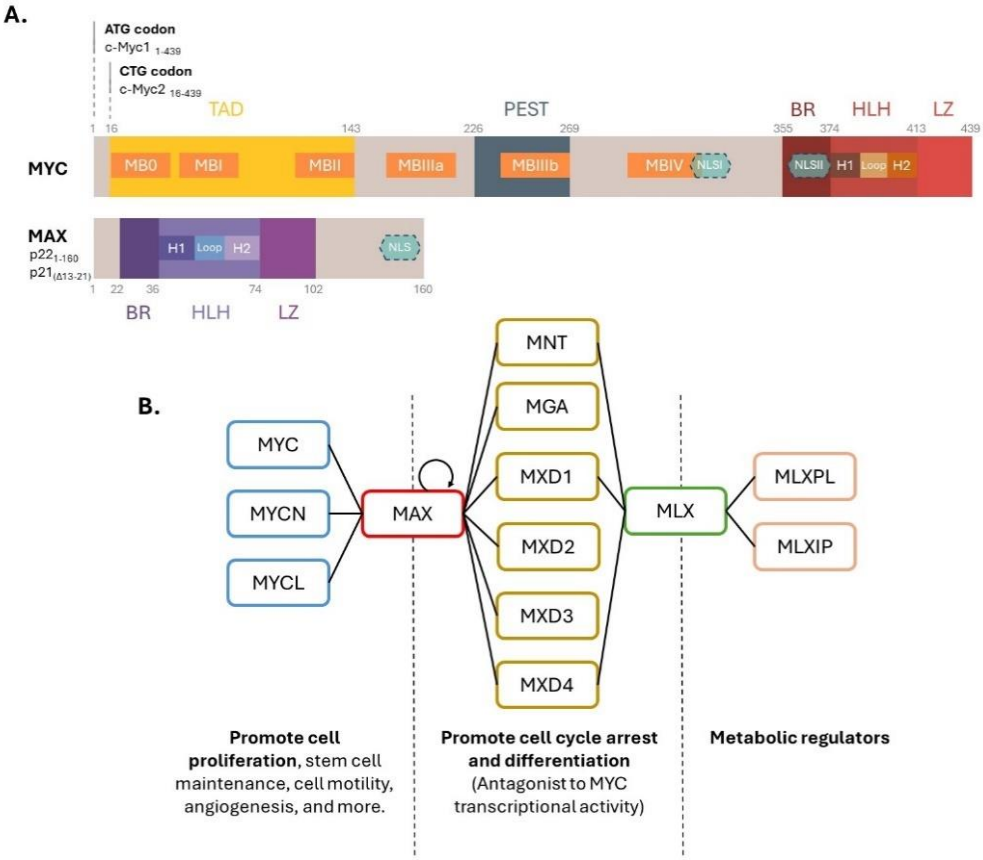


Figure 2, A. *Schematic representation of the MYC and MAX protein regions. Figure adapted from Beaulieu et al., 2020*[14]. ***B.*** *Representation of the extended MYC network and the potential heterodimerization formation among them.*

1.2. The physiological activity of MYC

The *myc* gene can be found in the genome of virtually all vertebrates[38] and plays a crucial role in a broad range of cellular processes encompassing metabolism, protein synthesis, cell cycle progression, chromatin modulation or ribosomal synthesis, among others[39]. Its activity, strongly supporting cell proliferation, is critical during embryogenesis and in processes such as wound healing or homeostatic tissue maintenance.

Interestingly, MYC was seen to induce epidermal cells dedifferentiation upon wounding or mechanical stimuli[40]. Of note, c-*myc* is actually part of the Yamanaka four genes, together with *Oct3/4*, *Sox2* and *Klf4*, which, when activated simultaneously, induce dedifferentiation of mouse fibroblasts into a pluripotent state[41]. Additionally, MYC and MYCN also play important roles in pluripotency maintenance by downregulating differentiation-inducing genes, such as the *HOX* genes[42], lineage-specific genes, such as *GATA6*[43], or by inhibiting differentiation-promoting pathways like the mitogen-activated protein kinase (MAPK) pathway[44,45]. Interestingly, although it can have a clear upregulating effect on dedifferentiation, MYC was also shown to have the opposite activity and support stem cell differentiation in specific contexts[46,47].

Rapidly upregulated after injuries[48], MYC orchestrates cell proliferation by directly inducing the expression of cyclins and cyclin-dependent kinases (CDKs), particularly cyclins A, B, D, and E, as well as CDK-4 and CDK-6[49,50]. These components are essential for driving cells through the different phases of the cell cycle, particularly the G1/S and G2/M transitions[49–53]. Transcription of these cyclins and CDKs facilitates the phosphorylation of the retinoblastoma protein (Rb), which releases E2F transcription factors that activate genes necessary for DNA replication[54,55]. Additionally, MYC represses cyclin-dependent kinase inhibitors (CDKIs) such as p15, p21, and p27 through various mechanisms, including direct transcriptional repression via its interaction with MIZ-1[56–58], induction of miRNAs[59], chromatin remodeling[60] or DNA methylation[61], among others[49,62]. This repression of CDKIs ensures that cyclin/CDK complexes remain active, driving the cell cycle forward. In addition to their direct regulation, MYC also enhances the cyclin/CDK complexes activity by indirectly promoting their activating phosphorylation patterns through the upregulation of the Cdc25 phosphatases and of the of Cdk-activating kinase (CAK)[53,63].

For what refers to DNA replication, MYC supports it by upregulating key components of the DNA replication machinery, including the origin recognition complex (ORC), Cdc6, and MCM proteins, which are necessary for initiating and progressing through DNA synthesis[49,64,65].

Moreover, MYC promotes ribosome biogenesis by upregulating components of the RNA polymerase I and III apparatus, enhancing the production of rRNA and tRNA, which support protein synthesis required for cell growth and division[38,66,67].

Through these combined actions, MYC acts as a central regulator of cell cycle progression, coordinating the expression of genes involved in proliferation, DNA replication, and ribosome biogenesis, while simultaneously suppressing cell cycle inhibitors to ensure robust cellular proliferation.

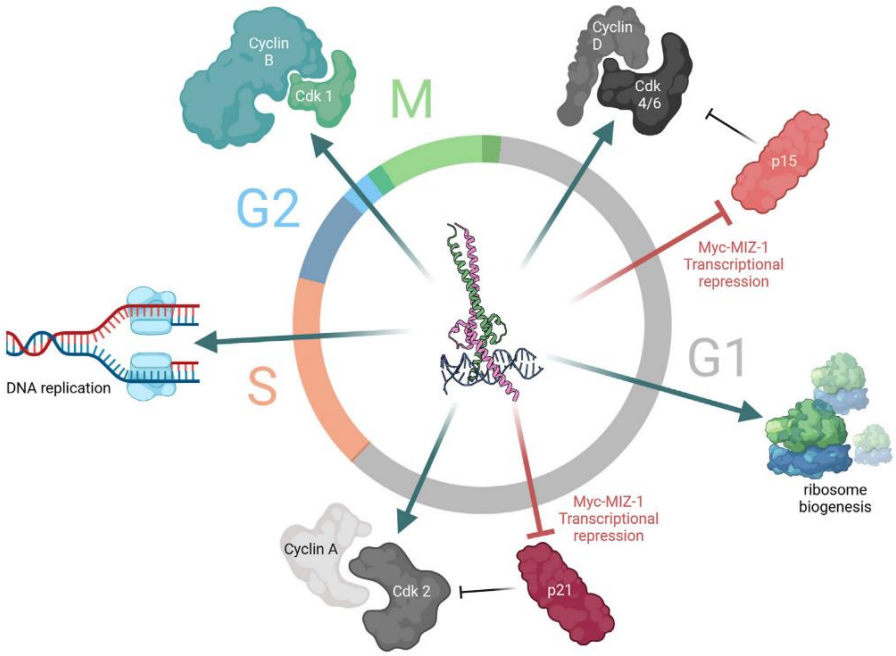


Figure 3, *MYC is a key regulator of the cell cycle and promote its progression at all stages through its transcription activity as activator of cyclins and CDKs, ribosome biogenesis and DNA replication, or as a CDK inhibitors repressor. Figure reproduced from Ascanelli et al. 2024.*[38]

At the metabolic level, MYC activation supports protein and DNA synthesis by both enhancing rapid adenosine triphosphate (ATP) production through the induction of glycolysis and lactic fermentation, and the upregulation of amino acids and nucleotides synthesis pathways. Moreover, MYC can induce both fatty acid synthesis and lipid catabolism. Therefore, MYC supports the generation of the energy and the building blocks necessary for cellular growth and division[49,68,69]. This aspect of MYC biology needs to be however considered with caution since it was mainly studied in pathological settings and might therefore not be completely relevant in physiological contexts where MYC is subject to tighter regulations.

Overall, MYC was shown to be important in the maintenance of the physiological homeostasis of organs that need constant cell renewal, such

as the intestine (maintenance of intestinal crypts' stemness and regeneration)[70,71], or the skin (keratinocyte proliferation, dermal stem cells maintenance, hair growth)[46,72].

In the context of tissue regeneration processes, through the induction of the mir-9 microRNA expression, MYC and MYCN downregulate the translation of E-cadherin, resulting in increased cell motility and invasion[73,74], which in turn enables efficient wound closure[75]. In addition, during the development or a wound healing process, MYC plays a crucial role in promoting angiogenesis mainly through enhancing the expression of vascular endothelial growth factor (VEGF). In this context, MYC also represses thrombospondin-1 (TSP-1), an inhibitor of angiogenesis, tipping the balance in favor of vessel growth[76].

Interestingly, although the expression of MYC and MYCN was shown to be indispensable during the embryogenic development[7,77], MYC overexpression in embryos endothelial cells led to their death due to multiple hemorrhagic lesions[78]. These experiments, suggesting that MYC not only induces angiogenesis but also regulates the vascular permeability[78], illustrate that the MYC family's broad functions are highly context-specific[9] and require to be tightly controlled to avoid any pathological development[79]. In fact, the MYC family proteins are strictly repressed in most of the cells of a healthy adult and punctual expression is only allowed in a few highly dividing cells or during tissue repairs.

1.3. Mechanisms of MYC regulation and loss of them in cancer

In physiological conditions, MYC levels are tightly regulated to enable its broad transcriptional activity only when it is needed.

At the transcriptional level, the regulation of MYC involves a multitude of signaling pathways such as Wnt, Notch, TGF-β or BMP, which can upregulate and/or downregulate MYC in a context-specific manner[80]. The MYC-coding mRNA transcription through activation of its main promoter, p2, requires the interaction with various transcription factors, such as CNBP[81], FBP[82], and TCF[83], as well as the Bromodomain-containing protein 4 (BRD4), a BET domain protein that facilitates transcriptional elongation through the promotion of the RNA polymerase II activity[84]. Several BET

inhibitors were shown to decrease MYC expression level, proving the importance of MYC regulation by BRD4[85,86]. Interestingly, the *c-myc* gene is rather isolated from other coding sequences within the genome[38], and the induction of non-conventional DNA structures, including Z-DNA[87], G-quadruplexes[88], and single-strand bubbles in the MYC promoter region are employed to regulate *c-myc* transcription[8].

After its transcription, *MYC* mRNA is then subject to stringent post-transcriptional/translational control. Notably, the 5' untranslated region (UTR) of *MYC* mRNA contains internal ribosome entry sites (IRES), which interact with RNA-binding proteins (RBPs) such as hnRNPA1, hnRNPK, PCBP1, among others[89]. These RBPs influence the initiation and efficiency of MYC translation. *MYC* mRNA includes a coding region instability determinant (CRD) that contributes to mRNA destabilization under conditions of tRNA shortage, leading to ribosomal stalling and subsequent degradation. This process is prevented if the CRD is bound by a CRD-binding protein, which is highly expressed in fetal tissues but significantly reduced in adults, allowing for more rapid *MYC* mRNA turnover[38,90–92]. Furthermore, RBPs such as CELF1 and HuR compete for binding to the *MYC* mRNA 3' UTR sequence. These two RBPs have opposite effects over the *MYC* mRNA stability: while HuR stabilizes the mRNA and promotes translation, CELF1 3' UTR binding downregulates MYC translational output[93]. Additionally, several miRNAs, namely *let-7a*, *miR-34b/c*, or *miR-145*, and ribosomal proteins, such as S14, L5 or L11, promote *MYC* mRNA degradation through binding to its 3' UTR region[89,94–98]. Altogether, these regulations result in a very short half-life of *MYC* mRNA, estimated to be between 10 to 20 min[99].

Finally, the MYC protein itself is subject to intense regulation through post-translational modifications, ensuring that its presence within the cell is short-lived, with a half-life of approximately 15 to 30 minutes[3]. The primary mechanism for MYC protein degradation involves a series of phosphorylation events that label it for recognition by E3 ubiquitin ligases, leading to its degradation via the ubiquitin-proteasome pathway. Phosphorylation of MYC at serine 62 stabilizes the protein during cell cycle re-entry, particularly in response to mitogen signals, and is mediated by kinases like ERK as part of the RAS/RAF/MAPK signaling cascade[100,101]. However, this phosphorylation was seen to stabilize MYC only temporarily, as it also serves as a prerequisite for subsequent phosphorylation at threonine 58 (T58) by glycogen synthase kinase 3β (GSK3β). These two

phosphorylations (S62 and T58) induce MYC degradation through its binding by the E3 ubiquitin ligase FBW7. Interestingly, these phosphorylations occurs at different times during the cell cycle, with S62 phosphorylation predominating during early G1 phase, and T58 phosphorylation occurring later[20,21].

The recognition of MYC by FBW7 can also involve additional phosphorylation sites at T244 and T248[102]. These sites are particularly relevant when MYC is overexpressed, and provide alternative recognition motifs for FBW7, leading to MYC's ubiquitination and subsequent proteasomal degradation. Of note, many other E3 ligases, such as Skp2, can also target MYC for ubiquitination[103]. Finally, additional MYC b-HLH-LZ phosphorylation at residues T358, S373 and T400 by the PAK2 kinase were shown to destabilize the MYC/MAX association and decrease MYC affinity for DNA[10].

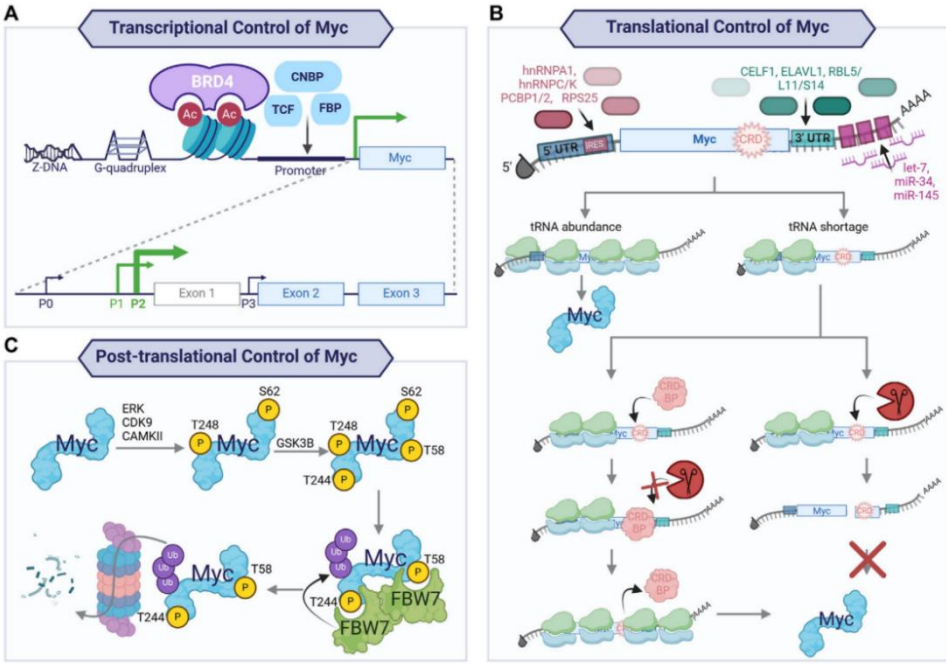


***Figure 4**, Schematic summary of the mechanism of MYC regulation in physiological conditions at (**A**) the transcriptional level, (**B**) the post-transcriptional / translational level and (**C**) at the post-translational level. Figure reproduced from Ascanelli et al. 2024.*[38]

In cancer, *c-myc* is one of the most, if not the most, deregulated genes[104,105]. While other oncogenes or tumor suppressors like *TP53* are typically mutated in cancer, MYC deregulation was shown to occur mainly through the alteration of its expression levels, through different mechanisms[104,105]. Amplification of the *c-myc*, *n-myc* or *l-myc* genes was observed in numerous cancer types and occurs through copy number amplification and chromosomal translocation. These events result in an increase of the *myc* gene copy number and/or place it under the regulation of other promoters, driving its aberrant expression. MYC, MYCN or MYCL overexpression are commonly observed in breast cancers (notably in triple negative breast cancer), ovarian cancers, prostate cancers or neuroblastoma, among many others, and correlate with poor prognosis[106]. A pan-cancer analysis revealed that MYC and/or its paralogs were amplified in 28% of all considered samples encompassing a total of 33 cancer types[107]. Interestingly, several studies have shown that elevated MYC levels are, however, not always essential for supporting tumor growth and that, perhaps, sustained expression even at relatively low MYC concentration, suffices to support tumorigenesis[108]. For example, mutations in the MYC sequence at residues T58 and S62, key phosphorylation sites involved in MYC degradation, have been observed to prolong MYC's half-life and contribute to tumorigenesis, stabilizing its levels over time, though not necessarily increasing its concentration[80].

Although deregulated in the majority of cancers, MYC is not always the driver oncogenic lesion. Indeed, MYC is positioned at a signaling bottleneck, downstream of other important oncogenes, such as TGF-β[109], EGFR, Wnt, or Hippo[110]. In cancer, their respective signaling pathways funnel through MYC and tend to upregulate its activity[110]. Consequently, even when *c-myc* is not the driver oncogenic lesion, it still has an important role in the progression of the disease. Its central position in oncogenic signaling contributes to making many tumors “addicted” to MYC, and even short interruption of its function have proven to be an efficient way to impair tumorigenesis[80].

Several studies suggest that deregulated MYC levels increase its propensity to bind non-canonical E-Boxes (CANNTG) or even “non-E-box” DNA sequences. Increased invasion of these alternative sequences by MYC/MAX dimers would promote the aberrant regulation of genes and contribute to give MYC its oncogenic functions[111–113].

1.4. MYC's role in cancer

In a context where MYC is deregulated and constantly activated, one can regard tumoral growth as an out-of-control and never-ending wound healing process[114]. Through its broad transcriptional activity, MYC's role in cancer is essential during both tumor establishment and maintenance. Although subject to scientific debate[115–118], most of the related publications support the idea that MYC also promotes metastasis, as illustrated in a recent study published by our group, showing that MYC inhibition results in a reduction of Triple-negative breast cancer (TNBC) metastatic colonization of the lungs[119].

In a pathological context, MYC physiological transcriptional activity in cell proliferation, dedifferentiation, metabolism or cell motility is amplified and fully exploited for tumoral growth. Hence, the pathological deregulation of MYC, potentially coupled with concomitant mutations commonly found in cancer, further potentiates disease progression through its transcriptional activity.

Interestingly, although MYC clearly induces cell proliferation, it was also shown to promote apoptosis, notably mediated by p53, when highly expressed in cells[120]. In this context, apoptosis is regarded as a safeguard mechanism against this oncogene excessive expression[121]. Illustratively, multiple studies of a MYC-induced murine lymphoma model (*Eμ-myc*) showed that most tumors developed by these mice displayed either a disruption of the $p19^{arf}$-MDM2-p53 pro-apoptotic pathway or an overexpression of the BCL-2 or BCL-xL anti-apoptotic proteins[122–125].

In normal conditions, MYC activity is repressed by the Hypoxia Induced Factor 1α (HIF1α), which silences MYC mitogenic signals and improves cell resilience in hypoxic conditions. In cancer, those conditions are frequent at the center of non-vascularized and rapidly growing tumors. In contrast to their physiological relationship, though, deregulated MYC and HIF1α can also actively collaborate in tumorigenesis, where HIF1α enhances MYC-dependent angiogenesis induction, notably through VEGF activation[126,127], and modulates cancer cells metabolism transformation by favoring glycolysis and lactate fermentation (Warburg effect)[128]. Therefore, akin to some p53 mutations, HIF1α deregulation switches it from a repressor of MYC activity in physiological conditions to a potential activator and partner in cancer.

Importantly, MYC also modulates tumor immunity, primarily by promoting an immunosuppressive tumor microenvironment[129,130]. Indeed, MYC influences both innate and adaptive immunity through various mechanisms[131]. For instance, MYC supports the expression of immunosuppressive molecules, such as PD-L1 and CD47[132,133], that help tumor cells evade immune detection and phagocytosis. Moreover, through the upregulation of interleukin-8 (IL-8), IL-10, TNF-α, CCL9 and IL-23, in cooperation with common mutations in KRAS[134], MYC contributes in creating a pro-inflammatory environment, favoring the tumoral growth by inducing the recruitment of regulatory T cells (Tregs) and tumor-associated macrophages (TAMs)[131,134,135].

To summarize, the overwhelming reach of the MYC transcriptional activity gives it an important role in almost all identified cancer hallmarks[136,137] (Figure 5). Moreover, several cases of resistance to immunotherapy, notably of B-cell lymphoma[138] (but not exclusively[139]), and to chemotherapy[138,140] were attributed to MYC, making this proto-oncogene a most-wanted target, not only for the promises of its inhibition on tumor development on its own, but also because of the potential of drastically improving the effectiveness of already approved cancer therapies.

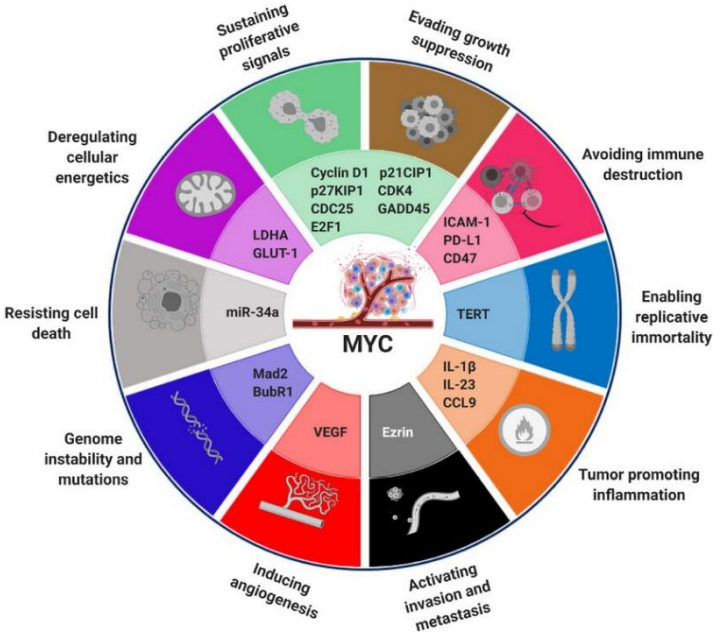


***Figure 5**, MYC is involved in virtually all the identified hallmark of cancer. Figure reproduced from Martínez-Martin and Soucek, 2021*[137].

1.5. MYC inhibition and its challenges

Over the last two decades, an astonishing variety of approaches, both direct or indirect, were explored to achieve therapeutic MYC inhibition, considered by many as one of the "Holy Grails" of cancer therapy[141–144]. The strategies employed to inhibit MYC, thoroughly reviewed in [141,144–147], include, among others, the use of small molecules[148–152], antisense oligonucleotides and siRNAs[153–156], as well as synthetic peptides or recombinant proteins[157–162]. Most of the direct MYC inhibitors developed so far aim at inducing MYC degradation, inhibiting its union to MAX and/or to E-Box DNA sequences by interacting with its b-HLH-LZ domain. Moreover, other strategies involve inhibiting the interaction between the various MYC box sequences and proteins involved in the recruitment of the transcription machinery[27] or MYCN stability[163].

Nevertheless, despite extensive efforts from the scientific community to develop therapeutic MYC inhibitors, only a handful of molecules have reached the clinical stage and no direct MYC inhibitor has ever received approval for commercialization as a therapeutic agent. Indeed, inhibiting MYC activity has proven to be challenging, to say the least. Three key factors make it a prime example of a target seemingly impossible to tackle through conventional strategies:

First of all, the MYC protein is intrinsically disordered, displaying a rather featureless surface lacking binding pockets that would be suitable for efficient and specific interaction with small molecules[144]. Nevertheless, many proof-of-concept MYC binding small molecules were tested over the years. The peptide mimetic IIA6B17 developed by Berg and colleagues in 2002 was the first small molecule direct MYC inhibitor and was described as an antagonist of MYC/MAX heterodimerization. Interestingly, this molecule turned out to also inhibit the Jun/Fos dimerization[148], suggesting that IIA6B17 might be inhibiting any similar b-HLH-LZ dimerization, to some degree.

Despite effectively disrupting MYC's activity, other small molecules such as 10074-A4, 10058-F4 and 10074-G5[164], or JKY-2-169[165] were also limited by low binding specificity, often generating undesired off-target effects and/or poor biodistribution.

Nevertheless, several independent studies conducted with the more recently developed Myci975 suggest that this compound might be the most

promising small molecule direct MYC inhibitor to date[152,166–168]. This molecule is an improved analog of the previous Myci361, which was shown to cause important off-target effects. Like its parent molecule, Myci975 was shown to bind the Helix1 of the MYC b-HLH-LZ, inhibiting MYC/MAX dimerization[152].

Interestingly, most small molecules were actually reported to bind only three distinct regions of the MYC protein, all located within the HLH domain (Figure 6), whether they were originally directed against them or not[14,144]. These sequences encompass the most hydrophobic parts of the MYC b-HLH-LZ and are located near some of the few MYC regions that display significant folding in the absence of MAX, even in the presence of chaotropic agents, as evidenced by nuclear magnetic resonance (NMR)[14]. The presence of persistent secondary structure elements organized around long, loose, and disordered stretches in the MYC protein might however not be enough for a small molecule to bind this protein with sufficient specificity. In fact, despite the existence of promising *in vivo* results for some of those molecules, to date there is no ongoing clinical trial for small molecules directly interacting with MYC, and we may be yet to discover one with the appropriate specificity and pharmacokinetics (PK) to be considered a serious clinically viable drug candidate.

The nuclear localization of MYC represents the second major obstacle to its inhibition. Indeed, developing a potent and selective MYC binder may not be sufficient if the molecule cannot reach cancer cells' nuclei appropriately. Typically, while antibodies excel in recognizing virtually any target protein with high affinity and specificity, they are not well-suited for cell penetration, let alone nuclear localization. Therefore, unless major advances in drug delivery are made, the development of an anti-MYC therapeutic antibody would not constitute a realistic path to MYC inhibition.

Finally, the MYC's central role in vertebrates' biology, illustrated by the fact that *c-myc* and *n-myc* (but not *l-myc*) knockout mice are not viable[7,77,169], suggested that its inhibition would cause severe side effects in patients. Research conducted by Dr. Soucek, though, strongly contributed to change this paradigm when she demonstrated that MYC systemic inhibition by a MYC dominant negative, called Omomyc, would cause only mild and reversible side effects in adult mice.[170]

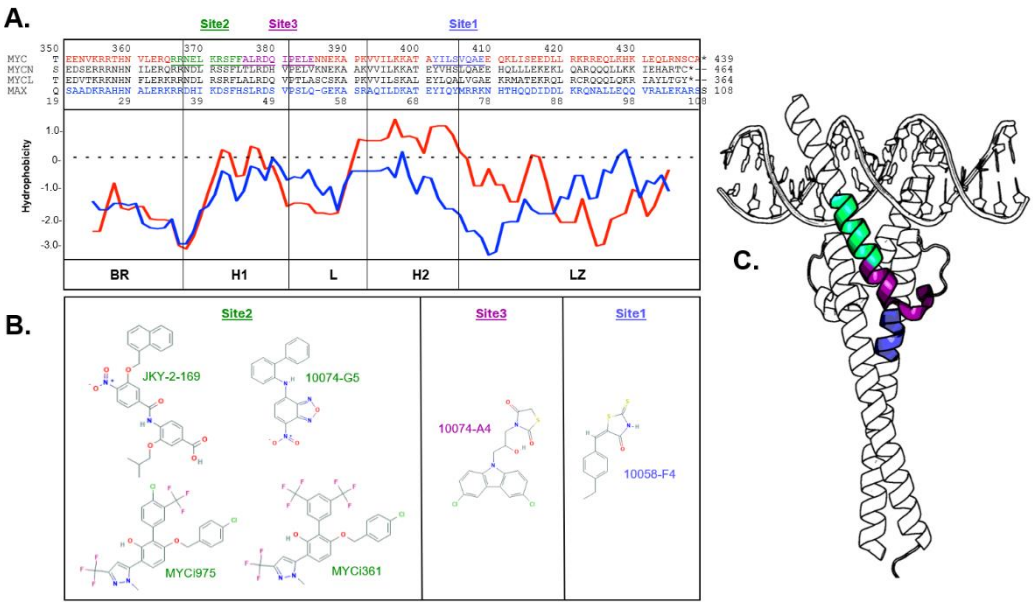


***Figure 6, A**. Top, Protein sequence alignment of MYC, MYCN, MYCL and MAX. The three binding sites in MYC's sequence are colored in blue (site 1), green (site 2) and purple (site 3). Bottom, hydrophobicity of MYC (red) and MAX (blue) respective b-HLH-LZ domain, calculated from the online tool ProtScale using the Kyte and Doolittle reference values (https://web.expasy.org/protscale/pscale/Hphob.Doolittle.html). **B**. Two-dimensional structure of small molecule MYC inhibitors grouped in function of their MYC binding site. **C**. Crystal structure of the MYC-MAX b-HLH-LZ heterodimer bound to canonical E-Box DNA. The three small molecule MYC inhibitor binding sites respective position was highlighted using the color code previously described. Figure adapted from Beaulieu et al. 2020.*[14]

2. Omomyc, a first-in-class MYC inhibitor

2.1. The use of the Omomyc transgene demonstrates the viability of anti-MYC therapies to treat cancer

Originally designed and validated in 1998 by Laura Soucek et al., Omomyc is a mini-protein of 91 amino acids corresponding to the MYC b-HLH-LZ region[171] bearing four mutations that differentiate it from its parent protein: E410T, E417I, R423Q and R424N. These mutations replace charged residues that would otherwise destabilize MYC homodimerization through electrostatic repulsion, with the corresponding residues from MAX. The E63T substitution is the only exception from this MAX homology rationale, since inserting a threonine residue at this position was found to be more favorable for complementarity than the asparagine found in MAX. These changes enable Omomyc to form stable homodimers as well as heterodimers with MYC and MAX. While, as a homodimer and a heterodimer with MAX, the translated Omomyc protein is capable of occupying MYC's DNA binding sequences efficiently, it curiously displays only marginal affinity for DNA as a heterodimer with MYC[171]. Importantly, lacking MYC's TAD, Omomyc is unable to recruit any transcription-related complexes, but instead has the ability to limit MYC transcriptional activity by sequestering it away from DNA and competing with it for MAX and E-Box binding.

Omomyc was initially extensively studied as a transgene and was shown to have anti-proliferative and pro-apoptotic effect in multiple models of cancer, including cell lines, transgenic mice and patient-derived xenograft models. Independent proofs of concept were collected in a broad range of tumor types, such as pancreatic ductal adenocarcinoma (PDAC)[172], neuroblastoma[173], glioblastoma[174], small cell lung cancer (SCLC)[175], or melanoma[176], among others. In SCLC, for example, Omomyc effectively inhibited the growth of nine cell lines with genetic inactivation of TP53 and RB1, most of which also had concurrent amplification of MYC, MYCL, or MYCN. Regardless of which MYC family member was amplified, Omomyc expression led to a reduction in cell number through induction of cell cycle arrest and/or apoptosis. The cell cycle arrest occurred in either the G1 or G2/M phases and was associated with the activation of cyclin-dependent kinase inhibitors p21 and p27, as well as a reduction in p16. Interestingly, treatment of various non-small cell lung carcinoma (NSCLC) cell lines with shRNA targeting *c-myc* replicated their respective Omomyc-induced

phenotype, proving that Omomyc therapeutic impact was based on MYC inhibition. Importantly, a cell line lacking detectable MYC levels showed the least response to Omomyc[175].

In a transgenic mouse model of KRAS-driven lung adenocarcinoma conditionally expressing Omomyc in all tissues except in the bone marrow, Omomyc caused tumor regression and eradication through decreased proliferation, and increased apoptosis and senescence[170]. In another study with the same mouse model, Omomyc expression was shown to significantly extend the animals' survival by efficiently reducing their tumor volume. New tumors arising after the end of a 4-week cycle of treatment were observed to be still as sensitive to MYC inhibition as the first ones, giving the opportunity for metronomic treatment. In fact, on and off treatment cycles in those mice reduced the number of relapsing tumors after each treatment round, until their complete eradication. Similar results were obtained in an aggressive p53-deficient tumor model, where no emergence of resistance to MYC-inhibition was observed. These results effectively reinforced the promise of MYC inhibition by Omomyc as a therapeutic approach in NSCLC[177]. They were later complemented by other studies in vivo demonstrating efficacy of MYC inhibition by Omomyc in breast cancer[178], melanoma[176] or glioblastoma, where the activation of the Omomyc transgene, for example, was shown to reduce the tumor vascularization, induce remodeling of the stroma or decrease the number of migrating cells[174,179].

Most importantly, all these models confirmed that Omomyc expression did not significantly affect body weight, blood biochemistry, or the general animal well-being. Moreover, histological examination of low proliferative tissues such as the pancreas, kidney, liver, heart, and lung did not reveal any phenotypic change[170]. Only highly proliferative tissues were shown to be mildly affected by Omomyc expression. In the testis, Omomyc induced atrophy of spermatogonia and a decrease in spermatocyte counts, and in the small intestine, villi were shortened, although this effect was not accompanied by an increase in apoptosis or any alteration in nutrient absorption or blood chemistry. Slightly more evident effects, although just transient, were observed in the bone marrow, where after 1 week of transgene induction a 50% reduction in proliferation leading to anemia and leukopenia onset were observed. Nevertheless, extramedullary hematopoiesis immediately ensued in the spleen, quickly restoring blood

counts to near normal levels after only one more week of Omomyc activation[170].

Despite all these results and the clear therapeutic potential of the Omomyc transgene, this molecule was still mainly considered an experimental tool to study MYC activity rather than a way to inhibit it in patients[180]. For context, at the time of most of the previously mentioned discoveries, no gene therapy had been ever approved, the first one being KYMRIAH® (tisagenlecleucel), which received the United States of America Food and Drug Administration (FDA) approval for commercialization in 2017 to treat acute lymphoblastic leukemia in pediatric and young adult patients[181].

2.2. The MAX b-HLH-LZ displays intrinsic cell penetration properties

An important breakthrough happened in 2012, when Montagne and colleagues demonstrated that the b-HLH-LZ domain of MAX displayed intrinsic cell penetrating properties[159]. This study was inspired by the identification of a putative nuclear localization signal in the MAX basic region sequence, and by an early study published by Noguchi and colleagues in 2005, which demonstrated that the b-HLH BETA2/NeuroD protein could efficiently translocate into cells[182]. Montagne and colleagues demonstrated that, upon treatment with micromolar concentrations of the MAX-derived mini-protein Max*, Hela cells proliferation was inhibited, as a consequence of cell cycle arrest in G0/G1 phase and an increase in apoptosis.

The cell penetration ability of Max* was observed to be inhibited at 4 ºC, suggesting that its translocation into cells happened through ATP-dependent mechanisms. Moreover, in the same cells, the expression of CyclinD1, upregulated by MYC, was repressed upon 72h of Max* treatment, while the expression of p21, downregulated by MYC, was restored in a dose-dependent manner[159]. Even though these studies were limited to an *in vitro* setting, and despite the claim within them that the MYC b-HLH-LZ domain could not penetrate cells spontaneously, the strong homology between Omomyc and MAX constructs motivated our group to test the cell penetrating ability of Omomyc as a purified mini-protein.

2.3. The Omomyc mini-protein is a viable anti-cancer drug candidate

As a consequence, in 2019, our group published for the first time the antitumoral activity of Omomyc as a purified mini-protein displaying cell-penetrating ability[160], which led to the foundation of Peptomyc S.L. in 2014 by Dr. Soucek and Dr. Beaulieu, the two co-directors of this doctoral thesis. In this study, Omomyc was successfully overexpressed in *E. coli* and purified by cation exchange chromatography (CEX). Similarly to what had been described for the BETA2/NeuroD and Max* b-HLH-LZ proteins, Omomyc displayed intrinsic cell penetration properties in a panel of NSCLC and various other cancer cell lines. In fact, an Alexa Fluor® 488 (AF488)-labelled Omomyc penetrated cells in a dose-dependent manner after a few minutes of exposure to low micromolar concentrations (Figure 7). The mini-protein was found in both endosomes and cell nuclei. Similarly to Max*, Omomyc cell penetration was inhibited when the incubation was carried out at 4 ºC. Comparing Omomyc cell penetration properties in the presence of various endocytosis inhibitors revealed that the mini-protein does not rely on a single entrance mechanism, but enters cells through multiple pathways, including clathrin-dependent endocytosis, macropinocytosis, and, to a lesser extent, caveolin-dependent entrance. Interestingly, the contribution of the different endocytosis pathways varied from one NSCLC cell line to another.

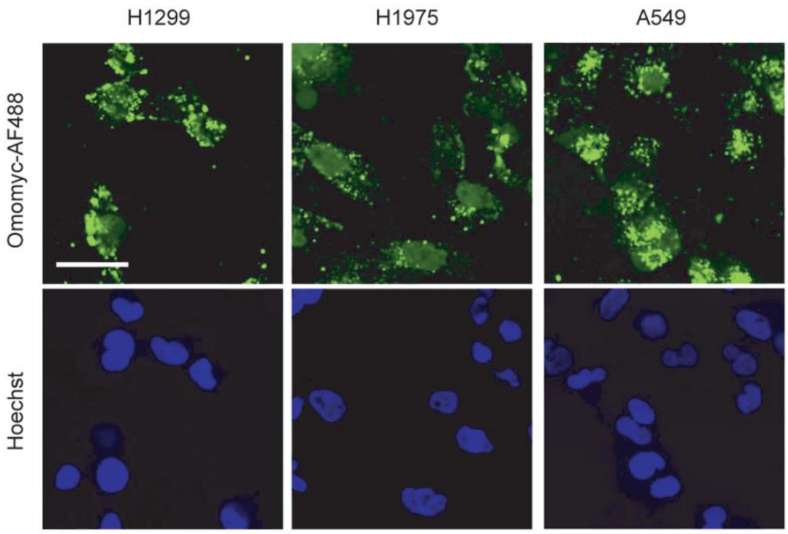


***Figure 7**, Omomyc conjugated to AF488 displayed intrinsic cell penetration properties, as observed in three NSCLC cell lines: H1299, H1975 and A549. Figure adapted from Beaulieu et al. 2019.*[160]

Furthermore, Omomyc was shown to lose its ability to penetrate cells when the arginines of its basic region were substituted with alanine residues, confirming their critical role and charge during the transduction processes.

The Omomyc mini-protein recapitulated all the main features of its transgenic counterpart:

Biophysical characterization of the Omomyc mini-protein by circular dichroism (CD) and nuclear magnetic resonance (NMR) confirmed its ability to bind MYC and MAX b-HLH-LZ regions. In contrast with MYC b-HLH-LZ sequence, Omomyc displays a high α-helix content which gets further stabilized upon DNA binding. The Omomyc crystal structure, solved and published by Jung and colleagues, further confirms the dimeric nature of Omomyc, as well as its high helicity both in the apo form and bound to DNA (accessible on the Protein Data Bank (PDB) with the codes 5I4Z and 5I50, respectively)[172]. Moreover, electro-mobility shift assay (EMSA) demonstrated high affinity of the Omomyc mini-protein for the canonical E-Box DNA sequence (CACGTG). Interestingly, when the assay was performed at 12 ºC, Omomyc homodimers displayed a higher affinity for DNA than Omomyc/MAX dimers, whereas the opposite was observed at physiological temperature. In contrast, Omomyc/MYC heterodimers only displayed intermediate DNA affinity at 12 ºC and were not seen to significantly shift DNA's migration at 37 ºC. As a whole, this set of experiments confirmed the mode of action of Omomyc hypothesized when studied as a transgene (Figure 8).

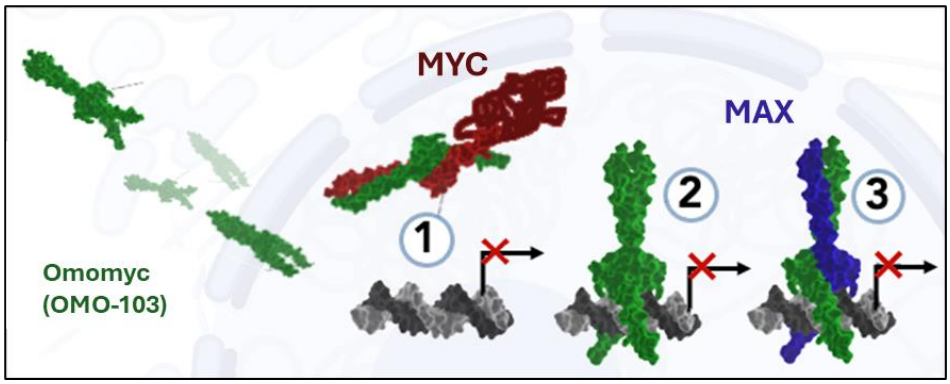


***Figure 8**. The Omomyc mini-protein inhibits MYC transcriptional activity by preventing it from binding to its target DNA sequences through three modes of action: (1) it binds the MYC protein and the resulting heterodimer displays very poor DNA binding, (2) the Omomyc mini-protein homodimerize and efficiently compete with the MYC/MAX dimers for E-box sequences, (3) through*

dimerization with MAX, it both depletes MAX amount available for dimerization with MYC and competes with MYC/MAX dimers for the MYC target DNA binding sites.

Treatment of NSCLC cells with micromolar concentrations of the Omomyc mini-protein led to a significant reduction in cell proliferation, mirroring the effects seen with the Omomyc transgene. In H1299, A549, and H1975 cells, Omomyc induced a dose-dependent decrease in cell number, with IC50 values in the low micromolar range, and caused cell cycle arrest predominantly in the G0/G1 phase (similar effects were also confirmed recently in breast cancer cell lines[119]).

Through microarray, Chromatin Immunoprecipitation-quantitative Polymerase Chain Reaction (ChIP-qPCR), and ChIP-seq analyses, the Omomyc mini-protein was confirmed to selectively inhibit MYC-driven transcriptional activity and downregulate gene signatures associated with poor prognosis in lung cancer. Importantly, the transcriptional programs of other b-HLH-LZ transcription factors were unaffected.

Omomyc exerted its effects by displacing MYC from its promoter sites, altering H3K27 acetylation, and reducing MYC's overall genomic occupancy, particularly at superenhancers. Remarkably, the study showed that MYC was at least partially displaced from 97.3% of its active promoter regions. Gene Set Enrichment Analysis (GSEA) further confirmed that Omomyc effectively disrupted MYC-regulated gene expression in NSCLC cells. As prominent examples, Omomyc downregulated the nucleolin (Ncl) and nucleophosmin 1 (Npm1) genes, which are typically upregulated by MYC and play roles in angiogenesis and cell division, respectively[160]. Furthermore, in breast cancer cells, a Proximity Ligation Assay (PLA) confirmed Omomyc's ability to heterodimerize with MYC and MAX, reducing the number of MYC/MAX dimers, consistently with previous biophysical findings[119].

An additional study was conducted independently from ours by Merck Sharp & Dohme (MSD) to further study Omomyc therapeutic potential[183]. In this study, Demma and colleagues could recapitulate most of our results and revealed that chemically synthesized Omomyc displayed a very similar activity than the recombinantly expressed counterpart. Interestingly, their PLA results revealed that both Omomyc homodimers and Omomyc/MAX heterodimers were detected in the nuclei, whereas Omomyc/MYC dimers were exclusively observed in the cytoplasm. By E-Box DNA pull down coupled to mass spectrometry (MS), they noticed that Omomyc was not

only competing for MYC binding sites with MYC/MAX dimers but also with other members of the PMN, such as MGA, MXI1 and MXD3. In a MYC, MAX and biotinylated Omomyc co-immunoprecipitation assay carried out in HCT116 lysate spiked with 10 µM of biotin-Omomyc, Omomyc was seen to co-precipitate with MYC but not MAX. The authors then hypothesized that MYC/MAX dimerization may occur co-translationally and that MAX would become unavailable. To further challenge this hypothesis, they performed a modified translating ribosome affinity purification as well as a ribonucleic acid (RNA) immunoprecipitation. In those assays, Omomyc treatment was found to interfere with the ability of MYC but not MAX, to bind MAX-coding RNA and the ribosomal machinery, observed in untreated cells. Interestingly, the union of MYC, MAX and Omomyc to MAX-coding RNA was lost when the ribosomes were disrupted with a puromycin treatment, or the addition of ethylenediaminetetraacetic acid (EDTA) after cell lysis. Altogether, these data provide convincing indications that MYC and Omomyc binding to MAX could be mediated by the union with intact ribosomes[183]. In the same study, the Omomyc treatments of HCT116 cells were observed to decrease cellular MYC levels, likely by inhibiting its interaction with MAX, thus promoting MYC's proteolytic degradation[183].

Work by our lab also showed that Omomyc intra-nasal[160] and intra-venous[119,160] administrations in mice were well tolerated by the animals and no significant loss of weight was ever observed upon Omomyc treatment. The intracellular presence of Omomyc in lung tissues was confirmed by immunofluorescence 4 hours after intra-nasal administration, and ^{89}Zr-radiolabelled Omomyc could be observed in tumor tissues for at least 48h. In NSCLC and TNBC cell line-derived xenograft (CDX) models, tumors growth was inhibited upon Omomyc administration, which conferred a survival advantage to the treated mice. Moreover, in highly metastatic TNBC models, Omomyc was seen to interfere with metastatic lung colonization and reduce the total number of tumors per mouse, suggesting that MYC plays an active role in metastasis establishment.

In 2023, our group published the results obtained during an Omomyc PK study in tumor-baring BALB/c-nude mice, using a liquid chromatography-tandem mass spectrometry-based (LC-MS/MS) method[184]. In contrast to the methods employed in other PK studies previously published[160,183], this technique does not rely on the use of fluorescently or radioactively labelled Omomyc , whose results may be impacted by the intrinsic labels' properties

that can affect the stability or biodistribution of the fused protein, and/or by the fact that the monitored signal may not correlate with the integrity of the studied protein. In this recent study, the presence of full length Omomyc in tumor tissues and serum could be asserted for the first time by determining the concentration of four Omomyc-corresponding proteolytic peptides encompassing the full-length protein. Two of them contain at least one of the 4 mutations differentiating Omomyc from the MYC's sequence. The last two correspond instead to a part of the basic region and loop common to both Omomyc and MYC. At every time point, the relative concentration of the 4 peptides were consistent among themselves.

In serum, the Omomyc concentration declined in two phases. First, a rapid decrease in concentration was observed, and about 5% of the amount detected at 5 min after injection was measured after 2 hours. After the 24 hours timepoint, the Omomyc concentration slowly decreases in the days following the injection; and after 72h, the last timepoint of the study, about 0.4% of the Omomyc concentration detected at 5 min was still present.

In contrast, in tumor tissues, significant Omomyc concentration was detected as early as 5 minutes post-injection, with little variability among samples. This observation indicates a quick and homogenous accumulation of the mini-protein within the tumor tissues, which persists for at least 2 hours following the injection. Interestingly, the subsequent decrease of Omomyc in tumors was much slower than the one observed in serum. Notably, from the 24-hours timepoint on, the mini-protein concentrations detected in tumor tissues become significantly higher than in serum, highlighting a long-lasting stability. Interestingly, this method did not allow us to detect any significant concentrations of MYC-derived peptides[184], which suggests that Omomyc concentrations may have also been underestimated due to an incomplete protein extraction and/or digestion during sample preparation.

In May 2021, the Omomyc mini-protein (a.k.a. OMO-103 for its clinical name) became the first direct mini-protein MYC inhibitor to reach the clinic in a phase I clinical trial in cancer patients with advanced solid tumors[185]. The results of this study conducted in a total of 22 patients demonstrated the excellent safety profile of Omomyc in human. OMO-103 was indeed well-tolerated, with most patients experiencing only grade 1 side effects, primarily mild infusion-related reactions commonly associated with biologic treatments. In agreement with previous studies, the mini-protein

also exhibited favorable pharmacokinetic properties up to the recommended dose, and little to no immunogenicity was detected even after long term treatment[185]. Moreover, although no objective response was obtained by RECIST criteria, the study revealed promising signs of OMO-103's effectiveness, with evidence of target engagement and biomarker response. Indeed, half of the patients, all with advanced cancers resistant to all previous lines of therapies, showed disease stabilization. Among those, there were two important notable cases: an advanced PDAC patient, who showed a 49% total tumor burden reduction after a first 2-month treatment cycle, and another patient with salivary adenocarcinoma, whose disease stayed stable for about 26 consecutive months. Importantly, this study was the first to demonstrate the suppression of MYC-driven transcriptional signatures in patient biopsies, underscoring the drug's activity and target engagement[185].

At the time of this thesis writing, OMO-103 is being tested in a new clinical trial phase Ib (Clinical trial ID NCT06059001) in PDAC patients as first line treatment in combination with standard of care chemotherapy.

From a drug development perspective, the publication of the OMO-103 phase I clinical trial results also indirectly demonstrated that a viable Good Manufacturing Practices (GMP)-compliant production process could be developed at industrial scale, and that this mini-protein displays sufficient solubility and stability for long-term storage and subsequent intravenous injections in patients.

As a whole, the thorough characterization of Omomyc so far has illustrated the potential of the b-HLH-LZ proteins for the inhibition of hard-to-reach transcription factors and demonstrated a capacity to bind specific DNA sequences and protein targets with an affinity in the nM range, rarely seen for short peptides or small molecules.

3. The emergence of the b-HLH format as a new class of therapeutic protein

3.1. Omomyc, a benchmark to surpass?

Before the publication of Montagne and colleagues demonstrating the intrinsic cell penetration ability of the MAX b-HLH-LZ[159], these protein domains were expressed recombinantly for *in vitro* mechanistic studies aiming to gain insight into their complex structures and biology, but never to study their potential therapeutic activity. However, the attention generated by the characterization of Omomyc and its MYC-inhibiting effects over the years undoubtedly raised the interest of the scientific community for b-HLH-LZ proteins.

For instance, Omomyc inhibitory effect over cancer cell growth was compared to the one displayed by ME47 (or Max-E47), a b-HLH construct shown to tightly interact with E-box sequences, displacing MYC/MAX dimers[186,187]. ME47 is a fusion protein displaying the MAX basic region and the HLH domain of the E47 transcription factor that enables it to form homodimers, ME47 was found to bind the canonical E-box with high affinity and specificity in a fluorescence polarization assay[186]. As a transgene and fused to a NLS sequence at its N-terminus, ME47 inhibited cell proliferation *in vitro* and tumor growth in a TNBC CDX model[187]. Although the authors reported a comparable effect on cell proliferation compared to Omomyc, it is not clear whether this was directly related to MYC downregulation. This type of analysis would have been of great value since ME47 only recapitulates one of Omomyc's three modes of action (occupancy of DNA). A ME47-derived variant with improved DNA binding and structural stability, named MEF, was described in a subsequent study published in 2021. Interestingly, in this work, Inamoto and colleagues achieved the optimization of ME47 through phage-assisted continuous evolution[188] and fusion of the resulting b-HLH to the FosW LZ domain, resulting in what the authors appropriately defined as a "franken-protein"[189]. Unfortunately, no efficacy data was shown in this publication, and it is therefore unclear whether the MEF enhanced DNA binding ability would correlate with improved cell proliferation inhibition. Hopefully, such data will be published in the future.

In a more complete publication, Wang et al. suggested that Omomyc activity could benefit from an improvement of its cell penetration and endosomal escape properties, which could be achieved by its fusion to a functional penetrating Phylomer sequence (FPP) at Omomyc's N-terminus. The resulting molecule, named FPPa-Omomyc, probably obtained through the ligation of recombinant Omomyc and synthetic FPPa moieties, induced convincing tumor growth inhibition and significant survival advantage in a murine T11 cell-based TNBC allograft model. Moreover, MYC targets inhibition was confirmed by histo-immunochemistry staining of the allografts tissues in FPPa-Omomyc-treated mice[190]. This research program, partially sponsored by the start-up company Phylogica (now renamed PYC Therapeutics), seems however to have been discontinued, judging by the information available in the PYC Therapeutics website, which appears now focused on RNA delivery[191].

Through a series of three publications originating from the same team[183,192,193], the pharmaceutical giant MSD also demonstrated its interest in Omomyc and MYC inhibition via b-HLH-LZ proteins. After a detailed characterization of Omomyc[183], Demma et al. studied the capacity of a MXD1-derived mini-protein, named Mad, to inhibit cancer cell proliferation by antagonizing MYC's activity[192]. Mad's sequence comprises the first 146 amino acids of MXD1, including its N-terminal mSIN3a binding domain and b-HLH-LZ. In this construct, the authors mutated the two MXD1 Cys residues to Ser to avoid any undesired disulfide bridge formation, and substituted the Ser145 with an Ala to improve protein stability, since Ser145 is a known MXD1 phosphodegron. Also, Mad displayed intrinsic cell penetration properties, and colocalized with Hoechst and histone 3 staining in HCT116 cells. In Ramos cells, Mad displayed superior growth inhibition when compared to Omomyc, as well as the capacity to efficiently bind MXD1 and UBF (a MXD1 interactor and repressor of rRNA transcription), by co-immunoprecipitation. Finally, Mad's ability to downregulate some of MYC target genes was demonstrated by ChiP-qPCR[192]. Surprisingly, despite showing very promising data *in vitro*, this study did not include any results obtained with *in vivo* models, and no follow up study was published yet. Although incomplete, Demma et al. demonstrated the intrinsic cell penetration properties of the MXD1 b-HLH-LZ, despite the fact that Mad is a significantly larger construct than Omomyc or Max*.

In a third publication, Brown and colleagues detailed the synthetic routes for Omomyc synthesis through solid-phase peptide synthesis (SPPS) and native chemical ligation strategies. Moreover, the authors explored the effects of LZ truncation on the Omomyc activity and performed a FRET-based high-throughput (HTP) mutagenesis study to identify sequences with optimized DNA binding affinity and specificity. The best variants identified through the HTP study mainly displayed point mutations in Omomyc's BR. Surprisingly, the Omomyc impact on cell proliferation was recapitulated and even seemingly improved by a LZ-truncated Omomyc variant whose sequence corresponds essentially to the MYC's b-HLH (Δ60-90, Q58C). Altogether the results of this publication suggest that the Omomyc inhibitory effect is mediated mainly through competitive binding to E-box sequences as a homodimer[193]. The authors' claims rely, however, on the sole observation of *in vitro* cell growth inhibition in two cell lines and should therefore be considered with caution. In fact, the improved variant's stability and *in vivo* PK characterization data promised by the authors in this 2020 publication are yet to be accessible, 4 years later.

In 2021, two novel studies conducted by the Massachusetts Institute of Technology group of B. Pentelute and partially financed by Bristol Myers Squibb™, presented some results obtained with synthetic Omomyc and MAX b-HLH-LZ homo- and heterodimers. Interestingly, this group is known for having developed a performant automated SPPS system enabling the rapid and efficient synthesis of large peptides and mini-proteins, a reference in the field of SPPS[194]. The work presented in Jbara et al., 2021[195] and Pomplun et al., 2021[196] is focused on the stabilization of b-HLH-LZ dimers through the substitution of the labile, spontaneously-forming disulfide bridge near the C-terminus of Omomyc and Max* by a synthetic permanent linker obtained through S-arylation[195] or parallel synthesis[196] approaches. In accordance with previous studies, both synthetic Omomyc and Max* homodimers could penetrate HeLa cells in a dose-dependent manner, partially colocalizing with the nucleus and inhibiting their proliferation *in vitro*. Also, the Omomyc oligomeric state was not seen to significantly impact its cell penetration properties[196]. In addition, they ligated the Omomyc dimer to a NLS peptide called Mach3, successfully improving its nuclear localization but not its overall cellular uptake. Nevertheless, the Omomyc-Omomyc-Mach3 construct displayed more potent cell proliferation inhibition in 3 cell lines, compared to the Omomyc and Max* homodimers[196]. Unfortunately, again no *in vivo* data are shown in those studies which would have been very

valuable to estimate the impact of the dimer stabilization over its pharmacokinetic properties and tissue distribution.

Recently, using an innovative chemical approach, Speltz and colleagues simplified the MAX b-HLH domain into a peptide composed of two ligated stapled helices, one corresponding to the MAX basic region / Helix 1 (termed basic helix) and the other to its Helix 2 (called zipper helix), and got rid of MAX's loop by replacing it with a more rigid thiol-maleimide ligation. The resulting constructs, called synthetic transcriptional repressors (or STR) could homodimerize, mimicking the structure of a natural b-HLH domain. Several variants of both moieties were independently associated to form a library of approximately a hundred molecules. Screening the library for optimal structural stability, resistance to tryptic degradation, DNA affinity and capacity to displace MYC/MAX dimers from DNA, the authors could identify two hits with significantly improved binding and stability. Both STRs demonstrated efficient transduction into cancer cells and inhibited proliferation in a MYC-level dependent manner. Perhaps more importantly, Speltz and colleagues showed that the modular design of these STRs also allows for the potential adaptation of these molecules to bind DNA sequences other than E-boxes, therefore targeting other transcription factors beyond MYC, offering a versatile platform for gene regulation[197]. This approach illustrates greatly the flexibility offered by synthetic peptides, both in terms of sequences and strategies. Synthetic peptide approaches may however suffer from technical challenges in later stages of development, when high amounts of material are required[198]. Nevertheless, those approaches have proven to become major contributors to the emergence of b-HLH and b-HLH-LZ as a new class of therapeutic molecules.

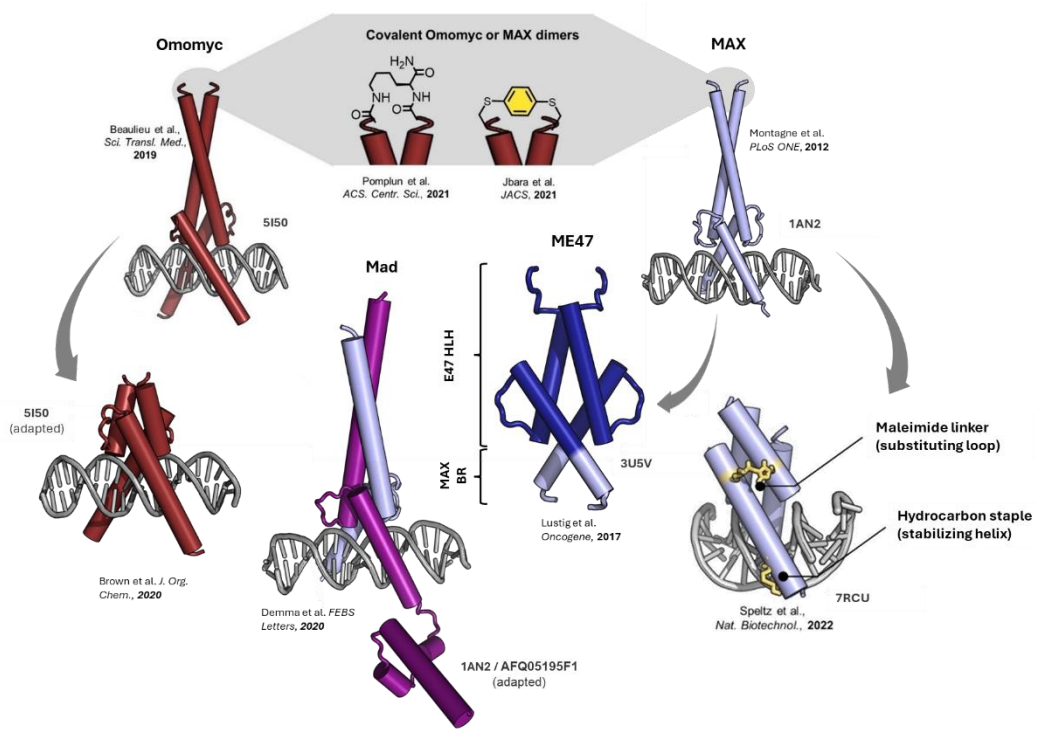


***Figure 9**, Illustrative summary of the strategies employed to inhibit MYC through b-HLH / b-HLH-LZ proteins. The Mad mini-protein representation was obtained by truncation of the full MXD1 structure predicted by AlphaFold2 (PDB entry: AFQ05195F1), and its alignment on one of MAX b-HLH-LZ monomer bound to DNA (PDB entry 1AN2). All the representations were generated using PyMOL. Figure adapted from Ellenbroek et al., Angew. Chem. Int. Ed., 2024.*[199]

3.2. The lessons learned from decades of characterization and discoveries

Decades of research and characterization of b-HLH-LZ proteins have generated an important quantity of information, but what is truly of interest for drug design and development processes?

As a start, some research groups have taught us the importance of studying the structural plasticity proper to this class of proteins. For instance, through multiple publications[11,12,15,19,200], the group of P. Lavigne demonstrated that the conditional folding of the MYC and MAX b-HLH-LZ was directly related to their oligomeric state, since their LZ domain would only fold in the presence of a partner for dimerization, and that the folding of

their BR is conditioned upon DNA binding. Of note, Lavigne's group made extensive use of CD experiments to demonstrate those points. CD methods are based on the proteins property to differentially absorb right-handed and left-handed circularly polarized light depending on their secondary structure and it is particularly indicated to evaluate a protein's α-helix content, which makes it a valuable tool to study b-HLH-LZ constructs. In fact, this technique not only can provide structural information at a given time in given conditions, but it can also be used to study the interactions of the b-HLH-LZ proteins with their partners and DNA by following the increase of their α-helix content, as well as their structural thermal stability. The fitting of CD data to thermodynamics equations can also allow the determination of a protein's thermodynamic features, using appropriate models. To that end, the same group developed a Microsoft Excel®-based workbook enabling the meaningful data fitting of CD thermal denaturation experiments of MYC and MAX-b-HLH-LZ derived proteins[201]. Moreover, CD experiments are relatively easy to set up and only require little material, which makes them compatible with early phases of molecular characterization. In contrast, NMR methods or X-ray crystallography, even though they provide meaningful structural data, require more material in a more concentrated form, as well as particular expertise for experimental setup and data analysis.

It is also clear that DNA binding is one of the most, if not the most, critical features of a b-HLH / b-HLH-LZ drug candidate, since its main mode of action often consists in occupying the binding sites of its target transcription factor(s). As reviewed previously, this feature was logically systematically studied and several strategies employed to improve Omomyc's efficacy consisted in stabilizing its affinity for DNA, directly through optimization of its BR, as in Brown et al.[193], or indirectly through stabilization of the Omomyc homodimer[195,196]. To characterize DNA binding, researchers have typically employed EMSAs or fluorescence polarization-based (FP) methods. Both techniques employ fluorescently-labelled DNA probes of short size. In an EMSA, formation of protein-DNA complexes is visualized by the difference of migration distances between the bands corresponding to free DNA and larger protein-DNA complexes through a native gel. In FP-based methods, the complex size increase caused by the protein union to DNA induces the slower tumbling of the latter (and of its fluorescent label, by extension), increasing the polarization levels of the fluorescence re-emitted after being excited with linearly polarized light (phenomenon called linear anisotropy). Since FP can be read by some plate readers, FP methods can be adapted to

384- or 1536-well plate formats and may therefore be the most indicated method for early quantitative characterization of the b-HLH-LZ-DNA interactions, if one can access such readers. Alternatively, surface plasmon resonance (SPR) has been used to study MYC and MAX interactions to DNA[24,202] and can provide the meaningful k_{on} and k_{off} binding kinetic parameters. However, this technique requires immobilizing at least one of the complex members, which may complicate the experimental setup and can limit the *in vivo* relevance of the findings.

Furthermore, intrinsic cell penetration appears to be a key and apparently common feature to all the b-HLH-LZ constructs displaying the MYC, MAX or MXD1 BR. Strategies that have improved Omomyc's ability to penetrate cells and reach their nucleus seemed to also correlate to a higher efficacy both *in vitro*[190,193,196] and *in vivo*[190]. Historically, the most common ways to determine the cell transduction capacity of a molecule relied on microscopy (detection of a fluorescently-labeled protein, immunofluorescence or immunohistochemistry) or flow cytometry. Those methods are limited by their reliance on the labeling of the molecule of interest, carried out either through chemical ligation of a fluorophore or recognition by antibodies. Both detection methods can therefore represent significant pitfalls in early drug development when resources are scarce and the labeling of a large panel of molecules may be particularly resource-demanding, especially in studies relying on recombinant expression. More importantly, the apparent intracellular localization of a molecule does not ensure its biological activity, since the protein can still be degraded or entrapped in endosomes. For those reasons, functional cell-based reporting assays may represent another meaningful approach to quantify cell penetration, although they might require significant characterization[203,204].

3.3. Paths and tools to better understand and optimize b-HLH-LZ drug candidates

The flexible and cooperative nature of b-HLH-LZ proteins, their complex mode of action, as well as their recent consideration as potential drug candidates are all factors contributing to our lack of full understanding of this new class of therapeutic proteins with potential in many key areas of drug development. Yet, paths and tools to explore aspects like cell penetration and PK can be found in the work carried out on other well-

established classes of therapeutic proteins. This last introductive section will be dedicated to giving an overview of the tools and concepts that served as a base for the work presented in this doctoral thesis.

3.3.1. The use of purification tags

Purification tags are powerful tools for the purification and detection of recombinant proteins and their number and diversity is ever-growing since the first use of the Staphylococcal protein A as a fusion partner for purification recombinant proteins in Uhlén et al. in 1983[205–207].

Among these tags, the HisTag is, without a doubt, the most commonly used purification tag. Its sequence typically consists of a 6 to 12 histidine tandem repetition and has strong divalent metallic ions chelating properties. This feature is exploited in HisTag affinity chromatography, where HisTagged proteins are captured through the interaction of their histidines residues with Nickel ions (Ni^{2+}) or less commonly Cobalt (Co^{2+}), Copper (Cu^{2+}) or Zinc ions (Zn^{2+}) immobilized on the stationary phase. The HisTag presents important advantages compared to other purification tags. First, it is relatively short, therefore its fusion to the Omomyc or the Max* sequences would not result in significantly larger constructs. It was also described as having a rather neutral effect on protein structure and solubility. Moreover, since the HisTag first description in 1988 by Hochuli and colleagues[208], a large selection of commercially available tools and reagents for purification and detection were developed. In fact, the two main reagents for HisTag affinity chromatography, the resin and pure imidazole for elution, are rather inexpensive, making them especially convenient for preparative-scale purifications of a large variant set.

Another prominent example of affinity tag is the Streptag-II (hereafter mentioned as StrepTag). In contrast to the monotonous composition of the HisTag, the StrepTag has a more equilibrated sequence and is composed of 8 amino acids: WSHPQFEK. This tag's sequence was obtained after optimization of a natural peptide identified in *Streptomyces avidinii* displaying good affinity for the biotin-binding pocket of the streptavidin protein[209]. A mutant of streptavidin called Streptactin was developed to bind StrepTag-II with even better affinity[210]. Its most optimized version, the Streptactin XT, displays nM affinity to this tag[211]. Immobilized on chromatography resin, Streptactin XT enables the efficient purification of StrepTag fusion proteins. The Streptactin XT-StrepTag complexes can be

disrupted by the introduction of competitive biotin. Compared to its parental protein streptavidin, Streptactin XT has the advantage of binding biotin in a reversible manner, allowing a possible regeneration of the resin, after elution of the protein of interest. Moreover, the Streptag has the advantage of including a Trp residue, which facilitates the quantification of StrepTag fusion protein by absorbance at 280 nm.

For their small size and extensive characterization, the HisTag and the StrepTag have the potential to be valuable tools for rapid purification and easy detection of a panel of protein, allowing for direct comparison.

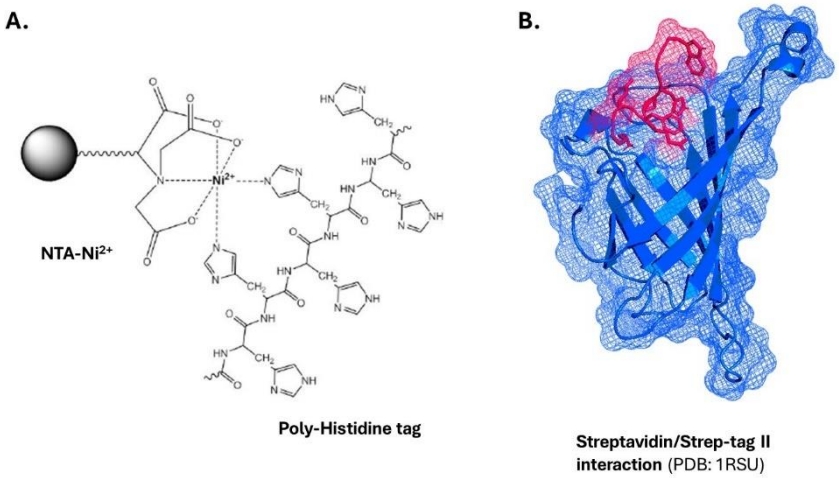


***Figure 10, A**. Interaction of NTA resin functionalized with Ni^{2+} interacting with a poly-histidine tag. Figure adapted from Magdeldin et al., 2012.*[212] ***B**. X-ray crystal structure of interacting Streptavidin (blue) with StrepTag-II (red) (PDB: 1RSU, for original publication, refer to Schmidt et al., 1996.)*[209]

3.3.2. The use of cell penetrating peptide and NLS as tags

As previously discussed, it was suggested that Omomyc may benefit from further improvement of its intrinsic cell penetration properties. A large body of literature has shown that this could be achieved through the fusion of the molecule of interest to an identified cell penetrating peptide (CPP)[213–217].

SynB3, Penetratin, and R9 are three short CPPs extensively described in the literature for their intrinsic cell penetration properties. Like many other well-described CPPs, they all share the common feature to have a sequence enriched in basic amino acids[218,219]. SynB3 is a 10-amino acid-long truncated version of its parent peptide SynB1 originating from the antimicrobial peptide protegrin 1[220–222]. Like the other members of the SynB peptide family,

SynB3 was described as a vector for both intracellular and brain delivery. In the early 2000's, Jamal Temsamami and his team published a series of articles demonstrating that the SynB CPPs could be used as carrier for small molecules, such as doxorubicin, to cross the blood brain barrier (BBB)[223–225]. While conjugated to the endomorphin-1 peptide (EM-1) via a labile disulfide linkage, SynB3 enabled BBB passage and intracellular delivery of an EM-1 active form, after physiological reduction of the disulfide[226]. Another recent proof of concept was published describing the fusion of SynB3 to the polyglutamine-binding peptide 1 (or QBP1). The fusion protein was shown to be efficiently delivered across an artificial BBB model into Huntington protein-expressing cells, where it prevented poly-Q induced protein aggregation[227].

Penetratin, also called pAntp-(43–58), was one of the first CPPs identified as such[228], little after the TAT peptide[229]. It is a 16 amino acid peptide and was isolated from the homeodomain of a *Drosophila* Antennapedia transcription factor. Several studies demonstrated Penetratin's ability to carry a broad spectrum of molecules into human cells cytoplasm, including siRNA and PNAs[230], or proteins such as the Green Fluorescent Protein (GFP)[231], Bovine Serum Albumin (BSA)[232] or INF-γ[233]. Like SynB3, the Penetratin CPP was shown to efficiently carry doxorubicin across the BBB[223].

The use of polyarginine peptides, such as R9, as cell penetrating peptides started to be formally studied in the year 2000, with a publication of Mitchell and colleagues, from Stanford University[234]. Building on research from the previous decade that saw the identification of various CPP, the authors hypothesized that the high cationic amino acid content of the CPP discovered so far was the molecular base for their cell penetration abilities. Comparison of polyhistidine (H9), polyarginine (R9) and polylysine (K9) showed that the polyarginine peptides had indeed intrinsic cell penetration properties, far superior that the ones displayed by H9 or K9[234]. Since then, polyarginine tags have been extensively studied as CPPs and conjugated to different molecules. For instance, an early study demonstrated the capacity of the R9-RNase A fusion protein to cross the human cells membrane[235]. In another study, R9 peptide decoration of nanoparticle enhanced drug accumulation in tumors and permeation in metastatic lymph nodes[236]. Interestingly, the co-administration of polyarginine peptides with insulin was shown to increase its intestinal absorption in rats[237]. Of note, negatively

charged PNAs could also be delivered intracellularly when conjugated to an R9 tag[238].

3.3.3. Net charge modulation can affect pharmacokinetics properties

A growing body of research in the field of monoclonal antibodies (mAb) suggest that a modulation of their net charge at physiological pH could significantly influences their PK, affecting factors such as systemic exposure, clearance, and tissue distribution. By adjusting the isoelectric point (pI) and net charge in the complementarity-determining regions (CDRs), it was observed that increasing positive charges led to higher tissue retention, particularly in well-perfused organs, and faster plasma clearance. Moreover, a bell-shaped relationship was identified between charge and clearance, with extreme charges (both negative and positive) resulting in rapid clearance[239]. This phenomenon was also shown in another study when selecting different antibody frameworks. Variations in the pI of therapeutic antibodies showed that higher pI values, which results from a higher content in basic residues, led to faster clearance due to increased catabolism in the liver and spleen[240–242].

Interestingly, another study used heparin chromatography, which is analogous to glycosaminoglycans found in tissues, as an in vitro predictor for antibody clearance rates through pinocytosis. Indeed, the authors could establish a correlation between heparin binding affinity and mAb clearance: the antibodies with higher positive charges showed stronger binding to heparin but also faster clearance from plasma[243].

II. Hypotheses and Objectives

The results of this thesis are divided into two different sections. In addition to partially addressing overarching questions, each section was conceived to pursue specific hypotheses and objectives.

General considerations:

Motivated by the therapeutic promises represented by MYC inhibition and following the Omomyc example, several proof-of-concept b-HLH-/b-HLH-LZ-based MYC inhibitors were described over the last decade. Yet, the most efficient path for b-HLH-LZ drug candidate characterization and development remains unclear. Thus, the development of a systematic approach to efficiently evaluate the potential of a b-HLH-LZ-based MYC inhibitor drug candidate could support and enable the discovery of a second generation of MYC inhibitors after Omomyc.

Hypothesis:

We hypothesize that the fusion of purification tags to candidate sequences will enable a straightforward purification and direct comparable detection of multiple constructs. In addition, we predict that a purification tag's impact on a construct biological and biochemical features will be independent of the b-HLH-LZ protein sequence.

Objectives:

- Evaluating the viability of the use of the HisTag and the StrepTag to streamline the purification and characterization of Omomyc-derived variants.
- Studying the impact of the HisTag and the StrepTag on Omomyc-derived constructs' biological and biochemical features.
- Exploring the use of relevant assays for the assessment of a peptidic direct MYC inhibitor key biochemical and biological features.

Section 1: Modulation of the cell penetration activity and nuclear localization of b-HLH-LZ proteins using short-stretch mutations and cell penetration tags.

Several publications suggested that Omomyc therapeutic activity could benefit from the improvement of its intrinsic cell penetrating properties and

nuclear localization. Moreover, three decades of research have enabled the discovery and thorough characterization of several short cell-penetrating peptides and nuclear localization signal sequences.

Hypothesis:

We hypothesize that constructs resulting from the fusion of Omomyc to CPP and short-stretch substitution to NLS will display improved intrinsic cell penetration properties.

Objectives:

- Express and purify a panel of Omomyc-derived variants displaying the SynB3, Penetratin or R9 sequences as an N-terminal tag.
- Investigate the impact of the CPP fusion and Omomyc sequence short stretch mutations over a construct secondary structure, as well as over its ability to dimerize with the MAX b-HLH-LZ domain and to bind DNA.
- Characterize the cell penetrating properties of the purified Omomyc-derived variant panel

Section 2: Impact of the net charge on the Omomyc-derived mini-proteins mode of action.

A growing body of research in the field of therapeutic mAb shows that modulation of their net charge can be a successful strategy to improve an antibody pharmacokinetic profile *in vivo*. Exploring a similar strategy for the b-HLH-LZ format will deepen our understanding of the role of the basic amino acids (including those not involved in DNA binding) in the underlying dynamics determining a b-HLH-LZ protein's solubility, stability and pharmacokinetic profile.

Hypothesis

We hypothesize that the impact of net charge modulation on PK profile observed in mAb can be recapitulated in b-HLH-LZ proteins. Moreover, we predict that this modulation can be implemented without significantly impairing the Omomyc charge variants' ability to bind MYC, MAX and cognate DNA sequences or their cell penetration properties.

Objectives

- Express and purify a panel of Omomyc charge variants

- Study the impact of net charge modulation over solubility in different conditions
- Explore the effect of basic residues mutation over structural stability as well as binding to MYC, MAX and cognate DNA
- Develop an assay for the determination of the charge variants PK profile *in vivo*.

III. Material and Methods

1. Bacterial transformation and protein expression

Competent *E. coli* BL21 (DE3) arabinose-inducible (Invitrogen®, Waltham, U.S.A.) bacteria were thawed from a glycerol stock and left on ice for 15 min prior to mixing with a pET3-a (Novagen) expression vector containing the protein of interest coding sequence, in sterile conditions. All the expression plasmids were purchased from Genscript (Genscript Biotech B.V., Rijswijk, Netherlands). The solution was then left on ice for 30 min, before being heated to 42 ºC for 45 seconds, using a thermoblock. The cells were placed back on ice for 5 min. 500 µL of Super Optimal Broth were then added to the mix, and was left incubating at 37 ºC, shaking at 300 rpm for 1 h. The solution was then centrifuged for 5 min at 10,000 rpm and 450 µL of the supernatant were discarded. The 50 µL left were re-suspended and plated on a Lysogeny Broth (LB)-agarose plate, supplemented with carbenicillin for transformed bacteria selection. The plates were then left overnight at 37 ºC and kept sealed at 4 ºC afterwards for a maximum of 1 month.

To start bacterial pre-cultures, 500 mL of sterile LB supplemented with carbenicillin were inoculated with a single colony of transformed *E. coli*. The shake flask was left incubating overnight at 37 ºC, under agitation. The next day, bacterial culture was started by addition of 20 mL of pre-cultured bacteria material per liter of 2x Yeast Extract Tryptone (2YT) culture medium, supplemented with carbenicillin. The resulting mixture is left incubating at 37 ºC, under 200 rpm agitation. When the OD_{600} of the solution reaches a value greater than 0.6, the recombinant protein expression was induced through the addition of 15 mM of arabinose in the bacteria culture medium. Bacteria were then left incubating overnight in the same conditions. The next day, after approximately 14 hours of expression, the cells were collected by centrifugation (10,000 x g for 10 min) and the cell pellet was resuspended with 3 mL / g of pellet of resuspension buffer (700 mM NaCl, 5 mM $MgCl_2$, 50 mM KH_2PO_4, pH 7) and kept frozen until purification.

2. Omomyc and Max* mini-proteins purification

The frozen resuspended cell paste was first thawed in a water bath at 37 ºC. The cell paste was then supplemented with 0.4% Triton X-100 and DNase I, left 60 min at 37 ºC with gentle shaking for incubation, and centrifuged at 19,000 rpm for 40 min at 4 ºC. Afterward, the supernatant was collected and diluted with two volumes of H_2Odd. The Omomyc or Max* mini-protein present in the diluted supernatant were captured by CEX, using prepacked columns of Capto S resin (Cytiva, Marlborough, U.S.A.) operated by an Äkta pure system (Cytiva). Fractions with sufficient purity were pooled and further purified by hydrophobic interactions chromatography (HIC). Traces of host DNA were eliminated by anion exchange chromatography (Capto Q, Cytiva) and traces of endotoxins were removed from the pure protein pool using a commercial kit (QIAGEN). Finally, buffer was exchanged to 25 mM sodium acetate, pH 4.8 (Omomyc) or 100 mM KH_2PO4, 100 mM KCl, pH 4.8 (Max*), each pool was then concentrated using Amicon® centrifugal filter (ACF) tubes with a 3 kDa molecular weight cut-off (MWCO), sterile filtered, aliquoted and conserved at -80 ºC.

3. HisTag purification

The frozen resuspended cell paste was first thawed in a water bath at 37 ºC. The cell paste was then supplemented with 0.4% Triton X-100 and DNase I and left 60 min at 37 ºC with gentle shaking for incubation. The cell paste was then diluted 1:1 with a lysis solution (8 M urea, 20 mM imidazole) and sonicated on ice for 4 min (4 intermittent rounds of 1 min). Cells debris were pelleted after a 20 min centrifugation at 18,000 x g, at 4ºC. The supernatant was then filtered using 0.22 µm polyethersulfone (PES) filters and loaded on prepacked HisTrap® columns (Cytiva), operated by an Äkta pure system, for capture of the protein of interest. Fractions with sufficient purity were pooled together and concentrated using ACF tubes. Using prepacked HiTrap Desalting columns (Cytiva), the buffer of the concentrated pool of protein was exchanged to a storage buffer (100 mM KH_2PO4, 100 mM KCl, pH 4.8). In a final step, the desalted pool of pure protein was concentrated again, filtered, aliquoted and conserved at -80 ºC.

4. StrepTag purification

After being thawed, the cell paste was diluted with two volumes of a lysis buffer (2 mM EDTA, 165 mM NaH_2PO_4, 20 mM $MgCl_2$, 8M urea, pH 7) and solid urea was added to reach a final urea concentration of 8 M, together with 0.2% of pDADMAC flocculating agent. The solution was left agitating at room temperature (RT, about 25 ºC) for 2 h and was subsequently diluted with 1 volume of a dilution solution (2 mM EDTA, 165 mM NaH_2PO_4, 20 mM $MgCl_2$, pH 7). The solution was then centrifuged for 30 min at 18,533 x g, at 4 ºC. The pellet was discarded, and the supernatant filtered with 0.22 µm PES filters before being loaded to StrepTrap XT prepacked columns (Cytiva) for the capture of the protein of interest. Fractions with sufficient purity were pooled together and diluted with 1 volume of a denaturing and reducing solution (8 M urea, 100 mM NaH_2PO_4, 20 mM dithiothreitol (DTT), pH 7). The diluted pool was then loaded in 30 kDa-MWCO ACF tubes to filter out high molecular weight contaminants. Permeate was collected and concentrated with 3 kDa-MWCO ACF tubes prior to final buffer exchange using prepacked HiTrap Desalting columns (Cytiva) to 50 mM sodium acetate, pH 5 (Omomyc variants SO, OLS, SLO, V108, V109, V110, V111 and V112) or 50 mM NaH_2PO4, pH 7 (V113). Finally, the desalted pool was concentrated again, sterile-filtered, aliquoted and conserved at -80 ºC.

5. Mini-protein solubility assay

30 mg of pure peptide samples were diluted in a denaturing buffer (8 M urea, 100 mM NaH_2PO_4, pH 7) to a concentration of 2 $g.L^{-1}$. The diluted solutions were then loaded in a 50 mL Superloop® (Cytiva), integrated to an Äkta pure system, to allow precise injections of 6 mg (3 mL) into 5 HiTrap® Desalting columns of 5 mL installed in a series. Each protein solution injection was preceded by a column equilibration step carried out with the buffer to be tested. Elution was allowed by injection of buffer of interest into the columns, after each peptide injection. Injection and elution steps were performed with a flow rate of 2.5 $mL.min^{-1}$, whereas the equilibration steps were carried out with a flow rate of 5 $mL.min^{-1}$. Protein stocks whose buffer was successfully exchanged were collected in a separate tube than the protein fraction eluting with the salts of its loading buffer. Both fractions

were then visually inspected regularly for signs of protein precipitation for 1 h following their elution.

6. Circular dichroism

All the CD data were acquired with a J-1500 CD spectrometer (JASCO, Tokyo, Japan) equipped with a multi-cell holder and a Peltier system coupled to a reflux water bath for temperature control. Unless indicated otherwise, all the samples were prepared in 50 mM KH2PO4, 50 mM KCl, 2 mM tris(2-carboxyethyl)phosphine (TCEP), pH 6.7 (CD 1x buffer), and placed in quartz cells of 1 mm optical path length for measurements. To perform DNA binding experiments, the double-stranded canonical E-box DNA probe was prepared by mixing equimolarly and heating the following single strand oligonucleotides: 5'-GCG CGG G**CA CGT G**GG CCG GGG-3' and its complementary reverse sequence at 95 ºC for 5 minutes followed by a slow cooling to room temperature. The double stranded non-specific DNA probe (forward sequence: 5'-GCG CGG G**GG ATC C**GG CCG GGG-3') was purchased already annealed. Both DNA probes were ordered from Integrated DNA Technologies (IDT, Coralville, U.S.A.).

Prior to analysis, oxidized Omomyc, Max* and variants stocks were systematically diluted to an appropriate concentration and reduced by incubation at 37 ºC in the presence of 2 mM of TCEP for 2 h. For heterodimerization assays, the reduced protein stocks were mixed 1:1 to allow the formation of heterodimers and left incubating for 15 min at RT. For DNA binding assays, the DNA probe was first diluted in ddH_2O, mixed to the reduced stocks of putative homodimers or heterodimers and incubated for 1 h at 37 ºC prior to analysis. For the optimization path leading to the determination of the sample preparation protocol for CD-based DNA binding assays, please refer to section IV.1.4.

6.1. Far-UV spectrum measurement

Samples were placed at least 15 min in the temperature-regulated holder, prior to analysis. Spectra were recorded with the following parameters: data pitch: 1 nm, D.I.T.: 1 sec, bandwidth: 1 nm, scanning speed: 100 nm/min. Each spectrum representation corresponds to the average of a 3-scan accumulation.

6.2. Thermal denaturation and protein melting temperature determination

Following a far-ultraviolet (UV) spectrum measurement at initial temperature, the protein unfolding was recorded following CD values at 222 nm, with a heating rate of 1 ºC/min, D.I.T. of 1 sec and bandwidth of 1 nm. Temperature control sensor was set to holder and monitor sensor to cell.

After acquisition, thermal denaturation data were formatted to be manually fitted using the Microsoft Excel®-based solution for b-HLH-LZ thermal denaturation data fitting, developed and kindly shared by the laboratory of P. Lavigne[201] (example shown in Figure 11). Melting temperature (Tm) was obtained by determining the highest point of the fitted data 1st derivative.

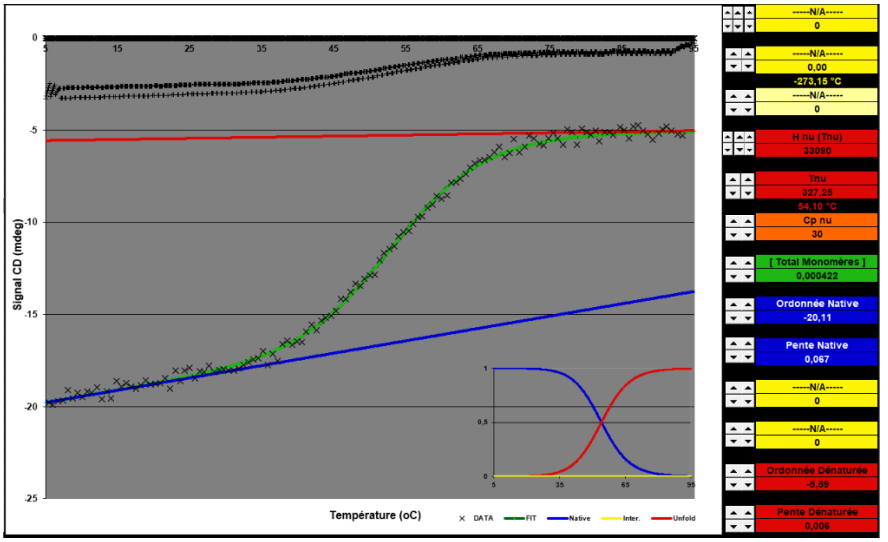


***Figure 11**, Example of manual thermal denaturation data fitting using the Excel® spreadsheet "CD_F". The fitting curve function was derived from the Gibbs-Helmholtz equations. Optimal fitting curves were obtained by incremental modification and visual determination of the transition thermodynamic parameters values. Those parameters include the denaturation enthalpy, temperature, or difference of heat capacity resulting from the structural transition. The Tm corresponds to the abscissa of the fitting data first derivative highest point. Back crosses: experimental data, green line: data fitting.*

7. Cell lines and cell culture

The A549 NSCLC cell line was purchased from the American Type Culture Collection (ATCC), the EFO-27 and A2780 ovarian cancer cell lines were provided by the laboratory of Manel Esteller (Josep Carreras Leukaemia Research Institute, Barcelona, Spain), the Normal Human Dermal Fibroblast (NHDF) cell line was provided by the laboratory of Gerard Evan (The Francis Crick Institute, London, United Kingdom) and the BZ1 mouse fibroblast was provided by the laboratory of Bruno Amati (European Institute of Oncology, Milan, Italy). The A549, EFO-27 and A2780 cells were grown in Roswell Park Memorial Institute (RPMI) medium (Life Technologies, Carlsbad, U.S.A.) supplemented with 10% fetal bovine serum (FBS) and 1% glutamine. NHDF and BZ1 cells were grown in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies) supplemented with 10 % FBS and 1% glutamine.

Cells were cultured in a humidified incubator set at 37 ºC with 5% CO_2. For cryopreservation, 10% dimethyl sulfoxide (DMSO) was added to the complete medium, and the cell vials were placed in a Mr. Frosty® container (Thermo Fisher Scientific, Waltham, U.S.A.) containing isopropanol and stored at -80˚C for a minimum of 48 hours. Afterward, the vials were transferred to liquid nitrogen storage. When thawing cells, the vials were briefly immersed in 37˚C water for 2 minutes, and the contents were then plated in a 10 cm dish with 20mL of complete medium. The next day, the medium containing DMSO was replaced with fresh complete medium. All cell lines were confirmed to be mycoplasma-free.

8. Endocytosis assay

For each condition tested, 200,000 A549 cells were first seeded in a well of a 24-well plate. One day after, the cell culture media was replaced with fresh media supplemented with 10 μM of Omomyc, Omomyc variant or a corresponding volume of vehicle (for non-treated, negative control cells). Prior to treating the cells, fresh supplemented media was pre-warmed to 37 ºC or kept at 4 ºC for control. Accordingly, cells were incubated for 15 min with supplemented media at 37 ºC or 4 ºC, or for 1 h at 37 ºC. The cells were then harshly trypsinized to detach them from the plate and digest any protein adhered to the membrane. After a 30 min incubation at 37 ºC in

presence of concentrated trypsin, the cells were collected in round bottom poly-propylene test tubes and maintained on ice during all the subsequent preparative steps until analysis, unless specified otherwise. After discarding the trypsin solution and a series of RMPI media / Phosphate-buffered Saline (PBS) washing steps, cells were incubated in the presence of the BD Horizon™ FVS-700 viability stain (BD Biosciences, Franklin Lakes, U.S.A.) in obscurity for 15 min at RT, labeling the cells with compromised cell membrane. The cells' membrane was then blocked for 10 min with a blocking solution (PBS, 2.5 mM EDTA, 5% (v/v) horse serum, 1% (w/v) BSA). Cells were then fixed and permeabilized using the BD Cytofix/Cytoperm™ Fixation/Permeabilization Kit (BD Biosciences) before being incubated again in blocking solution for 10 min, for intracellular blocking. Finally, intracellular mini-protein was detected using either the AF488-conjugated anti-HisTag (D3I1O) XP® Rabbit mAb (Cell Signaling Technology, Danvers, U.S.A.), or a combination of anti-Omomyc primary rabbit polyclonal antibody (pAb) and AF488-conjugated goat anti-rabbit antibody (ref. A11008, Invitrogen®). After washing the antibody excess, the cells were resuspended and left in PBS supplemented with sodium azide, at 4 ºC and protected from ambient light until analysis.

Flow cytometry experiments were carried out with a BD FACS Celesta flow cytometer (BD Biosciences). Data of 10,000 events (single living cells) were collected per tested conditions and were analyzed with the FCS Express 7 software (De Novo Software, Pasadena, U.S.A.). An example of gating that enables the selection of single living cells is shown in Figure 12.

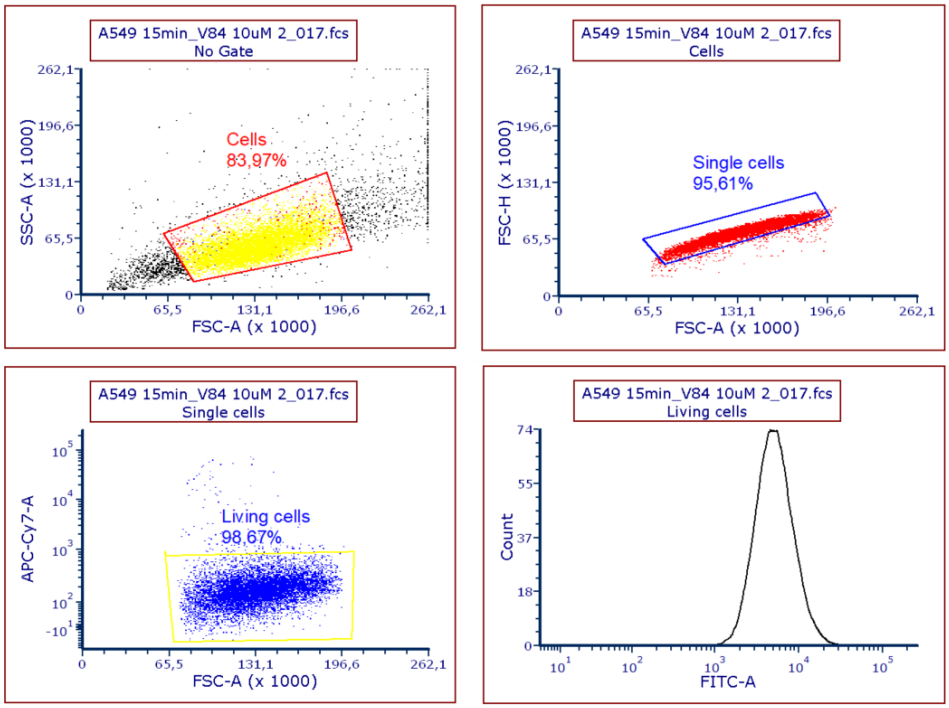


***Figure 12**, Example of gating to isolate single living cells, treated with 10 µM of V84 for 15 min at 37 ºC. Top-left, plot of the side scatter (SSC) peak area (SSC-A) (y-axis) of all events in function of forward scatter (FSC) peak area (FSC-A) (x-axis), events corresponding to cells were selected, filtering out cell debris, foreign particles or cells aggregates. Top-right: FSC peak height (y-axis) of cell-corresponding events in function of their respective FSC-A (x-axis), single cell events were gated to filter out cell doublets. Bottom-left, FV viability stain 700 fluorescence intensity (Sulfo-Cyanine 7 emission channel, y-axis) of single cell events in function of their respective FCS-A (x-axis), only the single cells with a fluorescence below a certain threshold were considered as living. Bottom-right, 1D histogram of single living cell events. Cell count (y-axis) is represented in function of AF488 fluorescence intensity (FITC emission channel, x-axis). Histogram representations were smoothed using the FCS express integrated data smoothing tool.*

9. Electro-mobility shift assay

9.1. Native gel composition

6%-acrylamide native gels were casted from a solution composed of: 50 % TA buffer 2x (80 mM Trizma® base, 40 mM sodium acetate, pH 8), 28.9 % H_2Odd, 20 % of 30% Acrylamide/Bis Solution, 37.5:1 (Bio-Rad), 1 % of a 10 % (w/v) ammonium persulfate solution, and 0.1% N,N,N',N'-

Tetramethylethylenediamine (TEMED), mixed in this order. Once polymerized, the gels were stored at least 1 day at 4 ºC before being used in an EMSA. Gels were casted and run with Mini-PROTEAN® Tetra Handcast Systems (Bio-Rad, Hercules, U.S.A.).

9.2. Samples preparation and assay conditions

Protein samples were first diluted to 12 µM (Omomyc variants) or 8 µM (Max*) in PBS + 1 mM DTT and left at 37 ºC for 2h for reduction. Reduced Omomyc variants sample were then mixed 1:1 with reduced Max* or with PBS + 1 mM DTT and left 15 min at RT for incubation. Homodimers and heterodimers samples were then mixed 3:1 with a 4 µM IRDye® 700-conjugated canonical E-Box dsDNA probes (forward: 5'-IRD700- GCG CGG G**CA CGT G**GG CCG GGG-3', purchased annealed from IDT) diluted in binding buffer 4x (80 mM Tris-HCl, 40 mM sodium acetate, 300 mM KCl, 10 mM DTT, 40 µg/mL poly dI-dC and 20% glycerol). Protein-DNA complexes were incubated for 30 min at RT prior to loading on the gel. Meanwhile the gels were run empty for 30 min at 100 V in TA buffer 1x, prior to loading. After sample loading, gels were run for 40 min at 100 V and were read using an Odyssey® imager (LI-COR Biosciences, Lincoln, U.S.A.).

10. Mini-proteins DNA-binding characterization by Enzyme-Linked Immunosorbent Assay (ELISA)

This assay relies on the use of clear 96-well Pierce™ Streptavidin High Binding Capacity Coated Plates (Thermo Fisher Scientific). First, the plate was washed with a lysozyme blocking buffer (or LBB: PBS, 0.5 % w/v chicken egg white lysozyme, 3 M urea). Biotinylated E-Box dsDNA probe (forward: 5'-Biotin dT- GCG CGG G**CA CGT G**GG CCG GGG-3', purchased annealed from IDT), diluted to 500 nM in LBB was then incubated on the plate for 3h at RT, rocketing. After washing steps with PBS + 0.3% (v/v) Tween™20 (PBST), the plate was then incubated with decreasing concentrations of variant (11 points, 2-fold serial dilution (500 nM to 0.5 nM homodimer)) for 2 h at RT, rocketing. The plate was then washed with PBST and subsequently incubated with a horseradish peroxidase (HRP) conjugated anti-StrepTag rabbit pAb (ref. A00875, Genscript Biotech B.V.), diluted in SuperBlock™ blocking buffer (or SBB, Thermo Fisher Scientific). Finally, after a last series

of washing steps with SBB, HRP activity was revealed using BD OptEIA™ TMB Substrate Reagent Set (BD Biosciences), and the reaction was killed by addition of 100 µL of a 0.18 M H_2SO_4 solution after 5 min. Absorbance at 450 nm was determined using a Spark® plate reader (TECAN, Männedorf, Switzerland).

11. Co-immunoprecipitation (co-IP)

11.1. Cell lysate preparation

NHDF cells were cultured until 70% confluency, were trypsinized to be collected and lysed on ice for 30 min with a mild lysis buffer (HEPES 20 mM, 180 mM NaCl, 10 mM MgCl2, 0,2% NP-40, 10% Glycerol, 1 mM EDTA, 1 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), pH 7.4) supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). The lysate was then centrifuged, and the resulting pellet was discarded. The supernatant protein content was quantified by DC protein assay (Bio-Rad). The lysate was then aliquoted and frozen at -80 ºC until the day of the co-IP assay.

11.2. Co-IP protocol

300 µg of total protein was used per conditions. The lysate was diluted with 1.25 volume of co-IP dilution buffer (HEPES 20 mM, 180 mM NaCl, 10 mM MgCl2, 1 mM EDTA, 1 mM EGTA, pH 7.4) and 50 µM or 10 µM final concentration of reduced Omomyc variant or a corresponding buffer volume was added and incubated at 4 ºC on a rotating mixer for 30 min. 7 µL of each sample were separated and frozen for future analysis of the whole spiked lysate. After being conditioned in co-IP dilution buffer, 20 µL of MagStrep® Strep-Tactin®XT (IBA Life Sciences, Göttingen, Germany) slurry were mixed to each sample and incubated for 30 min at 4 ºC. Using a magnetic tube rack, the beads were separated from the depleted lysate which was kept frozen until analysis. The beads were extensively washed with PBS buffer before being incubated with a StrepTag elution buffer (100 mM NaH_2PO_4, 75 mM NaCl, 4M urea, 50 mM biotin, 1 mM EDTA, pH 7). The elution fraction was kept frozen until analysis.

11.3. Western Blot analysis

Prior to gel loading, each of the three fractions generated for each tested condition were boiled in presence of reducing Laemmli buffer for 5 min at 95 ºC. All the fractions were then loaded on NuPAGE™ precast bis-tris 12% acrylamide gels (Invitrogen™, input: 10% of the spiked lysate, 10% of the depleted lysate and 50% of the beads elution fraction). Gels were run for 2h at 120 V and proteins were transferred to a PVDF membrane using the iBlot 3 Dry Blotting System (Life Technologies). The protein transfer was confirmed after staining the membrane with Ponceau red. After washing with PBS + 0.1% Tween and blocking with 5% milk, membranes were incubated overnight at 4ºC with a 1:1,000 diluted solution of the following primary antibodies: Anti-pan-myc rabbit antibody [EPR18863] (ref. ab195207, Abcam, Cambridge, United Kingdom), Anti-MAX rabbit antibody [EPR19352] (ref. ab199489, Abcam) and NWSHPQFEK Rabbit pAb (ref. A00626, Genscript). Membranes were washed with PBS-Tween® and incubated for 1 hour at room temperature with a 1:5,000 diluted solution of HRP-conjugated anti-rabbit IgG (ref. NA934, GE Healthcare, Chicago, U.S.A.). Membranes were washed twice with PBS-Tween and once with PBS, and then incubated for 5 minutes with SuperSignal West™ Pico Chemiluminescent Substrate (Thermo Fisher Scientific) before revealing. Enhanced chemiluminescence images were acquired using iBright™ CL1500 imaging system (Thermo Fisher Scientific).

12. Cell viability and proliferation assay

1,000 EFO-27, A2780 or BZ1 cells were initially seeded in each well of a 96-well cell culture plate. After 24 h, their respective culture media was replaced by fresh media supplemented with 20 µM of Omomyc, variants or a corresponding vehicle volume for control. After 5 days of treatment, 100 µL of alamarBlue™ Cell Viability Reagent (Thermo Fisher Scientific) was added to each well. After 3 h of incubation, the fluorescence intensity at 590 nm was measured (excitation at 560 nm) using a Spark® plate reader (TECAN).

13. Immunofluorescence (IF)

For each tested condition, 50,000 cells were seeded in a Millicell EZ slide 8-well glass well (Millipore, Burlington, U.S.A.). 24 h after seeding, the cell culture media was replaced by fresh media supplemented with 20 µM of Omomyc variants or an equivalent vehicle volume for control. After a 1-day treatment, the supplemented media was aspirated, and the cells were gently washed twice with PBS. A 4% perfluoroalkoxy alkane (PFA) solution in PBS was then used to fix the cells (15 min incubation at RT, rocketing), prior to their permeabilization with a 5 min incubation with 0.5% Triton™ X-100 in PBS, at RT. Subsequently, cells were incubated 40 min at RT with a blocking buffer (PBS, 2 % BSA, 0.2 % Triton™ X-100, 1% goat serum). Cells were then exposed to a primary antibody mix composed of m2 anti-Omomyc mouse mAb (1:100 dilution) and NWSHPQFEK Rabbit pAb (Genscript, 1:100 dilution), for 1 h at RT. Excess of antibody was washed, and cells were then incubated for 40 min at RT with a secondary antibody mix composed as follow: AlexaFluor 488-conjugated goat anti-rabbit IgG (H+L) (ref. 926-68020, LI-COR, 1:500 dilution) and AlexaFluor 568-conjugated goat anti-mouse IgG (H+L) (ref. A11004, Invitrogen®, 1:500 dilution). After washing of the secondary antibody excess, the chambers were removed, and a cover slip installed, using 4',6-diamidino-2-phenylindole (DAPI)-containing mounting media (ProLong™ Gold antifade reagent with DAPI, Invitrogen). The slides were left at least 24h at 4 ºC, in the dark, prior to analysis. IF images were acquired using a THUNDER Imaging System epi-fluorescence microscope (Leica Microsystems, Wetzlar, Germany). Images were analyzed and treated using the ImageJ Fiji package.

14. Competition assay for StrepTag concentration determination in blood-derived matrices

14.1. Blood-derived samples

All the work involving research animals was carried out strictly following all the ethical guidelines in force for the accommodation and care of animal used for scientific experimentation established by the Animal Research Ethics Committee of the Vall d'Hebron Institute of Research (VHIR)[244], following regional, national and European guidelines.

Human blood was collected from two healthy donors by the medical team of the Multiple Sclerosis Center of Catalonia (Cemcat). Rat blood derivatives were obtained through donation of blood left-overs, collected intracardially, from the Advanced Chronic Liver Disease Group (VHIR). Mice blood was collected through intracardiac punction by our group. Blood of either origin was collected in EDTA plasma tubes or in serum tubes and was processed through centrifugation at 10,000 x g for 10 min at 4 ºC. Supernatant corresponding to either plasma or serum was collected and supplemented with Omomyc variants or left untreated. All the frozen plasma or serum samples were conserved at -80 ºC until use.

14.2. Competition assay

Goat anti-rabbit IgG-coated black 96-well plates with non-proteic blocking (Biomat Srl, Ala, Italy) were used in this assay. Plates were washed with SB buffer, prior to incubation with a 1:1,000 dilution NWSHPQFEK Rabbit pAb (ref. A00626, Genscript) solution for 2 h, at RT. Meanwhile, a 70 nM GFP-StrepTag (purchased from IBA Life Sciences) solution was prepared in PBST and mixed 1:1 with the samples of interest, standard curve samples or control samples. After washing off the excess of NWSHPQFEK Rabbit pAb with PBST, the GFP-StrepTag-spiked samples were added to the plate and incubated for 1 h at RT. Plates were then washed with PBST and 100 µL of PBS added in each well for fluorescence reading. GFP fluorescence intensity at 535 nm (excitation at 485 nm) was determined using a Spark® plate reader (TECAN).

For the optimization path leading to the determination of the protocol of this assay, please refer to section IV.2.7.

IV. Results

Section 1:

Modulation of the cell penetration activity and nuclear localization of b-HLH-LZ proteins using short-stretch mutations and cell penetration tags

1.1. Design of the constructs

The work presented in this section aimed at understanding the effect of the fusion of CPP tags and short-stretch mutations on the Omomyc main biological features and stability.

The CPPs SynB3, Penetratin, and R9 were fused to the N-terminus of the original Omomyc sequence, the Max* sequence and 5 distinct mutated Omomyc sequences, namely OmoT3, OmoCrot, OmoNLS2, OmoEGFR and OmoNLS4. In the first result part of this thesis, I will describe the impact of the fusion of these three CPPs with the Omomyc sequence. Moreover, I will describe their association to 5 alternative Omomyc-derived constructs, namely OmoT3, OmoCrot, OmoEGFR, OmoNLS2 and OmoNLS4.

OmoT3 is a C-terminus truncated Omomyc-derived sequence. Its design was, in part, motivated by the claim of Brown and colleagues that shorter forms of Omomyc, conserving the MYC b-HLH domain but truncating its leucine zipper, could recapitulate Omomyc activity on 2D cell proliferation[193], although intriguingly, the MYC b-HLH-LZ domain alone did not display any cell penetrating activity when tested[245]. The OmoT3 deleted sequence corresponds to 11 amino acids or 3 complete α-helix turns, conserving the 2 last Omomyc amino acids, including the cysteine residue. The main goal of the OmoT3-based variants design in this study is to challenge the hypothesis that shorter Omomyc-based sequences could penetrate cells more easily.

OmoCrot comprises the first 7 amino acids of the Crotamine protein in its loop. Crotamine is a cell-penetrating protein found to be one of the main constituents of rattlesnake venom[246–248]. Interestingly, crotamine was found to specifically target rapidly dividing cells[249], giving it the potential to make it a good carrier to target tumoral cells. The additional substitution S51T carried by the OmoCrot sequence was made to prevent this threonine phosphorylation, since it was shown to be linked to MYC degradation[10]. Finally, with the same rationale than for the OmoT3 sequence, OmoCrot C-terminus was truncated to lose the 4 amino acids preceding the Omomyc cysteine residue.

The OmoEGFR sequence was obtained by the fusion of Omomyc C-terminus with an EGFR-targeting peptide sequence. This peptide was discovered by phage display and was shown to efficiently enter A431

squamous cell carcinoma cancer cells through EGFR interaction and clathrin-dependent endocytosis[250].

To design OmoNLS2, a large portion of the Omomyc loop region was substituted by a NLS sequence isolated from the SV40 large T antigen protein. Interestingly, this lysine-rich sequence was historically the first one to be defined as NLS by Kalderon and Colleagues in a key publication of 1984[251]. OmoNLS4 displays the NLS isolated from the Oct-6 protein on its C-terminus. In a study comparing the cellular uptake and localization of various NLS sequences (isolated from NF-κB, TFIIE-β, or SV-40, among others) the Oct-6 NLS was shown to be the most efficient at penetrating the MCF-7 breast cancer cells and was found to colocalize with both the cytoplasm and nuclei within the cells[252].

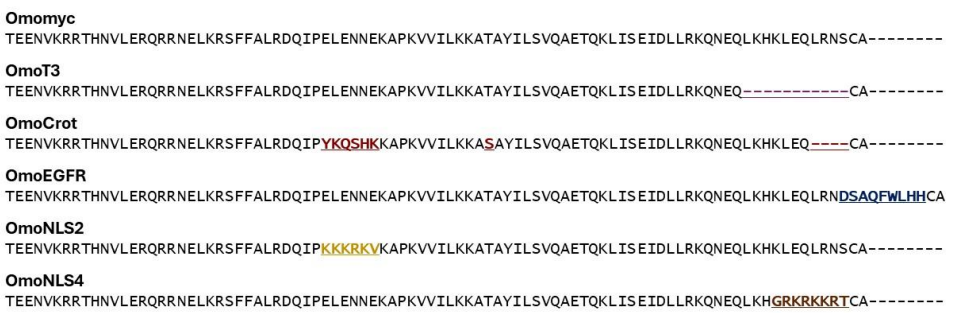


***Figure 13**, alignment of the Omomyc amino acid sequence with the sequences of 5 Omomyc-derived mini-protein: OmoT3, OmoCrot, OmoNLS2, OmoEGFR and OmoNLS4. Each respective mutation is highlighted with colored, bold and underlined characters.*

Additionally, a HisTag was added to the N-terminus of each construct to facilitate their purification and detection for direct cell penetration properties comparison (Figure 14). Each variant was given a unique numerical identifier to facilitate referencing, and a specific color designation for the representation of the corresponding results (cf. table 1). In addition to the variants described in Figure 14, another pair of variants, namely H8-Omo-R9 (or V90) and H8-Max*-R9 (or V91), was designed to study the effect of the R9 tag position over the variants properties by comparing them to their respective counterparts H8-R9-Omo (V88) and H8-R9-Max* (V89).

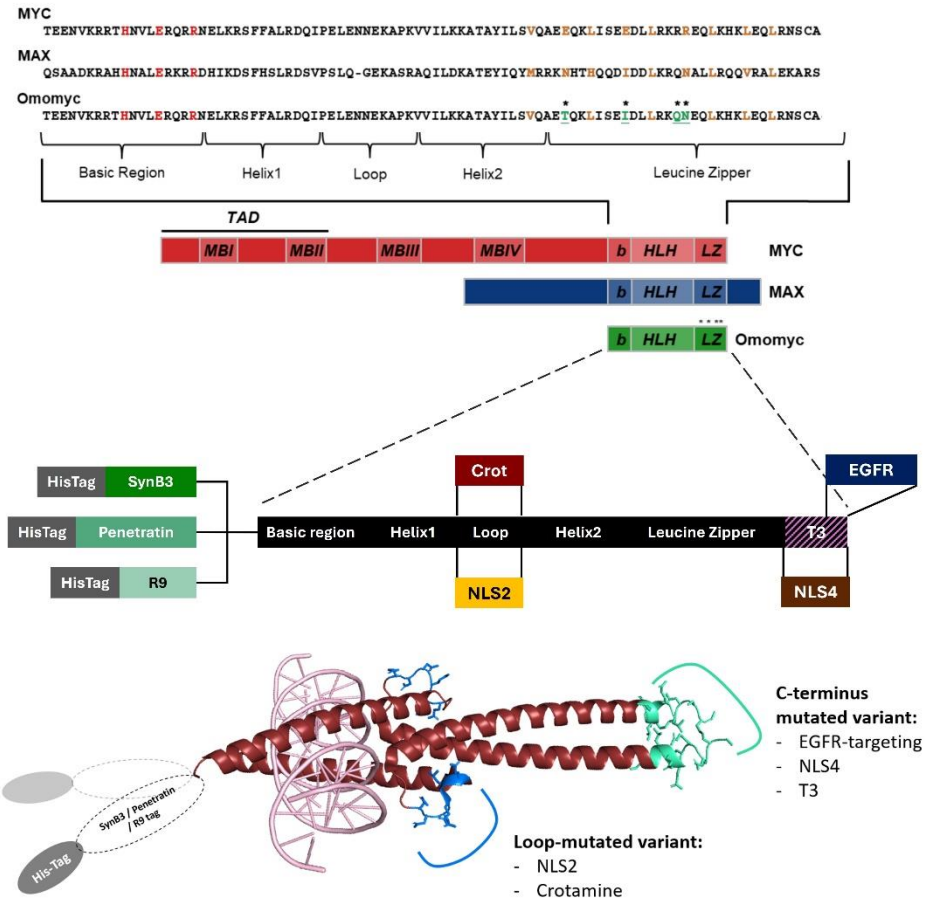


***Figure 14**, From top to bottom: MYC/MAX/Omomyc b-HLH-LZ region alignment; schematic description of the Omomyc regions and position of the variants substitutions/insertions; adapted X-ray crystallography structure of Omomyc (burgundy) bound to an E-Box DNA sequence (pink) (PDB: 5I50), the amino acids substituted in one or more Omomyc variants are colored in blue (loop) and turquoise (C-terminus).*

This Omomyc and Max*-derived variants set was originally designed and studied to support a patent application protecting the use of Omomyc-derived variants[245], crucial for the IP protection of Peptomyc's main asset and paving the way for the design of an Omomyc-based second generation of MYC inhibitors with improved properties. Having this large variant panel at hand, even though incomplete for a formal, publication-aimed study, represented a great opportunity to learn more about our main asset, and the effect of short stretch-mutations over its main features and stability. Hence, the characterization work presented in this thesis was carried out after the successful filing of this patent.

During this study, the work was divided into several steps. In order to get familiar with HisTag affinity chromatography as well as with the effects induced by the fusion of Omomyc with CPP sequences over stability, although for practical reasons the whole variants set was expressed at the same time, the variants with the simpler designs, namely V84 to V91, were studied first.

***Table 1**, variants set number identifier, description sequence length (amino acids) and theoretical average mass (Da).*

Name	Description	Sequence length	Mass (Da)
V84	8xHis-SynB3-Omomyc	109	13286.08
V85	8xHis-SynB3-Max*	107	12978.56
V86	8xHis-Penetratin-Omomyc	113	13895.02
V87	8xHis-Penetratin-Max*	111	13587.36
V88	8xHis-R9-Omomyc	108	13313.3
V89	8xHis-R9-Max*	106	13005.64
V90	8xHis-Omomyc-R9	108	13313.3
V91	8xHis-Max*-R9	106	13005.64
V93	8xHis-SynB3-OmoT3	98	11938.64
V94	8xHis-Penetratin-OmoT3	102	12547.44
V95	8xHis-R9-OmoT3	97	11965.72
V96	8xHis-SynB3-OmoCrot	105	12844.83
V97	8xHis-Penetratin-OmoCrot	109	13453.63
V98	8xHis-R9-OmoCrot	104	12871.91
V99	8xHis-SynB3-OmoEGFR	117	14321.35
V100	8xHis-Penetratin-OmoEGFR	121	14930.15
V101	8xHis-R9-OmoEGFR	116	14348.43
V102	8xHis-SynB3-OmoNLS2	109	13325.53
V103	8xHis-Penetratin-OmoNLS2	113	13934.32
V104	8xHis-R9-OmoNLS2	108	13352.61
V105	8xHis-SynB3-OmoNLS4	109	13328.36
V106	8xHis-Penetratin-OmoNLS4	113	13937.15
V107	8xHis-R9-OmoNLS4	108	13355.44

1.2. A simple, HisTag affinity chromatography-based process enabled the purification of V84 to V89

Expression of the variants was obtained transforming *E. coli* BL21 with a pET3a vector containing the sequence of interest, and culturing single colonies in LB medium supplemented with carbenicillin in shake flasks. This process was straightforward since excellent expression levels were obtained for almost all Omomyc variants without the need of any optimization of the standard expression protocol (Figure 15).

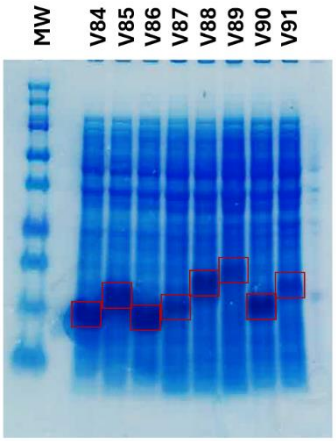


***Figure 15**, SDS-PAGE (Coomassie blue staining) of full protein extracts of E. coli cells after expression of their respective variant. The bands expected to correspond each of the protein of interest in each lane are circled in red.*

Interestingly, Omomyc-derived constructs seemed expressed at higher levels than each corresponding Max* variant. Similar expression levels were obtained for both V88 and V90, as well as for V89 and V91, suggesting that the His-tag position did not affect their capacity to accumulate within the cells.

After overnight expression, *E. coli* cells were centrifuged and resuspended in a hypertonic buffer before being frozen at -80 ºC. After being thawed, the resuspended cell pastes were incubated with Triton 100X and DNase I at 37 ºC. Cells were further lysed by going through several cycles of sonication on ice. Lysates were then centrifuged, and the supernatant collected, and the cells debris pellet was discarded.

A first purification attempt was carried out using the nickel-functionalized Bio-Rad IMAC resin at small scale, using a gravity column format. In this case, the HisTag position at the C-terminus seemed to be detrimental for V90 and V91 extractability and affecting their purification yields. On the

other hand, good binding to the resin was observed for all tested variants, since each respective flow-through fraction was depleted of the variant-corresponding band compared to the load-in lane. However, exceptions to this observation were noted for V84 and V90. In the case of V84, its presence in the FT fraction is likely due to a saturation of the resin, judging by its important concentration in the load-in fraction. As for V90, a strong band is observed in the FT fraction while no bands can safely be attributed to this variant in the elution fractions. These observations suggest that the purification of this variant may have failed due to poor extractability, solubility and/or degradation issues. Apart from the V90 construct, all variants could be successfully purified to some extent (Figure 16). These results proved the viability of the use of HisTag affinity chromatography for construct purification.

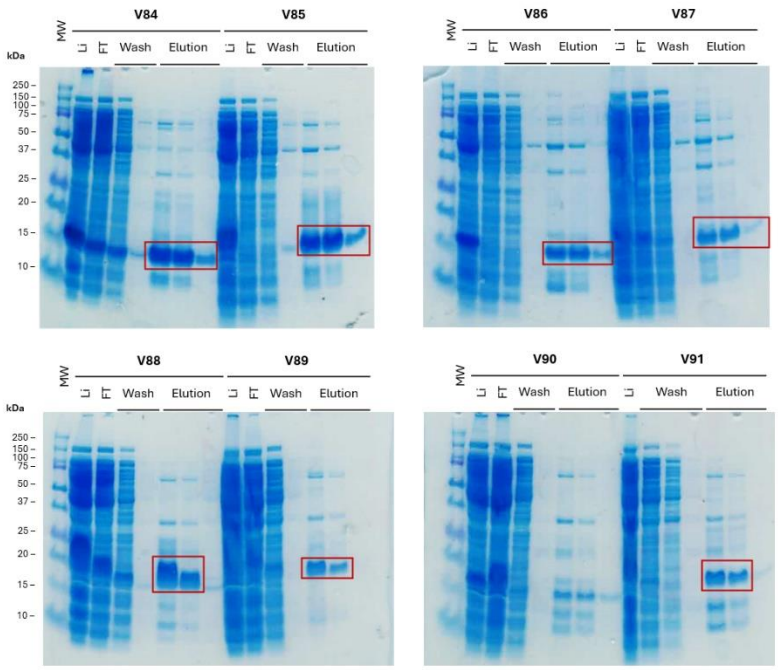


***Figure 16**, Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis SDS-PAGE (Coomassie blue staining) of the HisTag affinity chromatography purification fraction for variants V84 to V91. 5 fractions were collected for both the washing and elution steps: washing step fractions 1 and 5 were loaded (leftmost and rightmost, respectively), elution fractions 1, 3 and 5 were loaded (in this respective order, from left to right). Li: Load-in; FT: Flow-through. The V91 flow-through fraction was lost, the washing step fraction 3 was loaded instead for this variant. The same volume of every fraction was loaded per well, the difference of band intensity observed is therefore directly proportional to a difference of concentration. Bands suspected to correspond to the protein of interest are circled in red. The 4 gels presented in this Figure were all run and stained in parallel.*

Even if the gravity column format has the benefit of representing a faster approach to purify a large number of proteins simultaneously, we decided to move away from it and switched to HisTrap® HP prepacked columns operated by an Äkta Pure system. The reason for this choice was that, compared to gravity columns, this system allows for a tighter control over the purification process parameters, improving the variants-containing elution fractions purity and optimizing the recovery yields. Indeed, considering the high number of variants, developing an extraction and purification protocol compatible for all constructs appeared as a necessity to limit the time spent on troubleshooting to a minimum. Moreover, protein stocks obtained through the same process are likely to present similar impurities, therefore ensuring better comparability during the subsequent characterization steps.

The use of urea was then considered to enhance the extraction step, but also to improve the variants solubility in the conditions suitable for His-tag affinity chromatography. Moreover, according to the manufacturer, the HisTrap® HP columns are compatible with urea concentrations of up to 8 M.

The first V84 purification attempt demonstrated a relatively efficient binding, even in presence of 4 M urea, but ended with a long tailing peak indicative of a lack of solubility in the elution condition which did not include urea initially. In subsequent purification attempts, urea was progressively incorporated in the elution buffer until determining an optimal concentration of 4 M, allowing for more efficient elution steps. For each variant, the elution fractions with sufficient purity were pooled together and concentrated using 3 kDa molecular weight cut-off (MWCO) Amicon® Ultra Centrifugal Filters tubes. This step, although not crucial, was carried out to facilitate the following desalting / buffer exchange step.

Buffer was then exchanged to get rid of urea and imidazole, whose presence is unsuitable for biophysical characterization, also to be able to store each protein stock in conditions more adequate for long term stability. In fact, our extensive experience with MYC, MAX and Omomyc-derived variants has shown that those proteins all exhibit optimal solubility in acidic buffers. In addition, in prevision of the following characterization experiments, the elution buffer was exchanged for a storage buffer (100 mM monobasic potassium phosphate, 100 mM potassium chloride, pH 4.3), compatible with the subsequent characterization by circular dichroism. This buffer exchange step from a neutral pH buffer to an acidic buffer also has the

advantage of eliminating an important part of residual host DNA contamination, highlighted by an improvement of the A_{260}/A_{280} ratio of the purified protein pools.

Eventually, the variants V84 to V91 could be successfully expressed and purified following the workflow recapitulated in Figure 17D. However, yields obtained for variants V90 and V91 were very low, originating from much poorer extraction efficiency when compared to the other 6 constructs. Since the variants V90 and V91 did not seem to hold any advantage compared to their V88 and V89 counterparts, they were discontinued at this stage and were not further studied.

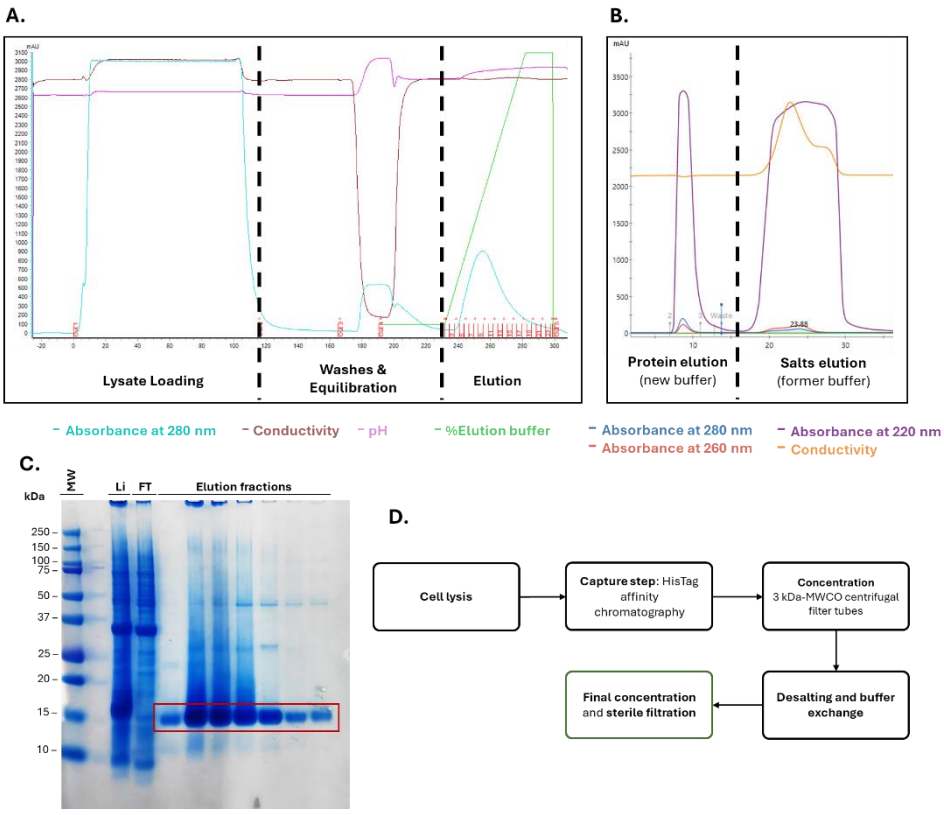


***Figure 17**, Representative example of the purification steps monitoring results obtained with V86. (**A**) Chromatogram of the HisTag affinity chromatography capture step. Absorbance at 280 nm is represented in blue, conductivity in brown, pH in pink and the proportion of elution buffer in the mobile phase in green. (**B**) Chromatogram of a buffer exchange step from the capture step elution buffer to storage buffer (100 mM monobasic potassium phosphate, 100 mM potassium chloride, pH 4.8). (**C**) SDS-PAGE (Coomassie blue staining) of the chromatography capture step fractions. (**D**) Variants downstream processing workflow.*

1.3. CPP-Omomyc fusion proteins have comparable secondary structures and retain the ability to bind Max*, but show signs of impaired DNA binding by CD

Circular dichroism (CD) was used to study the effect of the fusion of the different CPP tags with Omomyc or Max* on their secondary structure and Tm. All 6 constructs spectra displayed negative bands at 208 nm and 222 nm, sign of a rich α-helix content, similarly to their respective parent peptide. Interestingly, while the Omomyc variants presented a similar Tm than the one of Omomyc, the Max* variants displayed a slightly higher Tm than their parental sequence. Overall, though, these results show that the addition of none of the studied CPP tags has a dramatic effect on the secondary structure or Tm of the Omomyc or Max* b-HLH-LZ (Figure 18).

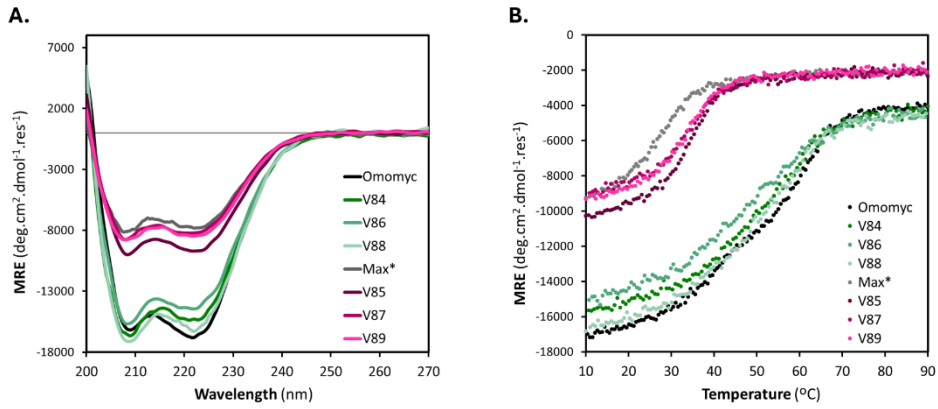


***Figure 18**, Overlay of the far-UV CD spectra at 20 °C (**A**) and thermal denaturation by CD at 222 nm (**B**) of Omomyc (black), Max* (grey) and the variants V84 (dark green), V85 (purple), V86 (green), V87 (magenta), V88 (light green) and V89 (pink) at 16 µM in CD buffer.*

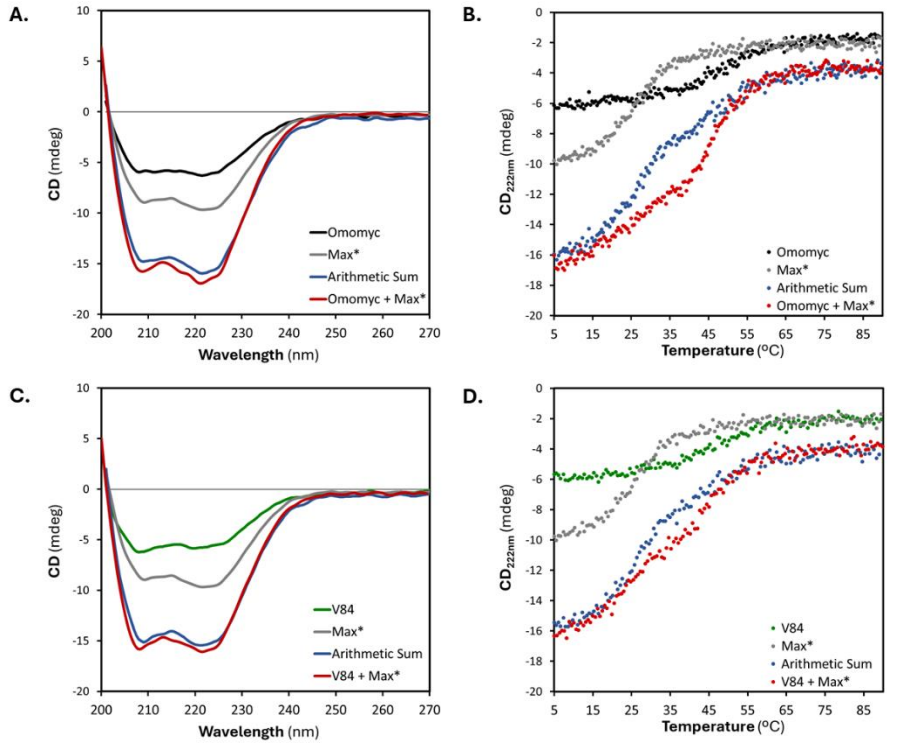
CD was also used to control whether the Omomyc variants' ability to form heterodimers with the Max b-HLH-LZ was conserved (Figure 19). In these experiments, the curve obtained for the Max b-HLH-LZ - here represented by the Max* construct - is arithmetically summed to the curve obtained for Omomyc and Omomyc variants. This arithmetic sum can then be compared to an experimental sum curve obtained by performing the measurements of samples containing both proteins. An overlap of both curves would be either indicative that both proteins do not interact with each other or that their interaction does not induce any change in secondary structure or thermal stability.

CD spectra of Omomyc, V84, V86 and V88 in the presence of Max* show only a modest increase of molar ellipticy, when compared to their respective arithmetic sum curve, at 5 ºC. In contrast, a more pronounced difference was observed during thermal denaturation of the putative heterodimers. In fact, the characteristic Max* transition visible between 15 ºC to 35 ºC is attenuated in each of the experimental sum data sets suggesting that the formation of putative heterodimers has a stabilizing effect on the Max* secondary structure. Despite non-optimal concentration conditions (in these experiments, 3 µM of either Omomyc or Omomyc variants were incubated with 8 µM of Max*), the CD results show convincing signs of heterodimeriation between Omomyc or Omomyc variants with the Max b-HLH-LZ. Moreover, unshown repetition of this experiment, carried out in the presence of a slight relative molar excess of the Omomyc variants, showed an even clearer disappearance of the Max* transition from each respective thermal denaturation curve. These results indicate that the Max*-corresponding transition disapearance is variant-concentration dependent and therefore most likely relates to Max* heterodimerization state. Since very similar results were obtained for all the Omomyc-related constructs during this experimental set, it was concluded that none of the CPP tag interfered significantly with the Omomyc variants ability to heterodimerize with Max.

The final step of this preliminary characterization aimed to understand the effects of the CPP tags on DNA binding. CD was once again used, following the same experimental rationale than for the Max-binding experiments. When bound to DNA, the helical secondary structure of b-HLH-LZ proteins such as Max* or Omomyc, is typically stabilized. This phenomenon, presentated in Figure 20A-B, can be seen in both the far-UV CD spectum in the shape of an amplification of the α-helix-corresponding minima at 208 nm and 222 nm, as well as in thermal denaturation experiments, during which significant increases in Tm (from 10 to 25 ºC) are observed. In this context, results obtained for the variants V84 to V89 suggest that DNA binding has a different effect on their structure compared to their parent peptide. Indeed, as presented in Figure 20C, the spectra shape at 10 ºC in the presence of DNA illustrates a destabilization of the helitpicity of all variants. Complex thermal denaturation curves observed for all variants are indicative of multiple transitions in secondary structure as a result of solution heating. These data, together with an overall reduction of signal monitored by an increase of the photomultiplicator (PMT) voltage, suggest

that the presence of DNA induced the precipitation of all Omomyc and Max* derived variants. Hence, considering that all His-Tag and CPP-tag bearing variants seem to display significant precipitation in the presence of DNA, it was unclear at this stage whether the precipitation was induced by either one of the tags or both of them. Of course, non-specific interactions between positively charged CPP-tags and negatively charged DNA could explain a certain aggregation leading to precipitation of the complexes. However, the role of the His-tag present on the N-terminus of each variant could not be discarded either, since histidine is known to play an important role in protein-DNA interactions[253]. As a consequence, the characterization of the Max*-derived variants was deprioritized at this point, and they were therefore not further studied, while we decided to continue only with the characterization of the Omomyc variants.

In order to individually characterize the effects of the His-tag on the Omomyc structure and on its interaction with DNA, we generated the His-Tag-Omomyc construct (or H8-Omo), an Omomyc derived fusion protein harboring an 8xHis-Tag at its N-terminus. Comparing H8-Omo to the Omomyc-derived variants would also allow us to clarify the individual role of each tag during variant-DNA aggregation events.



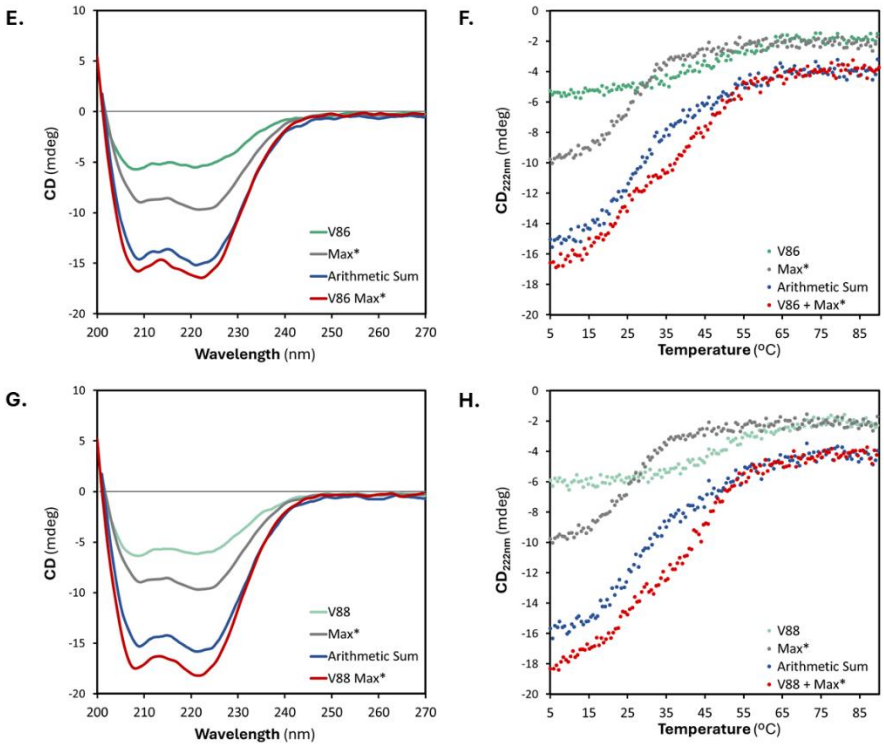


***Figure 19**, (**A**) Overlay of far-UV CD spectra and (**B**) thermal denaturation curves of Omomyc (3 μM, black), Max* (8 μM, grey), a mix of Omomyc (3 μM) and Max* (8 μM, red) and the arithmetic sum of the black and grey curves (blue). (**C**) Overlay of far-UV CD spectra and (**D**) thermal denaturation curves of V84 (3 μM, dark green), Max* (8 μM, grey), a mix of V84 (3 μM) and Max* (8 μM, red) and the arithmetic sum of the black and grey curves (blue). (**E**) Overlay of far-UV CD spectra and (**F**) thermal denaturation curves of V86 (3 μM, green), Max* (8 μM, grey), a mix of V86 (3 μM) and Max* (8 μM, red) and the arithmetic sum of the black and grey curves (blue). (**G**) Overlay of far-UV CD spectra and (**H**) thermal denaturation curves of V88 (3 μM, light green), Max* (8 μM, grey), a mix of V88 (3 μM) and Max* (8 μM, red) and the arithmetic sum of the black and grey curves (blue) in the presence of 50 mM potassium phosphate, 50 mM potassium chloride, 2 mM TCEP, pH 6.7.*

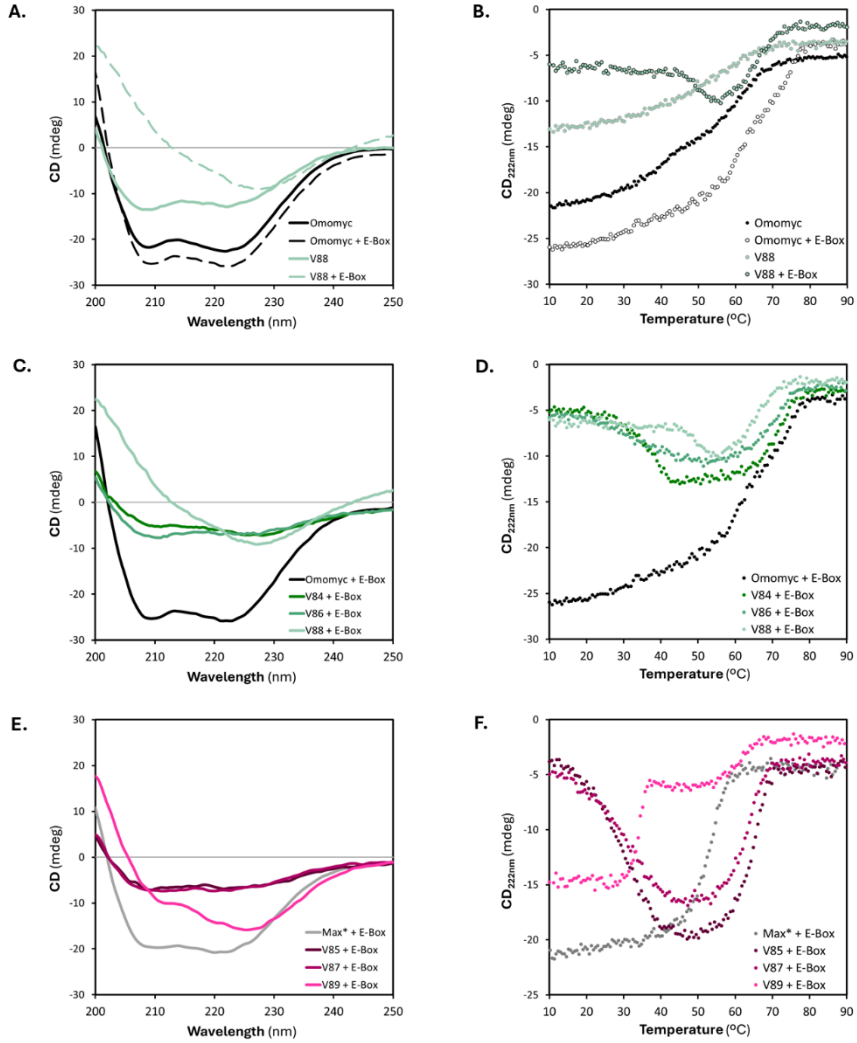


***Figure 20**, Overlay of (**A**) far-UV CD spectra and (**B**) thermal denaturation curves of Omomyc (black), Omomyc in presence of an E-Box containing DNA sequence (black dashed line / white rounded dot), V88 (light green) and V88 in presence of an E-Box containing DNA sequence (green dashed line / green rounded dots). Overlay of (**C**) far-UV CD spectra and (**D**) thermal denaturation curves of Omomyc (black) and the Omomyc-derived variants V84 (dark green), V86 (green), V88 (light green) in presence of an E-Box containing DNA sequence. Overlay of (**E**) far-UV CD spectra and (**F**) thermal denaturation curves of Max* (grey) and the Max*-derived variants V85 (purple), V87 (magenta), V89 (pink) in presence of an E-Box containing DNA sequence (8 µM). All experiments were carried out in CD buffer (50 mM KCl, 2 mM TCEP, 50 mM potassium phosphate, pH 6,7), in presence of 16 µM Omomyc or Max*-derived variants and 8 µM of DNA probe, when applicable.*

1.4. Both the HisTag and CPPs tag impair the variants DNA binding ability

Thanks to our previous experience with the other variants, the expression and purification of H8-Omo was straightforward and about 100 mg of pure protein could be obtained at the first attempt, following the same expression and purification protocol that were used for the obtention of the V84 to V89 variants.

By CD, almost identical spectra and thermal denaturation curves were obtained for both H8-Omo and Omomyc (Figure 21A-B). Our results show that the His-tag in the H8-Omo protein has little to no effect on the secondary structure of the rest of the construct or its thermal stability. They also indicate that the HisTag in H8-Omo has a similar propensity to form α-helices as the Omomyc sequence. Moreover, the Max*-corresponding transition disappearance from the H8-Omo + Max* denaturation curve presented in Figure 21C suggests that H8-Omo retains the ability to bind the Max b-HLH-LZ, as expected.

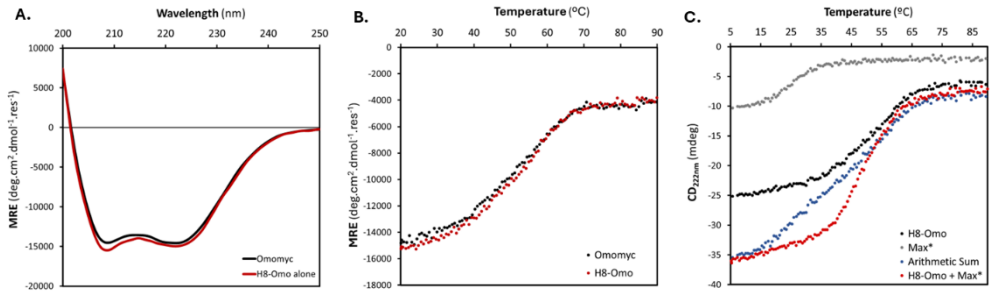


***Figure 21**, **A**. Overlay of the far-UV CD spectra of H8-Omo (red) and Omomyc (black), at 20 ºC; **B**. Overlay of the thermal denaturation curves of H8-Omo (red) and Omomyc (black) followed by CD at 222 nm. **C**. Overlay of the thermal denaturation curves of H8-Omo (black), Max* (gray), the arithmetic sum of the black and gray curves (blue), the experimental sum of H8-Omo and Max* (red) followed by CD at 222 nm. All the data sets presented in this Figure were obtained in CD1x buffer, in presence of 16 μM H8-Omo, 16 μM Omomyc, and 8 μM Max* when applicable.*

Perhaps surprisingly, though, a significant loss of ellipticity was observed at 10 ºC when H8-Omo was placed in the presence of either an E-Box-containing double stranded DNA sequence (referred to as E-Box DNA probe) or a non-specific DNA sequence probe, in the usual CD buffer conditions (or CD1x: 50 mM potassium phosphate, 50 mM potassium chloride, 2 mM TCEP, pH 6.7). In fact, thermal denaturation of H8-Omo in the presence of either DNA seems to depict a certain refolding occurring in the first place,

reaching a minimal CD_{222nm} value at 40 ºC in both cases. Interestingly, this putative refolding seems deeper in the presence of a canonical E-Box sequence than in the presence of the non-specific sequence (Figure 22B). These unusual denaturation curves, contrasting with the results typically obtained with Omomyc, share a strong resemblance to the denaturation curves obtained in the same conditions with the Omomyc variants V84, V86 and V88, depicted in Figure 20D. These observations led us to believe that the His-tag was at least partially responsible for the complex structural events that were observed with these variants in the presence of DNA, and that the experimental conditions used so far in all CD experiments would not be adequate to study the effect of Omomyc sequence mutations on DNA binding. Efforts were then dedicated to screening for conditions that could prevent non-specific HisTag / DNA interactions.

Since elevated temperatures seemed to allow for a certain refolding, the impact of the starting temperature on secondary structure was explored, as well as the effect of a short pre-incubation at 35 ºC. In the results depicted in Figure 22C, samples of H8-Omo in the presence of DNA were first cooled to 20 ºC, instead of 10 ºC as done previously. Incubation of 5 min at 35ºC led to a deepening of the 208 nm and 222 nm minima. Spectrum recorded after the subsequent cooling to 20 ºC showed a similar behavior as at 35 ºC, revealing a higher α-helix content than during the first read at 20 ºC. The thermal denaturation curve obtained from this sample showed decreasing CD values at 222 nm from 20 ºC to 40 ºC, reminiscent of the refolding observed in the previous experiment, but occurring to a lesser magnitude. Above 40 ºC, the denaturation curves of both Omomyc and H8-Omo in the presence of the E-Box DNA probe were found to be quite similar to each other, showing in both cases a more cooperative transition with a Tm higher than the one observed for each respective protein alone. It could also be noted that, in the presence of DNA, both proteins show an apparent lower residual ellipticity than when they are denatured in the absence of DNA. This result is likely to be explained by the tendency of Omomyc and H8-Omo to precipitate when being thermally denatured in the presence of DNA, since large amorphous aggregates are visible in the samples cuvette after denaturation.

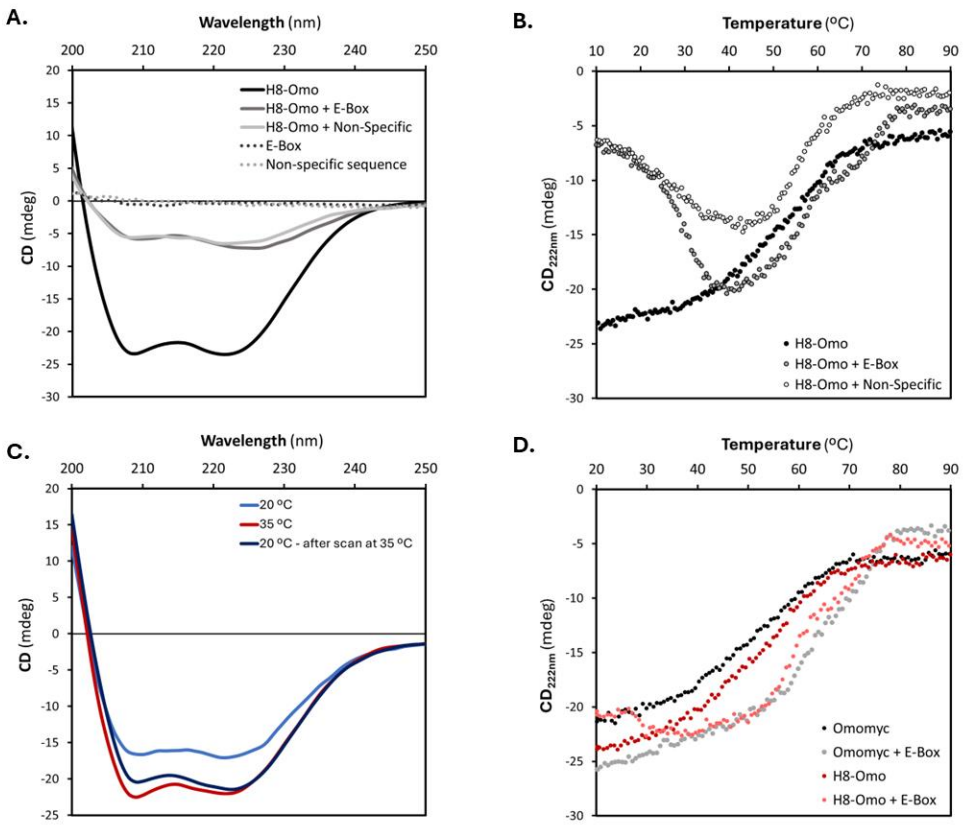
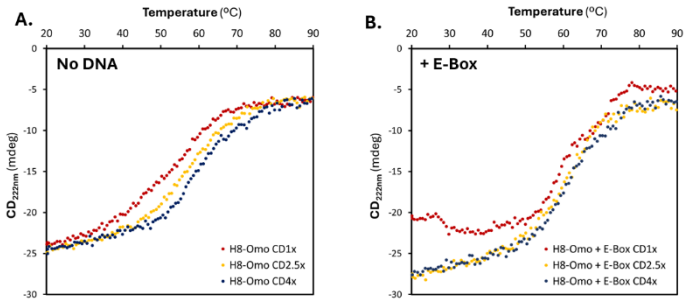


***Figure 22, A**. Overlay of the far-UV CD spectra of H8-Omo (black continuous line), H8-Omo in presence of an E-Box-containing DNA sequence (dark gray continuous line), H8-Omo in presence of a non-specific DNA sequence (light gray continuous line), the E-Box containing DNA sequence (dark gray dotted line) and a non-specific DNA sequence (light gray dotted line). All those spectra were recorded at 10 °C, **B**. Overlay of the thermal denaturation curves of H8-Omo (black dots), H8-Omo in presence of an E-box-containing DNA sequence (circled dark gray dots) and H8-Omo in presence of a non-specific DNA sequence (circled light gray dots). **C**. Overlay of far-UV CD spectra of the same sample of H8-Omo in presence of the E-Box DNA probe at 20 °C before incubation at 35 °C (light blue), after 5 min at 35 °C (red) and at 20 °C after the incubation at 35 °C (dark blue), each sample were maintained 5 min at the specified temperature before recording. **D**. Overlay of thermal denaturation curves monitored by CD at 222 nm of Omomyc (black), Omomyc in presence of the E-Box DNA probe (grey), H8-Omo (red) and H8-Omo in presence of the E-Box DNA probe. All results presented in this Figure were obtained in CD1x buffer, in presence of 16 µM of Omomyc or H8-Omo, 8 µM of DNA probe, when applicable.*

Since those experiments were carried out at a pH close enough to the histidine side chain theoretical pKa (~ 6.8)[254], the H8-Omo His-tag would be expected to be positively charged in the chosen experimental conditions. To test whether unspecific electrostatic interactions between the His-tag and DNA could be at the origin of the structural events observed below 40 ºC, we studied the effect of increased potassium phosphate and potassium chloride concentrations upon H8-Omo DNA binding.

In the next experimental set, H8-Omo thermal stability was studied in two additional buffer conditions: in the presence of 125 mM potassium phosphate and 125 mM potassium chloride (or CD2.5x) and in the presence of 200 mM potassium phosphate, 200 mM potassium chloride (or CD4x) (Figure 23). A pH of 6.7 and 2 mM of TCEP to ensure a reductive environment were maintained in all three buffer conditions. At 20 ºC, H8-Omo ellipticity was found to be unaffected by the increases in phosphate and potassium chloride concentration. In contrast, its thermal stability seemed to correlate with the CD buffer concentration. In fact, compared to the Tm obtained for H8-Omo in CD1x, the protein Tm was found to be 4.3 ºC higher in CD2.5x and 5.3 ºC in CD4x. In the presence of the E-Box DNA probe, the H8-Omo refolding events observable in CD1x below 40 ºC was no longer visible at increased salts concentrations. Interestingly, the thermal denaturation curves of H8-Omo in the presence of DNA in CD2.5x and CD4x were found to be almost identical (Figure 22B).



C.

Melting temperature (ºC)	No DNA	+ E-Box
CD1x	53.3	n.d.*
CD2.5x	57.6	61.7
CD4x	58.7	62.0

***Figure 23**, Overlay of the thermal denaturation curves of H8-Omo **A.** alone and **B.** in the presence of the E-Box DNA probe in different buffer concentration: 50 mM potassium phosphate, 50 mM potassium chloride, 2 mM TCEP, pH 6,7 (or CD1x, red); 125 mM potassium phosphate,125 mM potassium chloride, 2 mM TCEP, pH 6,7 (or CD2.5x, yellow); 200 mM potassium phosphate, 200 mM potassium chloride, 2 mM TCEP, pH 6,7 (or CD4x, blue). All the data sets presented in this Figure were recorded in presence of 16 µM of H8-Omo and 8 µM of DNA probe when applicable. **C.** Table summarizing the melting temperature of H8-Omo in the different presented conditions. *This*

Tm could not be determined due to an incapacity to achieve a satisfactory fitting of the corresponding thermal denaturation curve.

Furthermore, a strong resemblance was found between the denaturation curves of Omomyc and H8-Omo in CD2.5x either in the absence or in the presence of the E-Box DNA probe, displaying transitions of very similar cooperativity and Tm. In contrast, when exposed to the non-specific DNA probe, H8-Omo exhibited a thermal denaturation that was slightly more cooperative and had a lower Tm compared to Omomyc (Figure 24).

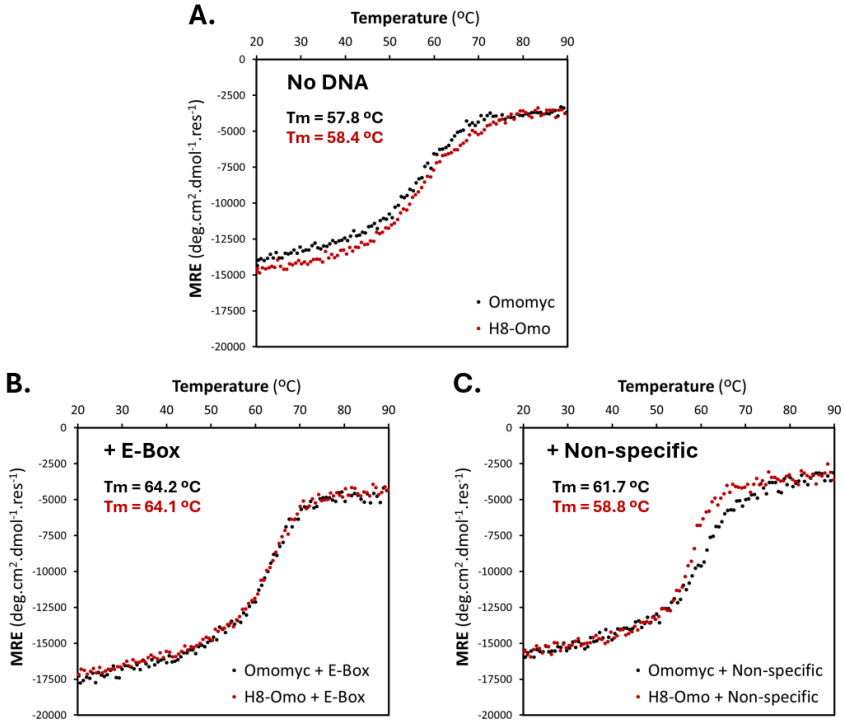


***Figure 24**, Overlay of the thermal denaturation curves followed by CD at 222 nm of Omomyc (black) and H8-Omo (red) in: **A**. the absence of DNA; **B**. the presence of the E-Box DNA probe; **C**. the presence of the non-specific DNA probe. All the data sets presented in this Figure were recorded in CD2.5x buffer, and in the presence of 16 µM of Omomyc or H8-Omo and 12 µM of DNA probe when applicable.*

While this high phosphates / high salt concentrations conditions seemed sufficient to prevent most of the His-tag non-specific interactions with DNA, new characterization of the Omomyc variants V84 and V86 in CD2.5x buffer in the presence of the E-Box DNA probe showed similar CD spectra at 20 ºC than the ones obtained previously in CD1x buffer (Figure 25), suggesting that both the SynB3 and the Penetratin tag interfere with the specific union of the

Omomyc basic region and DNA. The secondary structure of the two variants was then followed through successive far-UV CD spectra recording at increasing temperatures between 20 ºC to 55 ºC degrees, with one scan carried out every 5 ºC after a 2-min temperature stabilization. In the case of V84, a negative band at 227 nm is visible at 20 ºC, violet-shifting as the temperature increases, reaching 222 nm at 40 ºC. Concurrently, a negative band at 208 nm becomes apparent at the same temperature intervals, reaching its minimum (or maximum intensity) at 40 ºC, forming a typical α-helix spectrum together with the band at 222 nm. Beyond 45 ºC, the spectra of V84 indicate a progressive decline in α-helix content alongside an increase in random coil content, as evidenced by a reduction of CD signal at both 208 nm and 222 nm, a decreasing CD_{222nm}/CD_{208nm} ratio, and decreasing CD values at 200 nm. Although displaying a seemingly higher α-helix content than V84 at 20 ºC, V86 thermal denaturation resulted in a spectral shift between 20 ºC and 40 ºC similar to the shift previously described for V84 (Figure 25C-D).

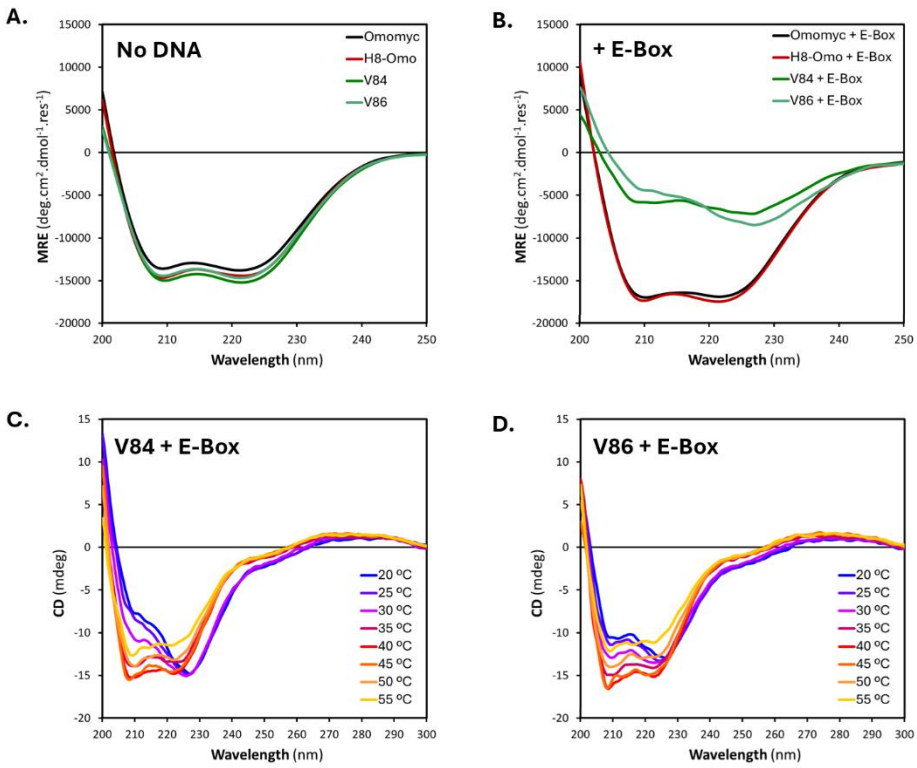


***Figure 25, A-B**. Overlay of the far-UV CD spectra of Omomyc (black), H8-Omo (red), V84 (dark green) and V86 (green): **A**. in the absence of DNA, and **B**. in the presence of the E-Box DNA probe at 20 ºC. **C-D**. Overlay of the far-UV CD spectra of V84 (**C**) and V86 (**D**) in the presence of the E-Box*

DNA probe recorded at increasing temperatures from 20 ºC (blue) to 55 ºC (yellow). All the data sets presented in this Figure were recorded in CD2.5x buffer, and in the presence of 16 µM of Omomyc, H8-Omo, V84 or V86 and 12 µM of DNA probe when applicable.

Since all previous results suggested that incubation at elevated temperature may help to limit non-specific tags-DNA interactions, the effect of a long incubation at 37 ºC of V84/V86-DNA samples prior to analysis was studied. Figure 26A highlights the effect of such incubation on the secondary structure of those variants in the presence of the E-Box DNA probe. At 25 ºC, samples that were incubated at 37 ºC prior to their analysis (referred to as incubated samples) showed a much higher helical content than the samples that were left incubating at room temperature (referred as non-incubated samples). In those conditions, the thermal denaturation curves of V84 and V86 in the absence of DNA showed similar mean residual ellipticity (MRE) at 222 nm and at 25 ºC than H8-Omo and seemingly identical Tm. In the presence of the E-Box DNA probe, V84 and V86 showed slightly lower MRE at 222 nm at 25 ºC compared to H8-Omo. However, they displayed very similar Tm values and exhibited no apparent refolding between 25 ºC and 40 ºC. Those results show that these specific experimental conditions enabled to study the DNA-binding abilities of V84 and V86, while preventing most of complex structural events caused by the His-tag, SynB3 tag, and Penetratin tag non-specific DNA interactions. In other words, considering that in those conditions H8-Omo, V84 and V86 gave very similar results in the presence of DNA, the tags effect can be neglected compared to the importance of the rest of the sequence. This experimental process, summarized in Figure 26D, can therefore be used to study the effects of the short stretch mutations introduced within the Omomyc sequence to design the variants V93 to V107.

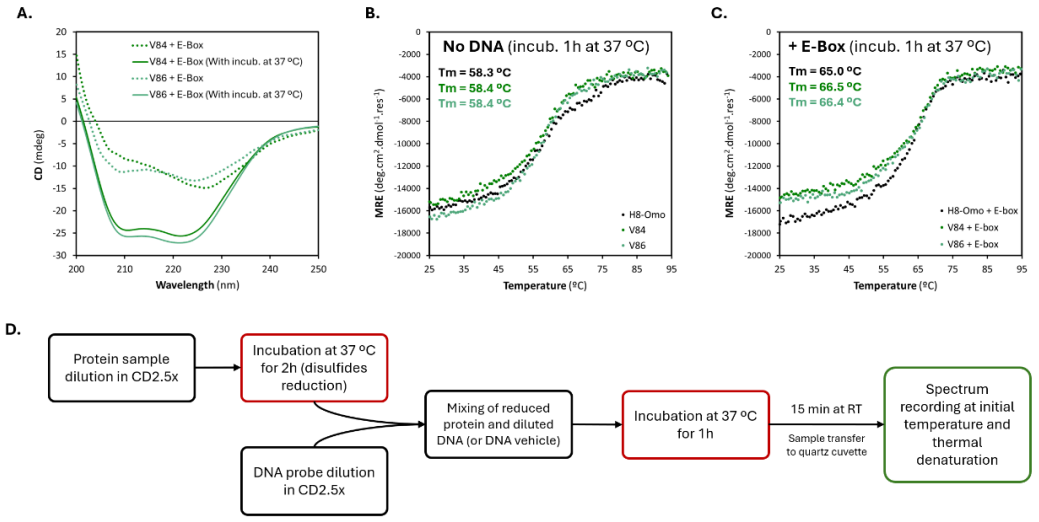


***Figure 26, A**. Overlay of the far-UV CD spectra at 25 °C of V84 (dark green) and V86 (green) in the presence of the E-Box DNA probe, without prior incubation at elevated temperature (dotted lines); and analyzed after incubation of 1h at 37 °C (continuous lines). **B-C**: Overlay of the thermal denaturation curves of H8-Omo (black), V84 (dark green) and V86 (green): **B**. in the absence of DNA, and **C**. in the presence of the E-Box DNA probe, after incubation for 1h at 37 °C. All data set presented in this Figure panel were recorded in CD2.5x buffer, in the presence of 16 µM of H8-Omo, V84 or V86 and 12 µM of DNA probe when applicable. **D**. Workflow chart of the optimized sample preparation for CD analysis of protein-DNA interaction.*

1.5. The R9 tag and Omomyc LZ truncation affect solubility and stability

All tested *E. coli* colonies transformed with the variants V93- or V94-expressing plasmids displayed poor expression levels. This significantly limited the amount of each protein that could be recovered after the purification capture step. Indeed, for both constructs, the recovered protein amounts after capture were too little to consider further characterization. For this reason, and since poor expression often correlates with poor stability, these variants were discontinued and not further studied. Interestingly, colonies expressing V95, displaying the same truncated Omomyc moiety as V93 and V94, showed instead good expression levels, and satisfactory expression levels were obtained for the rest of the variant set too (Figure 27).

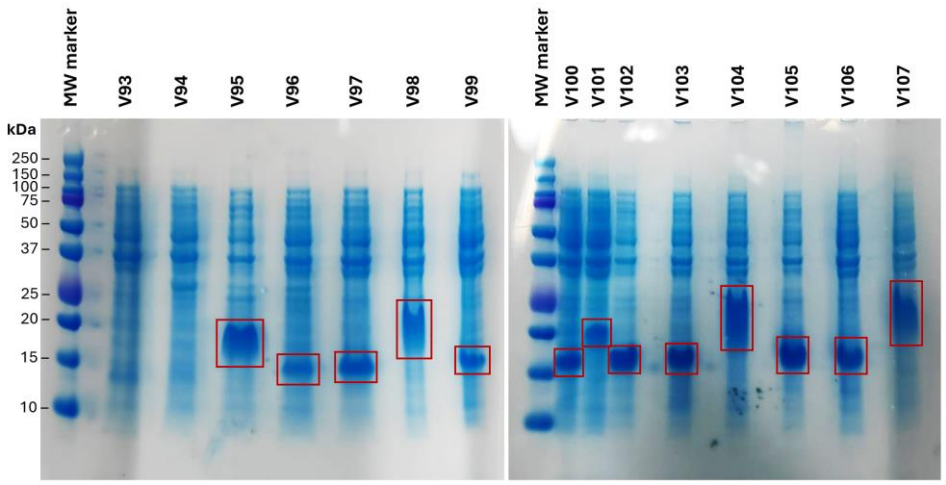


***Figure 27**, sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) of full protein extracts of transformed E. coli cells, collected and lysed after induction of the expression of each respective variant (V93 to V107). The bands expected to correspond to the protein of interest in each lane are circled in red. Both gels presented in this Figure were run and stained in parallel and in the same conditions.*

When ran on an SDS-PAGE, the R9-displaying variants consistently appeared as long bands, migrating to distances corresponding to significantly higher molecular weight than their respective theoretical size spanning from 12 to 15 kDa. This unusual migration pattern was particularly observed when concentrated solutions were loaded for analysis, and the length of the band was proportional to the loaded amount of R9-variant.

The R9-variant family showed an overall reduced solubility while compared to their respective SynB3 or Penetratin-conjugated counterparts. This low solubility was observed at both pH 8, since a long peak tail was observed during their elution after capture, even in presence of 4 M urea, and at acid pH, in potassium phosphate-based storage buffer in which most R9-variants could not be concentrated to satisfactory levels. Similar observations were made for the variant V106 (H8-Penet-OmoNLS4), which displayed a limiting solubility of about 25 µM in storage buffer. For this reason, all R9-conjugated variants and V106 were excluded from subsequent studies (the status of all the variants at this point of the project is recapitulated in Table 2).

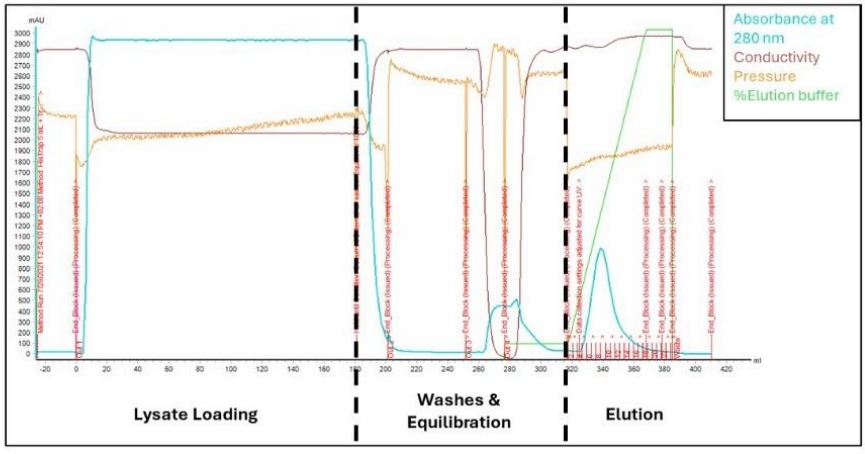
Using the previously described purification process, variants V96, V97, V99, V100, V102, V103 and V105 could be successfully purified, buffer exchanged and concentrated to concentrations over 100 µM. Since signs of

degradation were observed in both V84 and V86 variants stocks, new batches of both variants were produced at the same time. A representative example of key V99 purification and quality control steps adopted are represented in Figure 28.

***Table 2**, Status of each Omomyc-derived variants after purification and concentration attempts.*

Name	Description	Status
H8-Omo	8xHis-Omomyc	Purified / concentrated
V84	8xHis-SynB3-Omomyc	Purified / concentrated
V86	8xHis-Penetratin-Omomyc	Purified / concentrated
V88	8xHis-R9-Omomyc	**Discontinued**
V93	8xHis-SynB3-T3Omo	**Discontinued** – Very poor expression levels
V94	8xHis-Penetratin-T3Omo	**Discontinued** – Very poor expression levels
V95	8xHis-R9-T3Omo	**Discontinued** – Poor solubility
V96	8xHis-SynB3-OmoCrot	Purified / concentrated
V97	8xHis-Penetratin-OmoCrot	Purified / concentrated
V98	8xHis-R9-OmoCrot	**Discontinued** – Poor solubility
V99	8xHis-SynB3-OmoEGFR	Purified / concentrated
V100	8xHis-Penetratin-OmoEGFR	Purified / concentrated
V101	8xHis-R9-OmoEGFR	**Discontinued** – Poor solubility
V102	8xHis-SynB3-OmoNLS2	Purified / concentrated
V103	8xHis-Penetratin-OmoNLS2	Purified / concentrated
V104	8xHis-R9-OmoNLS2	**Discontinued** – Poor solubility
V105	8xHis-SynB3-OmoNLS4	Purified / concentrated
V106	8xHis-Penetratin-OmoNLS4	**Discontinued** – Poor solubility
V107	8xHis-R9-OmoNLS4	**Discontinued** – Poor solubility

A.



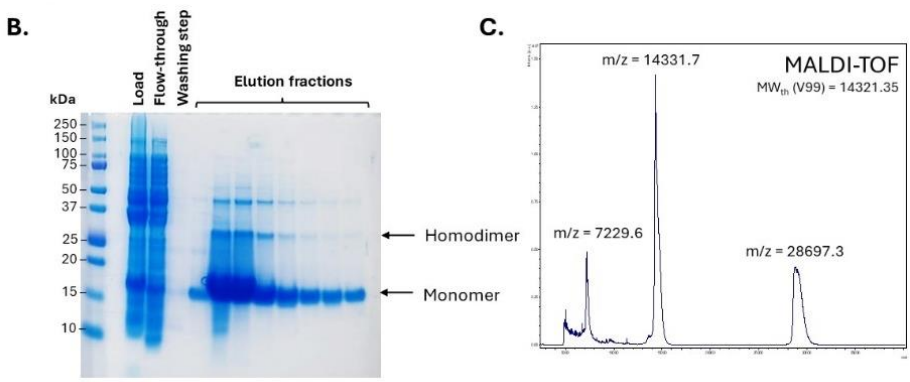


***Figure 28**, **A**. Capture step chromatogram of variant V99 (absorbance at 280 nm: blue, conductivity: brown, pressure: orange, proportion of elution in buffer in mobile phase: green); **B**. SDS-PAGE in reducing conditions of the V99 capture step chromatography main fractions; **C**. Matrix-Assisted Laser Desorption/Ionization mass spectrometry (MALDI-TOF) results of the V99 pure fractions after buffer exchange, concentration and final filtration.*

1.6. Short stretch mutations may strongly affect Omomyc structure and its ability to bind DNA and MAX

CD was used to study the effect of the various mutations on the CPP-fused variants secondary structure. Far-UV CD spectra recorded at 25 ºC showed that all variants had reduced ellipticity compared to their respective parent peptides, V84 or V86. More in detail, despite suggesting a certain α-helix content, the spectra of the loop-mutated variants show a particularly higher random coil content, highlighted by a violet shift of their respective 208 nm negative band to 206 nm and 205 nm for V102/V103 and V96/V97, respectively.

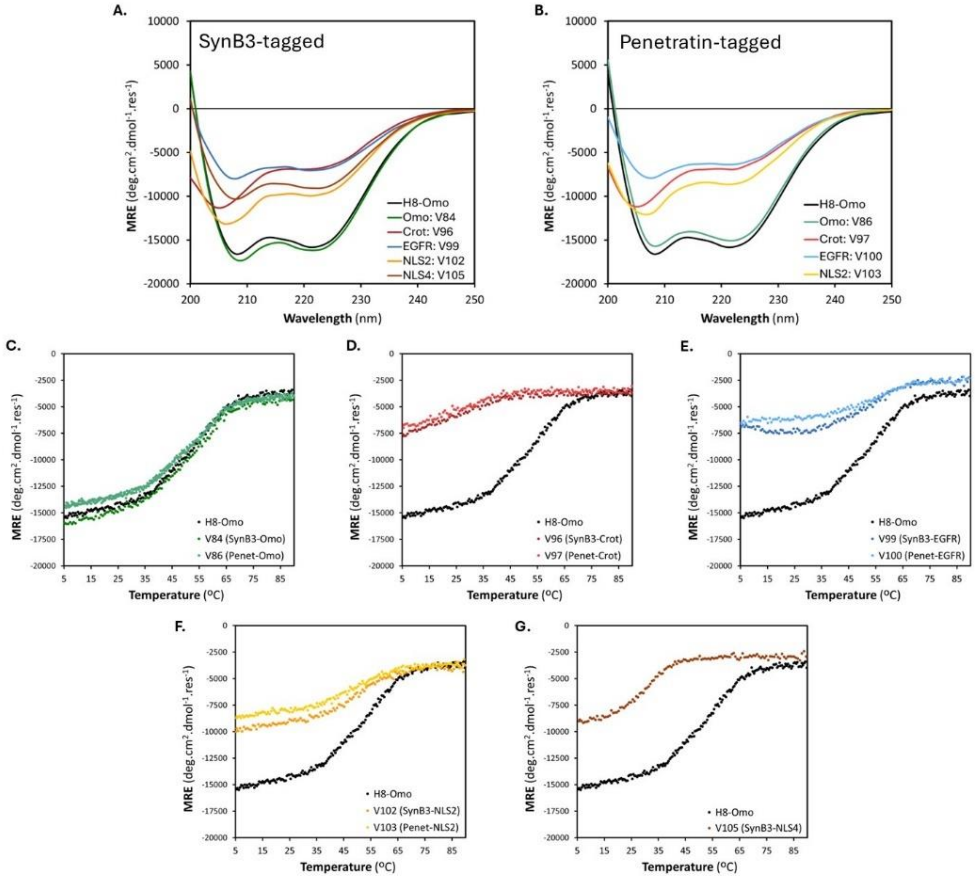


***Figure 29**, **A-B**: Far-UV CD spectra at 25 ºC of (**A**) the SynB3-tagged variants and (**B**) Penetratin-tagged variants with H8-Omo. **C-G**: Overlay of the thermal denaturation curves, followed by CD at 222 nm of (**C**) V84 and V86, (**D**) the "Crot" variants: V96 and V97, (**E**) the "EGFR-targeting": variants V99 and V100, (**F**) the "NLS2" variants: V102 and V103, and (**G**) the "NLS4" variant: V105. All data set presented in this Figure were obtained in presence of 16 μM of each respective construct diluted in CD1x buffer.*

In this dataset, the thermal denaturation of V84 and V86 were apparently composed of two overlapping transitions with Tm of about 42ºC for the first one, and 59 ºC for the second. This two-transitions thermal denaturation was also visible in the H8-Omo curve, yet they were too overlapping for accurate independent Tm determination and were therefore considered as a single one. The presence of two transitions could suggest the independent denaturation of the basic region, followed by denaturation of the LZ at higher temperatures.

Thermal denaturation analysis of the variant set by CD at 222 nm highlighted a significant reduction of Tm in variants V96, V97 and V105, while the other variants conserved a Tm very similar to the one observed for their parent peptide (Figure 29, Table 3). The OmoCrot-derived variants displayed a particularly reduced Tm (lower than 25 ºC in both cases), which is probably caused by their LZ truncation, considering that OmoNLS2 variants, that also display a basic amino acid-rich loop region, conserved a Tm similar to H8-Omo.

Once again, pairs of identical sequences fused to either the SynB3 or Penetratin CPP tag exhibited remarkably similar secondary structure content and stability. These observations further reinforce the hypothesis that both tags have little effect over the structure of Omomyc-derived variants.

***Table 3**, Descriptive name, mutation position, and melting temperature in CD1x for each Omomyc-derived variant characterized. *The lack of lower plateau at 5 ºC did not enable an accurate determination of the melting temperature for these variants.*

Name	Description	Mutation Position	Tm in CD1x
H8-Omo	8xHis-Omomyc	-	53.3 ºC
V84	8xHis-SynB3-Omomyc	-	42.8 ºC / 59.7 ºC
V86	8xHis-Penetratin-Omomyc	-	41.9 ºC / 58.5 ºC
V96	8xHis-SynB3-OmoCrot	Loop substitution / C-terminus truncation	*<25 ºC**
V97	8xHis-Penetratin-OmoCrot	Loop substitution / C-terminus truncation	*<25 ºC**
V99	8xHis-SynB3-OmoEGFR	C-terminus end substitution / elongation	53.1 ºC
V100	8xHis-Penetratin-OmoEGFR	C-terminus end substitution / elongation	55.2 ºC
V102	8xHis-SynB3-OmoNLS2	Loop substitution	51.7 ºC
V103	8xHis-Penetratin-OmoNLS2	Loop substitution	50.5 ºC
V105	8xHis-SynB3-OmoNLS4	C-terminus end substitution	31.5 ºC

The next experimental set focused on determining whether those variants conserved the ability to bind the Max b-HLH-LZ region, following the same experimental strategy previously used for V84, V86 and H8-Omo characterization (Figure 30). Among all the studied variants, the EGFR-

targeting variants exhibited the most stabilizing effect on the Max* secondary structure: comparison between the arithmetic and experimental sum curves shows a complete shift of the Max-corresponding transition.

Results obtained with the NLS2 variants suggest incomplete heterodimerization, as the Max-corresponding transition remains visible in the experimental sum curves obtained with these variants. Another possible explanation is that the heterodimer formation has only a minimal stabilizing effect on the Max* melting temperature.

The interaction between V105 and Max* appears to result in a slight increase in Max* stability.

The arithmetic and experimental sums obtained for each respective "Crot" variant largely overlapped, indicating that these variants have likely lost their ability to bind the Max* protein efficiently.

Finally, the binding capacity of these variants to the E-Box and a non-specific DNA probe was studied under optimized conditions previously determined with H8-Omo (Figure 31). The behavior of the variants in the presence of DNA appeared to be influenced differently depending on the site of their respective mutations. As described earlier, binding of variants V84 and V86 to the E-Box DNA probe has a stabilizing effect on their melting temperature. Interestingly, interactions with the non-specific DNA probe seem to have the opposite effect, destabilizing the melting temperature and secondary structure of both variants at room temperature. Loop-mutated variants were observed to precipitate heavily in the presence of both DNA probes. Similarly, aggregation of the C-terminus-mutated variants occurred with both DNA probes, but the presence of the canonical E-Box sequence seemed to allow partial refolding, whereas no such refolding was observed with the non-specific probe. In the presence of the E-Box, an initial transition from 40 ºC to 50 ºC toward more negative CD values at 222 nm was observed, followed by a second transition to a denatured state ending at about 60 ºC for V105 and 65 ºC for V99 and V100.

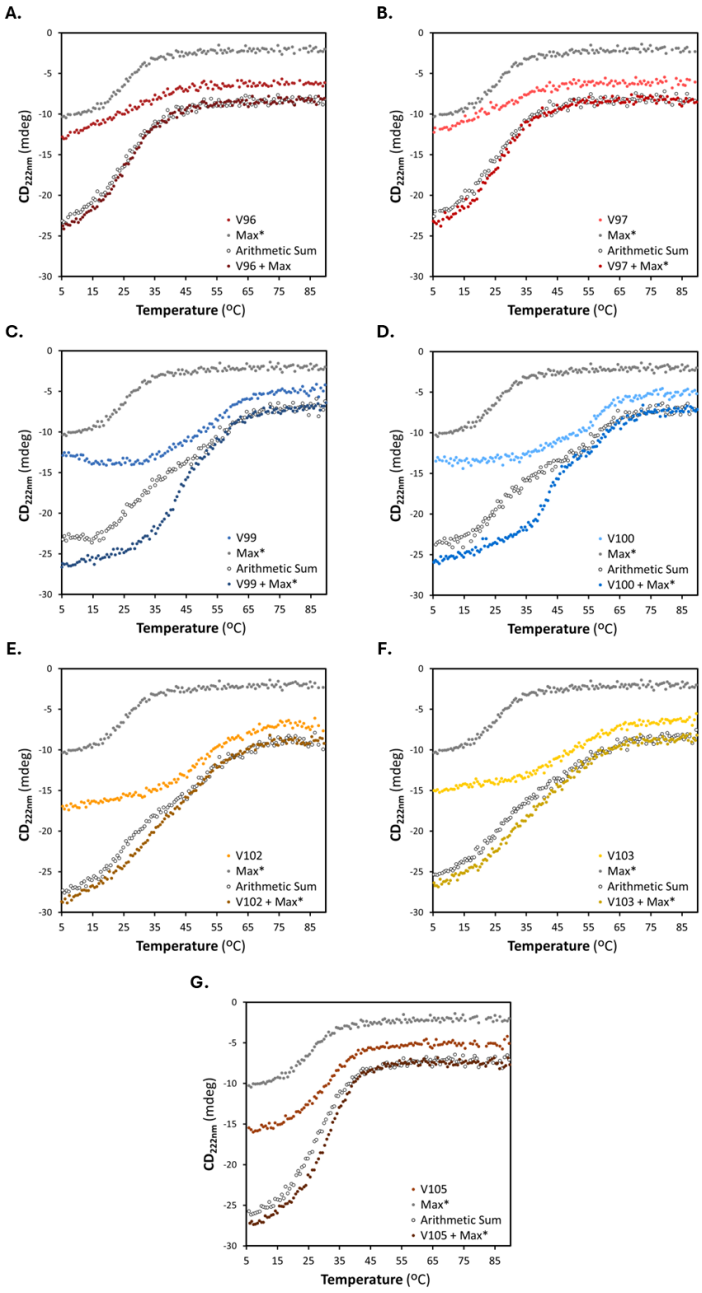


***Figure 30**, Overlays of the thermal denaturation of Max* (grey) with (**A**) V96 (red) and the experimental mix of V96 and Max* (dark red), (**B**) V97 (light red) and the experimental mix of V97 and Max* (red), (**C**) V99 (blue) and the experimental mix of V99 and Max* (navy blue), (**D**) V100 (light blue) and the experimental mix of V100 and Max* (electric blue), (**E**) V102 (gold yellow) and the experimental mix of V102 and Max* (orange), (**F**) V103 (yellow) and the experimental mix of V103 and Max* (dark yellow), (**G**) V105 (brown) and the experimental mix of V105 and Max* (dark brown),*

together with the arithmetic sum of the Max and each respective variant thermal denaturation curve (black-circled white dot). All data set presented in this Figure were obtained in presence of 16 μM of variant, and 8 μM of Max* diluted in CD1x buffer.*

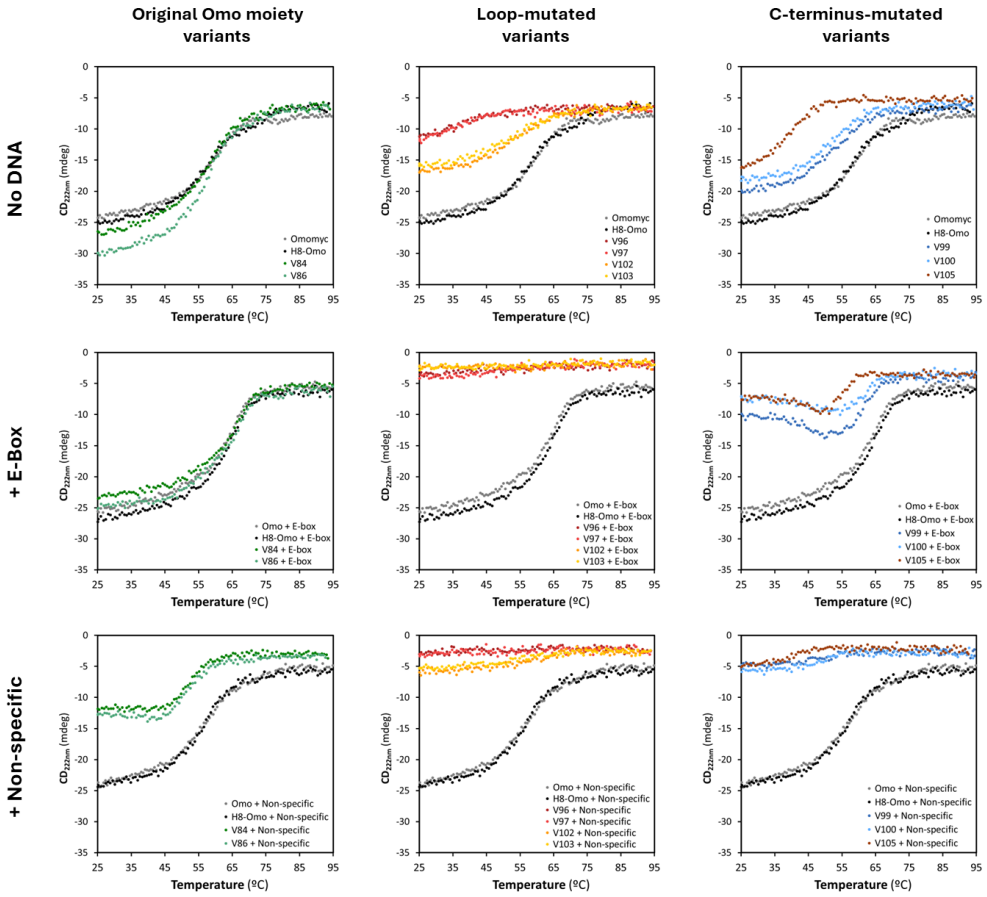


***Figure 31**, Thermal denaturation curves of Omomyc-derived variants in the presence or absence of DNA, overlaid with the curves obtained for Omomyc and H8-Omo under the same conditions to facilitate comparison. Top row, variants in the absence of DNA; Middle row, variants in the presence of the E-Box DNA probe; Bottom row, variants in the presence of the non-specific DNA probe; Left column, data sets obtained with the variants V84 (green) and V86 (light green); Center column, data sets obtained with the loop-mutated variants (V96: red, V97: light red, V102: gold yellow, V103: yellow); Right column data sets obtained with the C-terminus-mutated variants (V99: blue, V100: light blue, V105: brown). All data were obtained following the protocol summarized in Figure 26D, in presence of 16 μM of variant and 12 μM DNA probe, when applicable, in CD2.5x buffer.*

1.7. All variants display the capacity to penetrate cells through ATP-dependent routes and the CPP might affect their cell penetration kinetics

To study whether the cell penetration tags and mutations could affect the intracellular localization of the variants, anti-HisTag immunofluorescence was performed on A549 non-small cell lung cancer cells after treatment with 10 μM of Omomyc-derived variants for 4h. While results indicated that all variants could penetrate the cells, they were mainly located in the cytoplasm (Figure 32). Precipitation patches were also observed for all variants, indicating their limited solubility in cell culture media.

In any case, these results did not clarify whether the various mutations could in any way improve Omomyc intrinsic cell penetration properties. It was therefore decided to develop a flow cytometry-based assay to gain a more quantitative notion of the cell penetration abilities of each variant.

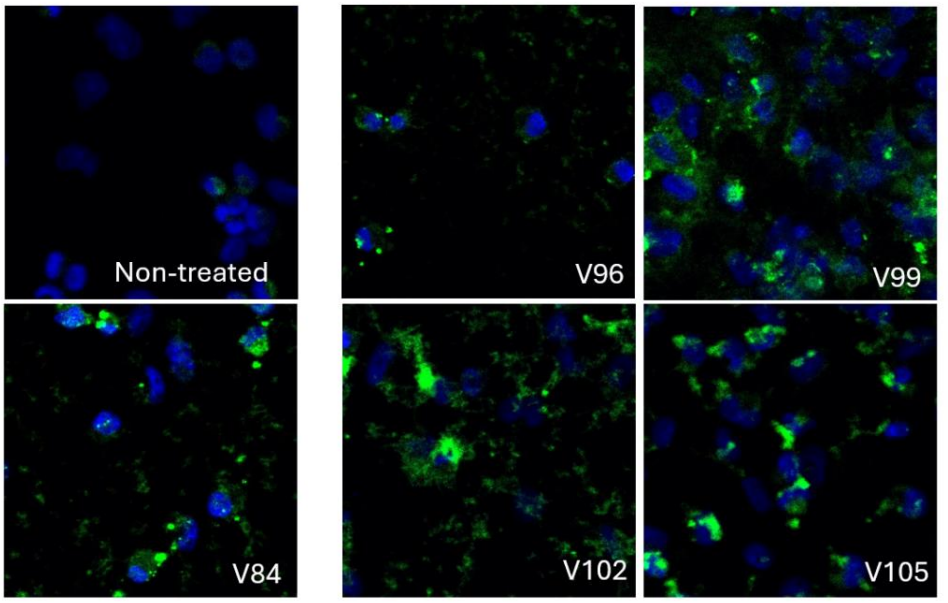


***Figure 32**, Immunofluorescence staining of A549 non-small cell lung cancer cells. Cells were treated with 10 μM of either V84, V96, V99 or V102 for 4h prior to slide preparation and staining. Cells nuclei was stained by DAPI (blue) and an AlexaFluor® 488-conjugated anti-HisTag antibody was used to label the Omomyc variants (green). Images were captured by confocal microscopy and obtained through the merge of both blue and green channels. Each image is a representative section of each respective slide.*

The first step in the assay development was to determine whether the His-Tag displayed by each variant could affect its cell penetration ability. Results obtained following the protocol described in Figure 33 suggested that H8-Omo exhibits similar, or even slight superior cell penetration capabilities compared to Omomyc after a 10 µM treatment for 15 minutes at 37 ºC. The detection of intracellular Omomyc and H8-Omo was characterized by an increase in the median fluorescence intensity per cell, which is proportional to the concentration of intracellular mini-protein. Additionally, a higher fluorescence heterogeneity was observed in the treated cells, resulting in broader histogram peaks compared to the non-treated cells (Figure 34A).

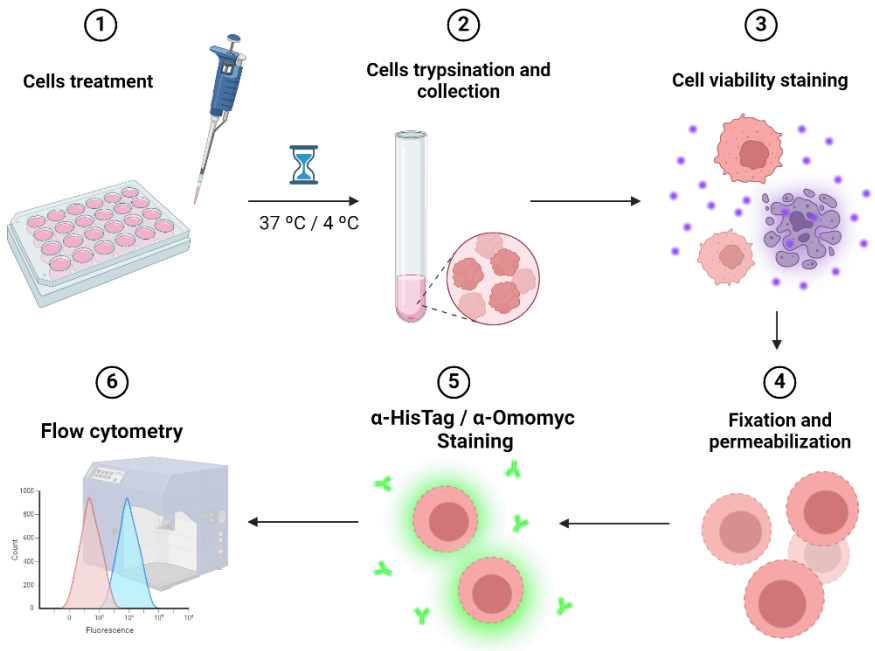


***Figure 33**, Schematic representation of the flow cytometry-based cell penetration assay protocol. Cells were initially cultured in a 24-well plate before being treated with 10 µM of an Omomyc variant. The cells were then harshly trypsinized to detach them from the plate and digest any protein adhered to the membrane. To filter out dead cells that could uptake variants passively, a viability dye was used for staining. Following this, the cells were fixed and their membranes permeabilized to allow the staining of intracellular variants by anti-HisTag or anti-Omomyc antibodies. After the final washing steps, the fluorescence distribution of each sample was measured by flow cytometry. In this experiment, the fluorescence intensity is therefore proportional to the concentration of the variant uptaken by the cells.*

No significant cell death was detected upon either of these two short treatments.

The cell penetration properties of the variants were then compared to H8-Omo and each other. To ensure comparable labelling across all variants, we opted to use an AlexaFluor® 488-labelled, anti-HisTag antibody. Despite some optimization attempts, the difference of signal between non-treated and H8-Omo obtained using anti-Omomyc antibody could not be reproduced using an anti-HisTag detection, probably illustrating a poorer selectivity of the anti-HisTag antibody when compared to the anti-Omomyc.

After a 15 min incubation at 37 ºC, V84 and V86 presented a slightly higher, yet not significant, fluorescence median than H8-Omo. Very similar results were obtained with the variants V96, V97, V102, V103 and V105.

The EGFR-targeting V99 and V100 variants displayed statistically significant increases in cell penetration, compared to H8-Omo. This increase seemed particularly pronounced for the variant V99 and was also found to be statistically significant when compared to results obtained with its parent peptide V84. Overall, the increase of fluorescence found in variants-treated cells compared to the non-treated group was much more profound when the cells were incubated at 37 ºC, compared to 4 ºC (Figure 34C-D). This difference suggests that the A549 cells variants uptake is done mainly through active and ATP-dependent processes. Interestingly, a significant increase of median fluorescence was observed for H8-Omo comparing the effect of 15 min-treatments to 1 h-treatment, while little to no increase was observed for variants V84 and V99. Hence, in this protein model, cell penetration tags may improve uptake kinetics, improving cell penetration of Omomyc-derived proteins upon short exposure, but they do not seem to make any difference at later time points (Figure 34E).

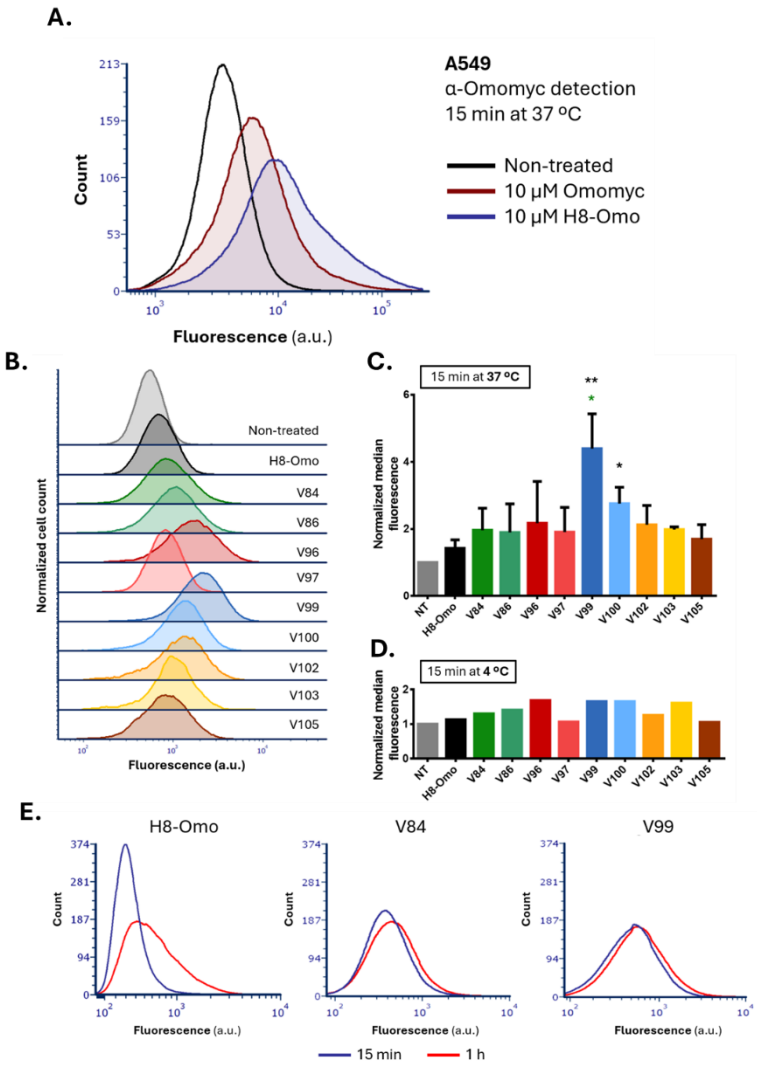


***Figure 34, A**. Overlay of 1D histograms obtained for A549 cell pools treated for 15 min with 10 µM Omomyc (red), 10 µM H8-Omo and non-treated cells (black). The x-axis represents the AlexaFluor® 488 fluorescence intensity (logarithmic scale), and the y-axis shows the cell count. Cells were processed following the protocol described in Figure 33 prior to analysis and 30,000 events (single cells) were analyzed per condition. A combination of a mouse anti-Omomyc primary antibody and a goat AlexaFluor® 488-labelled anti-mouse secondary antibody was used for Omomyc and H8-Omo detection. **B**. Overlay of representative 1D histograms obtained for A549 cell pools non-treated or treated with 10 µM of Omomyc-derived variants for 15 min, incubated at 37 °C. **C-D**: Quantification of the intra-cellular uptake of the variants by A549 cells after a 10 µM treatment for 15 min at (**C**) 37 °C (average median of biological duplicates are represented, the bars representing the standard deviation across replicate) and (**D**) 4 °C. Each median fluorescence intensity was normalized against the median fluorescence value obtained for the non-treated sample. (* p-value < 0.05; ** p-value < 0.01). **E**. Overlays of 1D histograms obtained for A549 cell pools treated with 10 µM of H8-Omo, V84 or V99 for 15 min or 1h at 37 °C.*

Section 2:

Impact of the net charge on the Omomyc-derived mini-proteins mode of action

2.1. Design of the constructs

As described in the introduction, a growing body of studies with monoclonal antibodies suggests that high net charge in physiological conditions favors their rapid clearance, whereas antibodies displaying a neutral net charge are prone to display better pharmacokinetics. The main objective of the work presented here is to study whether those observations, made in monoclonal antibodies, would be valid in a completely different protein model such as Omomyc.

The Omomyc mini-protein is characterized by a high isoelectric point of 9.8, originating from its high basic amino acid content, which represents more than 20% of its sequence. Approximately one tier of those residues is concentrated in the Omomyc BR, which confers to this mini-protein the ability to bind DNA with high affinity, to penetrate cells and reach their nuclei. 9 lysines, 4 arginines and one histidine residues are spread throughout the rest of the Omomyc sequence, helping to stabilize its secondary structure as well as allowing the formation of a homodimer. While the basic residues of MYC's BR are highly conserved throughout the PMN, only a few are well conserved in the Helix-2 and Leucine zipper regions, suggesting a higher sequence flexibility in terms of charge repartition in those regions.

6 Omomyc charge variants were designed through 3 different approaches (Table 4).

The first pair, namely V108 and V109, were designed after examination of the two Omomyc crystal structures available (PDB: 5I50, homodimer bound to DNA; 5I4Z, apo form[172]). First, two positively-charged residues were identified as non-interacting with any residues of the same monomer or the other chain in the homodimer. Hence, in the V108 variant, both residues were substituted with acidic amino acids. In the V109 variant, instead, the two basic amino acids were replaced by residues with uncharged, polar side chains, and two original polar amino acids with side chains pointing outward from the dimer were substituted with negatively charged residues.

The next variant pair sequences were obtained after alignment of the members of the PMN: basic residues that were not conserved among them were systematically substituted by polar amino acids in V110, and negatively charged amino acids in V111.

A similar strategy was used to obtain the last variant pair, V112 and V113, but substitutions were inspired solely by the MAX b-HLH-LZ sequence. With the primary objective of decreasing Omomyc's net charge, non-conserved basic amino acids were this time substituted with their respective counterparts in the MAX sequence. While only point mutations were made to design variants V108 to V112, a short stretch mutation was made to create the V113 sequence. Following their respective design strategy, only a single amino acid was mutated in the MYC basic region in V110 to V113, while the basic region of V108 and V109 remained unchanged. The residues that were seen to directly contribute to DNA binding were left untouched in all 6 sequences.

As summarized in Table 5, the design of these charge variant sequences allowed to lower Omomyc's pI (and, by extension, its net charge at physiological pH) to different extent and through different substitutions, both in terms of nature and position in the sequence.

***Table 4**, Aligned sequences of the Omomyc moiety of each charge variant. The substitutions in each variant are represented in bold, underlined and red characters.*

	Basic region	Helix1	Loop	Helix2	Leucine zipper
Omomyc	MTEENVKRRTHNVLERQRRN	ELKRSFFALRDQI	PELENNEKAP	KVVILKKATAYILSVQ	AETQKLISEIDLLRKQNEQLKHKLEQLRNSCA
V108	MTEENVKRRTHNVLERQRRN	ELKRSFFALRDQI	PELENNEKAP	KVVILKKATAYILSVQ	AETQ**E**LISEIDLL**E**KQNEQLKHKLEQLRNSCA
V109	MTEENVKRRTHNVLERQRRN	ELKRSFFALRDQI	PELENNEKAP	KVVILKKATAYI**E**SVQ	AETQ**Q**LI**E**EIDLL**Q**KQNEQLKHKLEQLRNSCA
V110	MTEENVKR**S**THNVLERQRRN	ELKRSFFALRDQI	PELENNEKAP	KVVILKKATAYI**Q**SVQ	AETQ**Q**LISEIDLL**Q**KQNEQLKH**Q**LEQLRNSCA
V111	MTEENVKR**E**THNVLERQRRN	ELKRSFFALRDQI	PELENNEKAP	KVVILKKATAYI**E**SVQ	AETQ**E**LISEIDLL**E**KQNEQLKH**E**LEQLRNSCA
V112	MTEENVKR**A**THNVLERQRRN	ELK**D**SFFALRDQI	PELENNEKAP	KVVIL**D**KATAYILSVQ	AETQ**T**LISEIDLLRKQNEQLKH**Q**LEQLRNSCA
V113	MTEENVKR**A**THNVLERQRRN	ELK**D**SFFALRDQI	PELENNEKAP	KVVIL**D**KATAYILSVQ	AETQ**T**LISEIDLLRKQNEQL**E**QQL**R**QL**E**NSCA

***Table 5**, Omomyc and Omomyc-derived charge variants design description and charge characteristics. *Theoretical values of pI and net charge calculated using the online tool Prot pi (www.protpi.ch).*

	Net charge at pH 5.5*	Net charge at pH 7.4*	pI*	Number of mutation	Design inspiration	Nature of the substitutive amino acids
Omomyc	+8.4	+5.2	9.77	-	-	-
V108	+5.6	+1.2	8.79	2	Structural	Acidic
V109	+5.5	+1.2	8.79	4	Structural	Polar / Acidic
V110	+5.4	+1.2	8.82	5	PMN homology	Polar
V111	+0.8	-3.7	5.67	5	PMN homology	Acidic
V112	+2.5	-1.7	6.30	5	MAX homology	Miscellaneous
V113	-0.2	-3.8	5.41	9	MAX homology	Miscellaneous

2.2. The StrepTag-II is a viable replacement for the HisTag

Previously, we saw that the presence of a HisTag on the N-terminus of the Omomyc sequence could affect its DNA binding. Moreover, since we ought to study the effect of change in net charge over the Omomyc characteristics, using a tag made of basic amino acids may complicate the interpretation of the results. Therefore, the use of HisTag was considered inadequate in this new study. StrepTag was therefore considered as a possible replacement, since only three charged amino acids are present in its sequence (one His, one Lys and one Glu), so that its impact on the net charge of a fusion-protein at physiological pH would be relatively low.

Our previous experience with HisTag showed that its position could have an effect over the structure of a protein. Hence, we decided at this point to conduct a short study to determine the viability of StrepTag for protein purification and characterize its effect over the Omomyc main features. Two constructs were designed bearing a StrepTag on their N-terminus, separated or not from their sequence by a glycine – serine linker (respectively Strep-Linker-Omomyc or SLO, and Strep-Omomyc or SO). A third construct was designed to display a StrepTag on its C-terminus, separated by a glycine – serine linker (Omomyc-Linker-Strep or OLS) (Figure 35).



***Figure 35**, Graphic representation of the three StrepTag control constructs. The effect of the StrepTag-II sequence on the Omomyc mini-protein properties was studied at its N-terminus (SLO and SO) or at its C-terminus (OLS), separated by a short glycine-serine linker or not.*

The three fusion proteins were successfully overexpressed in *E. coli* BL21 transformed with pET3a vectors containing each sequence of interest. After overnight expression, the bacteria were collected and lysed in a high concentration urea buffer. All three proteins could be captured and purified using Streptactin XT pre-packed columns (Figure 36A). After a short condition screening phase, the use of phosphate and magnesium chloride in the lysis buffer was judged optimal to limit the presence of host DNA, co-eluting with the Omomyc variants. Although the variants seemed relatively

pure after a simple capture step, the presence of high-molecular weight host cell proteins (HCPs) was identified by SDS-PAGE in each elution pool (Figure 36B). These HCPs were filtered out using centrifugal filter tubes with a MWCO of 30 kDa (Figure 36C). This operation was found to be much more efficient when carried out in reducing conditions and in the presence of a high urea concentration. Finally, the three control variants underwent a buffer exchange step to a sodium acetate buffer, suitable for long-term storage of Omomyc (Figure 36D). No significant difference was noted between the three constructs in terms of expression or purification.

The secondary structure, the capacity to bind the Max* protein, and the ability to bind DNA were then studied by CD. While OLS and SO displayed a very similar ellipticity, the presence of a glycine – serine linker on the N-terminus of SLO seemed to decrease its α-helix content at 25 ºC, as deduced from its lower MRE at both 222 nm and 208 nm, as well as a lower CD222nm / CD208nm ratio. Tm of the three fusion proteins were however found to be almost identical. In the presence of DNA, all three proteins displayed an increased α-helix content and Tm, signs of DNA binding. Interestingly, the Tm of OLS in the presence of the E-Box DNA probe was found to be about 4 ºC superior to the one obtained for the two constructs displaying a StrepTag on their N-terminus. Lastly, all three constructs demonstrated the ability to bind the MAX b-HLH-LZ domain (Figure 37).

Subsequently, the respective cell penetration capabilities of the variants were studied using the flow cytometry protocol previously described, in Figure 33. All 3 variants were seen to penetrate the A549 cells after a 20 µM treatment for 15 min, as evidenced by a statistically significant fluorescence increase compared to the vehicle-treated cells. However, the fluorescence level observed for the original Omomyc was always seen as superior compared to the one obtained with the streptag fusion proteins, indicating that the streptag may impair the cell penetration properties of its carrier (Figure 38).

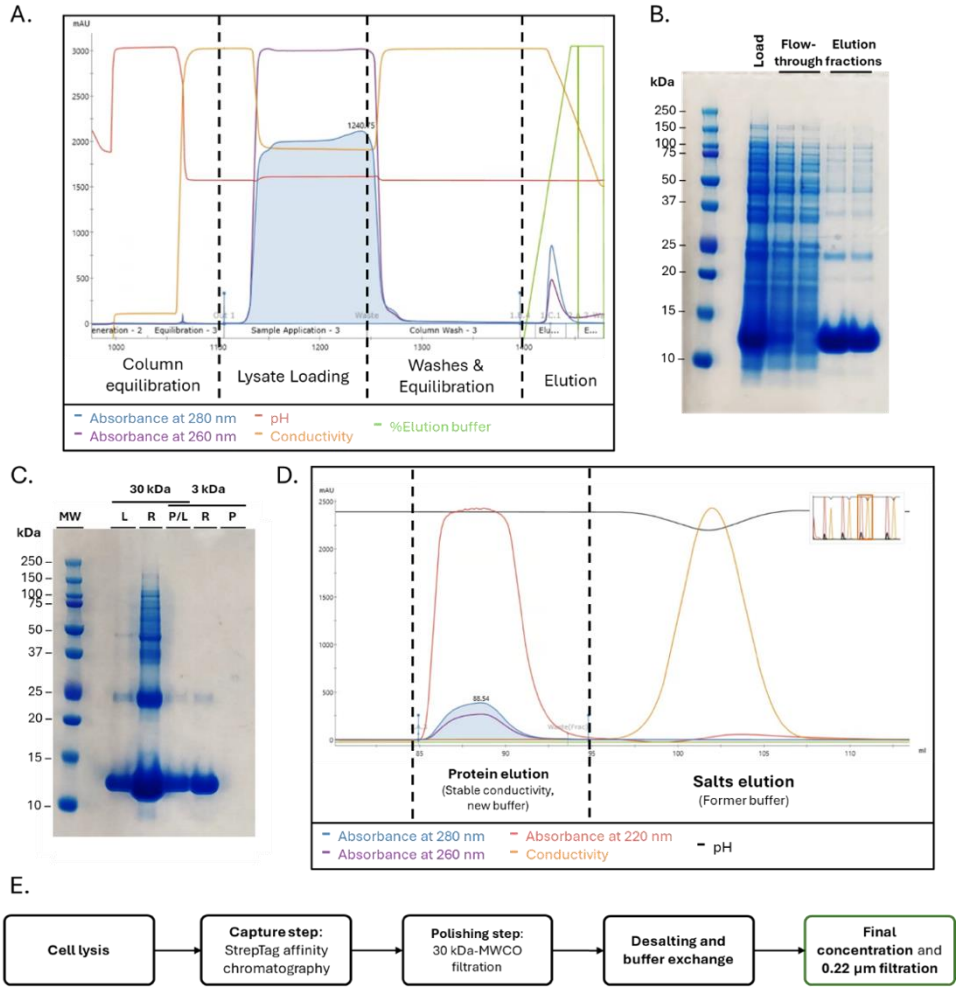


***Figure 36**, **A**. Chromatogram of OLS capture by StrepTag affinity chromatography. **B**. SDS-PAGE (Coomassie blue staining) of the chromatography fractions obtained during the run depicted in (A). **C**. SDS-PAGE (Coomassie blue staining) of the OLS 30-kDa filtration step and subsequent concentration using 3-kDa MWCO centrifugal filtration tube. The same volume of each fraction was loaded. L: load; R: retentate; P: permeate. **D**. Chromatogram example of the buffer exchange from acidified capture step elution buffer to 50 mM sodium acetate, pH 5. **E**. Downstream processing workflow for the obtention of pure stocks of Omomyc variants fused to a StrepTag.*

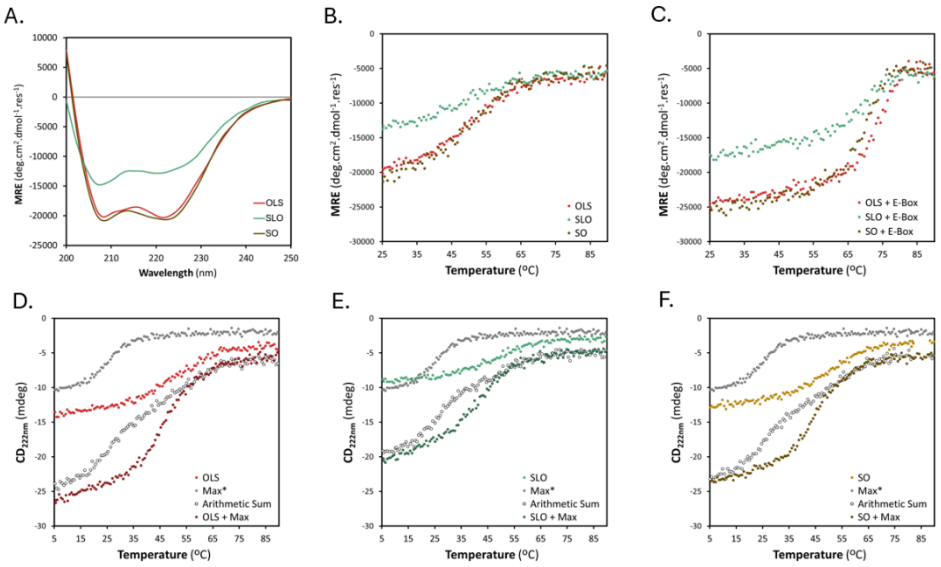


Figure 37, A*. Overlay of the far-UV CD spectra of OLS (red), SLO (green) and SO (brown) at 25 ºC.* ***B-C****: Overlay of the thermal denaturation curve by CD at 222 nm of OLS (red), SLO (green) and SO (brown) (****B****) in the absence of DNA and (****C****) in the presence of 12 µM of the E-Box DNA probe.* ***D-F****: Overlay of the thermal denaturation curve of Max* (grey) with (****D****) OLS (red) and the experimental mix of OLS and Max* (dark red), (****E****) SLO (green) and the experimental mix of SLO and Max* (dark green), (****F****) SO (yellow) and the experimental mix of SO and Max* (brown), together with the arithmetic sum of the Max* and each respective variant thermal denaturation curve (black-circled white dot). All data set presented in this Figure were obtained in the presence of 8 µM of variant and 8 µM of Max* when applicable, in CD1x buffer.*

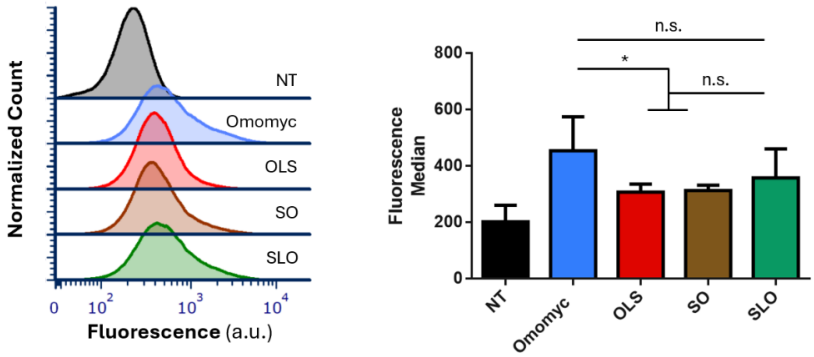


Figure 38*, Left, overlay of the 1D histograms obtained for A549 cell pools treated for 15 min with 20 µM Omomyc (blue), 20 µM OLS (red), 20 µM SO (brown), 20 µM SLO (green) or with vehicle (black). 30,000 events (single cells) were analyzed per condition. A combination of a mouse anti-Omomyc primary antibody and a goat AlexaFluor® 488-labelled anti-mouse secondary antibody was used for Omomyc and Omomyc variant detection. Right, Quantification of the intra-cellular uptake of the variants by A549 cells after a 20 µM treatment for 15 min at 37 ºC (average median of a technical triplicate is represented, the error bars representing the standard deviation across replicate). * p-value < 0.05.*

In conclusion, this short study enabled us to confirm the viability of the use of StrepTag for purification purposes. We observed that the glycine – serine linker placed on the protein N-terminus affected its folding. Also, both the N-terminus StrepTag-displaying constructs had a slightly less stable interaction with the E-Box DNA probe (Table 6). In light of these observations, we determined that placing a StrepTag on the C-terminus of the Omomyc charge variants, following the OLS model, would be the most adequate strategy to follow for subsequent studies.

***Table 6**, StrepTag control variant comparison summary table.*

	Folding	Max* binding	Tm in presence of DNA	Cell penetration
Omo-GGGS-Strep (**OLS**)	OK	OK	73.9 °C	Worse
Strep-GGGS-Omo (**SLO**)	Less folded	OK	69.4 °C	~
Strep-Omo (**SO**)	OK	OK	69.9 °C	Worse

2.3. Net charge modifications may impact the variants solubility profile

E. coli BL21 cells were transformed with pET3a vectors containing each of the Omomyc charge variant coding sequence. Excellent expression levels were obtained for all 6 charge variants. The purification process developed previously (Figure 36E) was successfully employed to purify the whole variant panel. Before proceeding to the final buffer exchange step, the effect of the charge variants mutations over their solubility in different buffers at various pHs was studied. To that end, a controlled amount of mini-protein formulated in an 8 M urea, phosphate-based buffer was injected into prepacked desalting columns installed in an Äkta Pure system. Importantly, our experience with protein buffer exchange has shown that the efficiency of this step is predictive of protein solubility in the target buffer. 10 mg of the 6 variants and OLS were therefore sequentially injected into desalting columns equilibrated with 5 different buffers:

- 50 mM sodium acetate, pH 5
- 100 mM potassium phosphate, 100 mM KCl, pH 4.8
- 50 mM sodium phosphate, pH 7
- 50 mM TrisHCl, pH 6.5

- 50 mM TrisHCl, pH 8.3

The efficiency of the buffer transfer was evaluated according to the shape and symmetry of the protein elution peak in the stable conductivity phase, and according to the proportion of protein eluting with the salts of its former buffer. Each fraction was then collected and inspected visually for signs of precipitation. As illustrated in Figure 39, each variant buffer exchange was empirically categorized as 1. efficient (symmetrical peak, all the injected protein elutes in the new buffer), 2. inefficient and predictive of limited solubility (the elution peak is tailing, a small proportion of the injected protein elutes with the salts of its former buffer), or 3. ineffective and predictive of very low solubility in the target buffer (most of the injected protein elutes with its former buffer salts).

As expected, OLS and the charge variants displayed the best solubility in the two buffers with acidic pHs. The only exception was found for V113, the variant with the lowest pI, which showed signs of limited solubility in both the acetate and phosphate-based buffers. Solubility in the alkaline buffer was improved for the three variants with the lowest pI: V111, V112 and V113. Interestingly, OLS, the protein with the highest pI of the set, in basic conditions displayed better solubility than V108, V109 and V110. Solubility at neutral pHs seemed more sequence- and buffer-dependent. Interestingly, V108, whose sequence differs from OLS's by only 2 residues, displayed very poor solubility in all buffers that did not have an acidic pH, whereas OLS displayed good solubility in both neutral phosphate and tris HCl-based buffers. Overall, the neutral pH phosphate-based buffer allowed a better solubility than the Tris HCl buffer, despite having a slightly higher pH, typically unfavorable for Omomyc-derived variants. Although V111 buffer exchange chromatograms suggested a good solubility in all tested buffers, visual inspection of the collected fraction after buffer exchange in neutral pH buffers revealed a certain cloudiness after a few minutes. This result illustrates well the main limitation of this assay: a change of the elution profiles is only seen in case of fast aggregation events within the desalting column, whereas slower aggregation events are not detected without further studies. The results of this assay are summarized in Table 7.

Learning from the results of this study, the sodium acetate buffer was chosen as storage buffer for the variants V108 to V112, whereas V113 was formulated in 50 mM sodium phosphate, pH 7. In those conditions, the charge variants could be concentrated to over 600 µM in their respective storage buffer and could be stored at – 80 ºC for at least 1 year without showing signs of precipitation or degradation after a single freeze/thaw cycle.

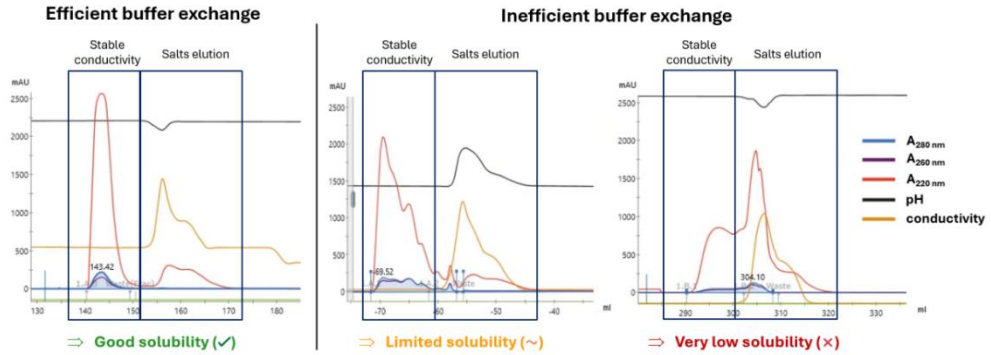


***Figure 39**, The buffer exchange efficiency is indicative of the protein solubility in the target buffer. Examples of chromatogram obtained during three distinct buffer exchange steps representative of an efficient buffer exchange (left): all the protein injected is eluted in the new buffer, during the stable conductivity phase. In this case the elution peak is relatively symmetric; an intermediate situation (middle) where the elution peak is tailing, and a small portion of the injected protein elutes during the salt elution phase; a completely inefficient buffer transfer (right) during which most of the injected protein elutes in the salt elution phase, results indicative of a very low solubility of the protein in the target buffer.*

***Table 7**, summary table of the results obtained during the formulation assay. The charge variants solubility in various buffer conditions was categorized as good (green tick mark), limited (yellow tilde) or very low (red cross).*

	Acetate pH 5	Phosphate pH 4.8	Phosphate pH 7	Tris HCl pH 6.5	Tris HCl pH 8.3
OLS	✓	✓	✓	✓	~
V108	✓	✓	×	×	×
V109	✓	✓	~	~	×
V110	✓	✓	✓	~	×
V111	✓	✓	~	~	✓
V112	✓	✓	✓	~	✓
V113	~	~	✓	✓	✓

After the final 0.22 µm filtration step, pure stocks of 50 mg to 140 mg could be obtained for each variant (Figure 40), their respective concentration was determined by UV spectrophotometry and confirmed by amino acid analysis. Their identity was confirmed by mass spectrometry (MALDI-TOF).

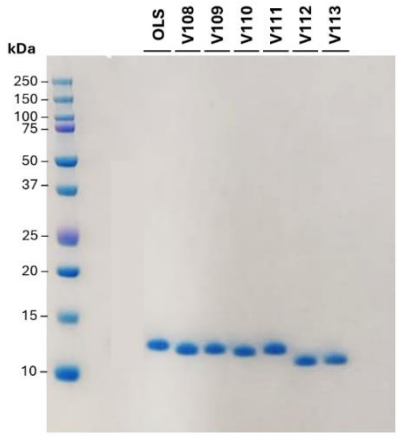


***Figure 40**, SDS-PAGE (Coomassie blue staining) of OLS and each charge variant pure stock after final sterile filtration. 500 ng of each stock were loaded after being boiled for 5 min in Laemmli buffer.*

2.4. The charge variants retained the ability to bind MYC and MAX, despite a destabilized structure

After ensuring the quality of each protein stock, their secondary structure was investigated by CD. As previously predicted, V108 solubility was found to be very low in the standard CD experimental conditions (pH 6.7) and clear signs of aggregation were observed as soon as it was diluted in CD buffer (pH 6.7). After an incubation of 2h at 37 ºC in reducing conditions for disulfide bridge reduction, there was no sufficient soluble V108 left that would allow the gain of an interpretable CD signal. Nevertheless, all the CD characterization experiments were carried out with the whole variant set, including V108, keeping in mind this caveat.

The spectra recorded at 5 ºC for each charge variant suggested that their secondary structure retained a high proportion of α-helix, yet in those conditions, all of them displayed a lower ellipticity than OLS, judging by the lower intensity of the characteristic negative bands at 222 nm and 208 nm (Figure 41A). Of all the charge variant set, V110 retained most of the α-helix content of OLS but displayed a Tm about 12 ºC lower. Interestingly, the basic to acidic residues substitution seemed to further destabilize V111 structure and Tm. Of note, despite different Tm, the transitions of OLS, V110 and V111 displayed similar cooperativity. On the other hand, V109, V112 and V113

transitions were observed to be much less cooperative. Among all variants, V112 and V113 formed the pair with the most altered secondary structure. Both displayed a significantly reduced ellipticity and very long transitions, with a low Tm, suggesting that those variants are less prone to form homodimers in these experimental conditions. For what refers to V109, its thermal denaturation curve starts with a long monotonous transition, followed by another more cooperative one starting at about 42 ºC (Figure 41B).

Clear signs of heterodimerization with Max* were observed by CD for V109, V110, V112 and V113 (Figure 41D-F, H). Not surprisingly, precipitated V108 was not resolubilized by the presence of Max* (Figure 41A-C). In the case of V111, only a mild increase of apparent Tm was observed between the arithmetic sum and the experimental sum curves. This may indicate that this variant does not dimerize with the MAX b-HLH-LZ domain as efficiently as the other variants, or that V111/Max* heterodimerization only has a limited stabilizing effect on either of their secondary structures (Figure 41G).

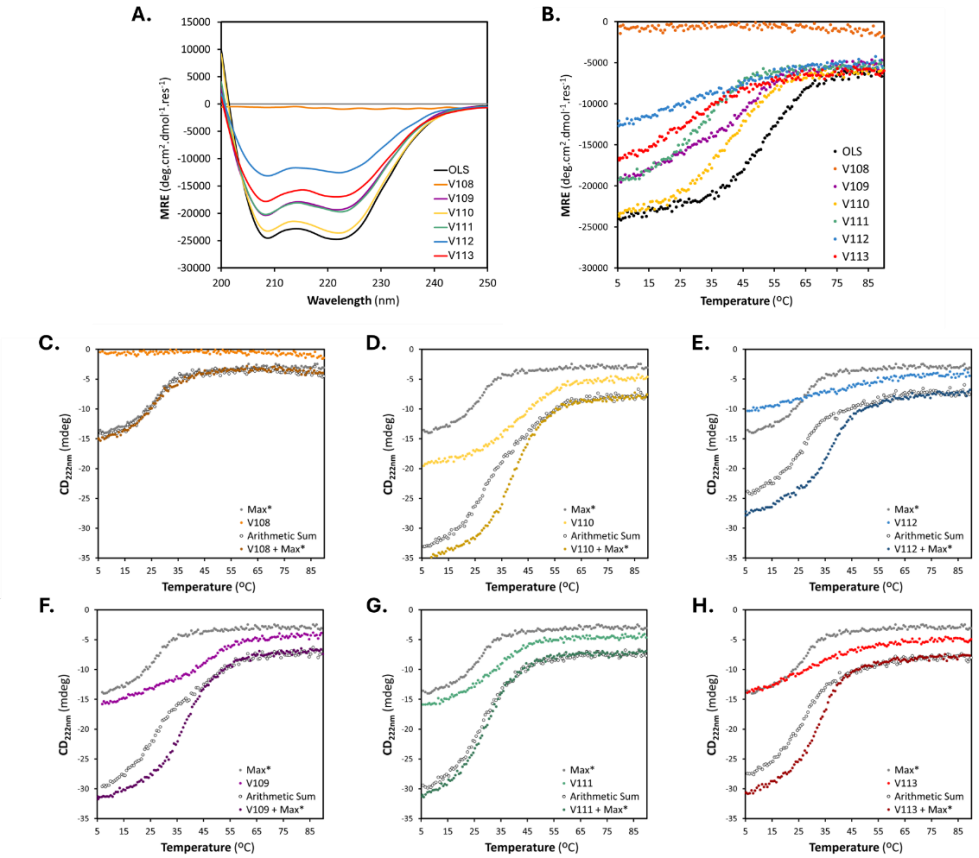


***Figure 41, A-B**: Overlay of (**A**) the far-UV CD spectra obtained at 5 ºC and (**B**) the thermal denaturation curves by CD at 222 nm of OLS (black), V108 (orange), V109 (purple), V110 (yellow), V111 (green), V112 (blue) and V113 (red). C-H: Overlay of the thermal denaturation curve of Max* (gray) with (**C**) V108 (orange) and the experimental mix of V108 and Max* (brown), (**D**) V110 (yellow) and the experimental mix of V110 and Max* (dark yellow), (**E**) V112 (blue) and the experimental mix of V112 and Max* (navy blue), (**F**) V109 (purple) and the experimental mix of V109 and Max* (dark purple), (**G**) V111 (green) and the experimental mix of V111 and Max* (dark green), and (**H**) V113 (red) and the experimental mix of V113 and Max* (dark red) together with the arithmetic sum of the Max* and each respective variant thermal denaturation curve (black-circled white dot). All data set presented in this Figure were obtained in CD1x buffer, in the presence of 8 μM of variant and 8 μM of Max* when applicable.*

The variants' ability to bind the MYC and MAX proteins was also studied in more biologically-relevant conditions, through a co-immunoprecipitation assay (co-IP). A summary of this assay's protocol is depicted in Figure 42. Two MAX-corresponding bands, likely to be the p21 and p22 MAX isoforms, were shown to co-precipitate with OLS and every charge variant in the presence of both 50 and 10 μM of each Omomyc variants in NHDF cell lysate (cells expressing a MYCER fusion protein in addition to the endogenous MYC). OLS, V110 and V113 seemed to be the most efficient variant at co-precipitating MAX. Interestingly, the p21 isoform seemed to be captured slightly more efficiently in some conditions (OLS – 50 and 10 μM, V108 – 50 μM, V110 – 10 μM, V111 – 50 μM), although additional repetition of this experiment should be carried out to determine whether this observation is significant or not. Although not as clear as MAX, at least two MYC-corresponding bands were found in the elution fractions variant-spiked lysates. Judging by their elution fraction content, OLS, V109 and V110 displayed the highest affinity for MYC. In contrast, only faint bands could be observed in V111, V112 and V113-treated lysates. However, results obtained with V108, and V111-treated lysates are to be considered with caution, since clear signs of precipitation were observed after mixing the pure protein solution with NHDF diluted lysate. Overall, those results suggest that all variants conserved the ability to bind the MAX protein, even in a complex matrix. In contrast, the ability to efficiently bind MYC may have only been retained by the variants OLS, V109 and V110 (Figure 43).

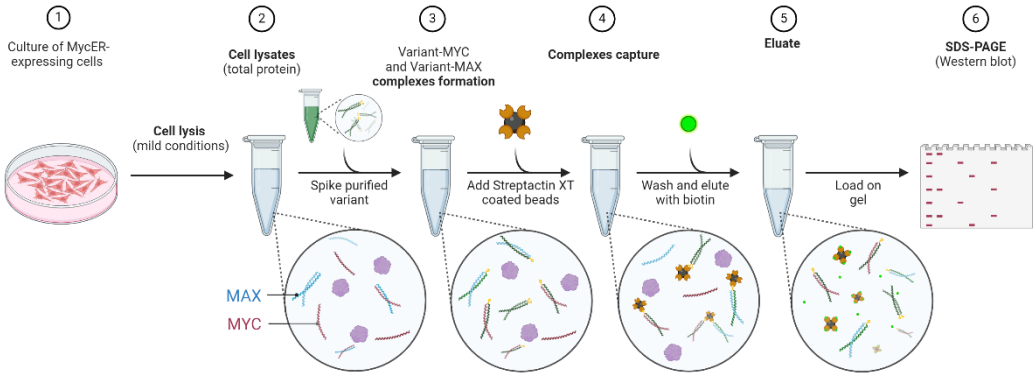


***Figure 42**, Graphic representation of the co-IP protocol. NHDF cells were cultured up to 70% confluency, trypsinized to be collected, and lysed on ice. After separation from the cell debris, the lysate was aliquoted, diluted, and spiked with 50 µM or 10 µM of OLS, either of the charge variants, or an equivalent volume of sodium acetate vehicle. After incubation, magnetic streptactin XT-coated beads were added to each spiked lysate. Following incubation, the beads were separated from the depleted lysate, and the complexes bound to the beads were eluted using a biotin-containing solution. A western blot of each spiked lysate, bead elution fraction, and depleted lysate was carried out to detect the presence of MYC, MAX, and variant in those samples.*

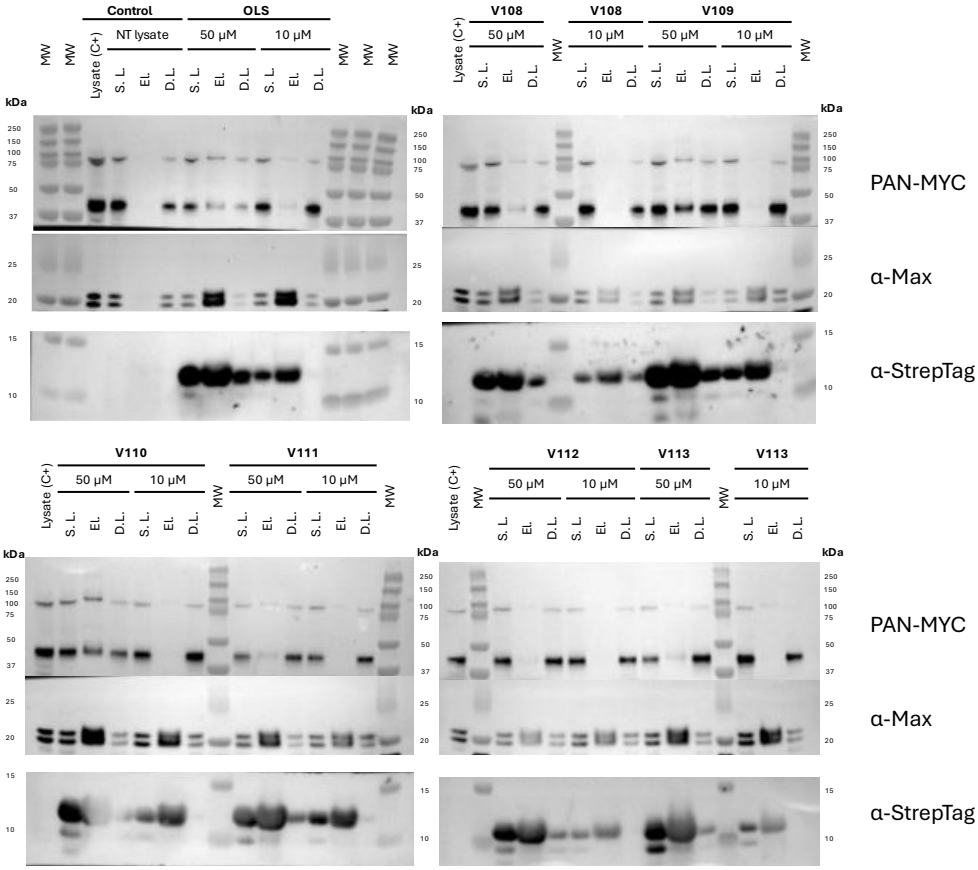


***Figure 43**, Western blot of the MYC and MAX co-IP experiment fractions. Each membrane was separated into 3 parts, enabling the detection of MYC (250 to 37 kDa), MAX (37 to 17 kDa) and Strep-tagged proteins (mass lower than 17 kDa). Each of the 4 corresponding membrane parts were revealed at the same time. MW: molecular weight ladder, Lysate (C+): non-diluted NHDF lysate (positive control for MYC and MAX detection), S.L.: spiked lysate, El.: beads elution fraction, D.L.: depleted lysate after separation from beads.*

2.5. The charge variants retained the abilities to bind DNA and to discriminate specific from non-specific DNA sequences, as homodimer or heterodimer with Max*

In the previous CD experiments, we observed that the substitution of basic residues for polar amino acid could destabilize Omomyc secondary structure. V112 and V113 even showed signs that they could behave as independent monomers rather than homodimers.

Although the basic region of the charge variants was preserved during their design, we hypothesized that their mutated Helix 2 and LZ domain may affect their DNA binding abilities. Each variant was studied in the presence of either the E-Box DNA probe or a non-specific DNA probe. All the charge variants were shown to interact with DNA, as evidenced by systematic increases of their ellipticity at 5 ºC and by increased resistance to thermal denaturation, as well as transition cooperativity. Interestingly, the extent of the structure stabilization was, for each variant, more profound in the presence of the E-Box DNA probe than in the presence of the non-specific one. These results suggest that each variant conserved the ability to properly discriminate specific DNA sequences despite having a destabilized leucine zipper.

Binding to the non-specific DNA sequence stabilizes OLS and its variants to a seemingly intermediate form in between their apo form and their hypothetically most folded state when bound to E-Box DNA. Of note, a relatively conserved and significant increase of about 20 ºC in Tm was observed in all variants upon binding to the E-Box DNA probe. However, differences of Tm among the various variant homodimer/DNA complexes remained significant, and followed the same trend observed in the absence of DNA (Figure 44C). Therefore, we may conclude that the basic region stabilization through DNA binding has also a stabilizing effect on the rest of the sequence. We also demonstrate here that the LZ domain stability

affects the variants DNA binding, so that both protein domains stabilize each other in a bi-directional and cooperative manner.

This same DNA-binding experiment was repeated with each variant/Max* heterodimer (Figure 45). The Max* protein homodimer as well as each variant/Max* heterodimer showed very similar, sequence-specific, DNA binding-induced structure stabilization than the Omomyc variants homodimers. In contrast, Max* heterodimerization had a different effect over the variant DNA-binding stability: on the one hand, the interaction with E-Box DNA of the Max* heterodimers containing OLS and V110 was less stable than with their respective homodimers; on the other hand, the Max* heterodimers containing V109, V111, V112 and V113 showed more stable DNA binding than each respective homodimer. This effect was particularly striking for variants V112 and V113, which were observed to have the lowest Tm, even in the presence of DNA.

***Table 8**, apparent Tm (in ºC) determined for OLS and the charge variants alone (8 µM monomer) or in the presence of Max* (8 µM) and/or in the presence of the E-Box / non-specific DNA probe (12 µM) by CD at 222 nm. *Estimations made considering the end-point temperature of the respective related curve and the fact that their respective onset temperature cannot be clearly determined at 5 ºC.*

	Variant alone			**+ Max***		
	No DNA	+ NS	+ E-Box	No DNA	+ NS	+ E-Box
OLS	52.7	60.9	71.5	48	60.9	63.2/70.8
V108	-	-	-	-	-	-
V109	49.8	53.9	62.3	38.3	53.3	65.8
V110	40.8	58.4	65.4	39.9	55	63.6
V111	32.5	40.8	52.9	30.6	47.6	56.2
V112	< 30*	37.4	46.6	37	49.4	57.4
V113	< 30*	31.6	42.0	33.2	48.5	53.9
Max*	-	-	-	26.2	45.5	53.1

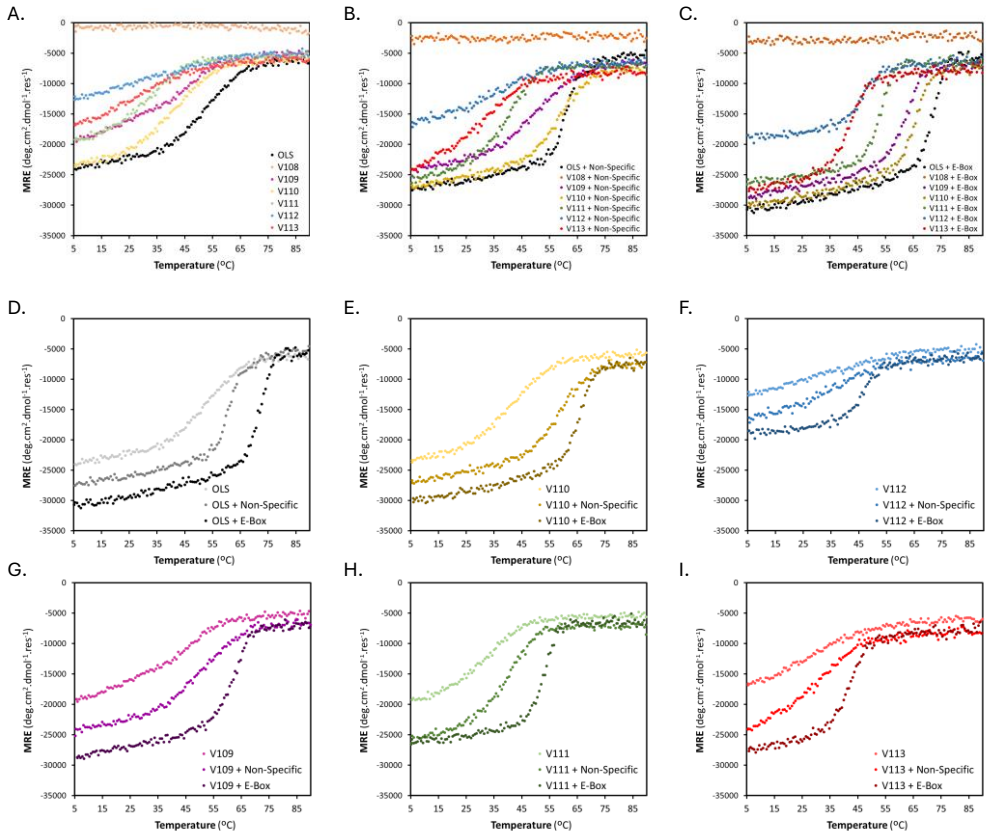


***Figure 44, A-C**: Overlay of the thermal denaturation curves of OLS and the charge variants (**A**) in the absence of DNA, (**B**) in the presence of the non-specific DNA probe, (**C**) in the presence of the E-Box DNA probe. **D-I**: Overlay of the thermal denaturation curves, in the absence or presence of DNA; (**D**) OLS, (**E**) V110, (**F**) V112, (**G**) V109, (**H**) V111, and (**I**) V113. In each Figure the OLS datasets are represented in gray or black, V109 in purple, V110 in yellow, V111 in green, V112 in blue, and V113 in red. Light color: absence of DNA; medium color: presence of the non-specific probe; dark color: presence of the E-Box probe. All data sets were recorded in CD1x buffer in the presence of 8 µM of variant and 12 µM of DNA probe when applicable.*

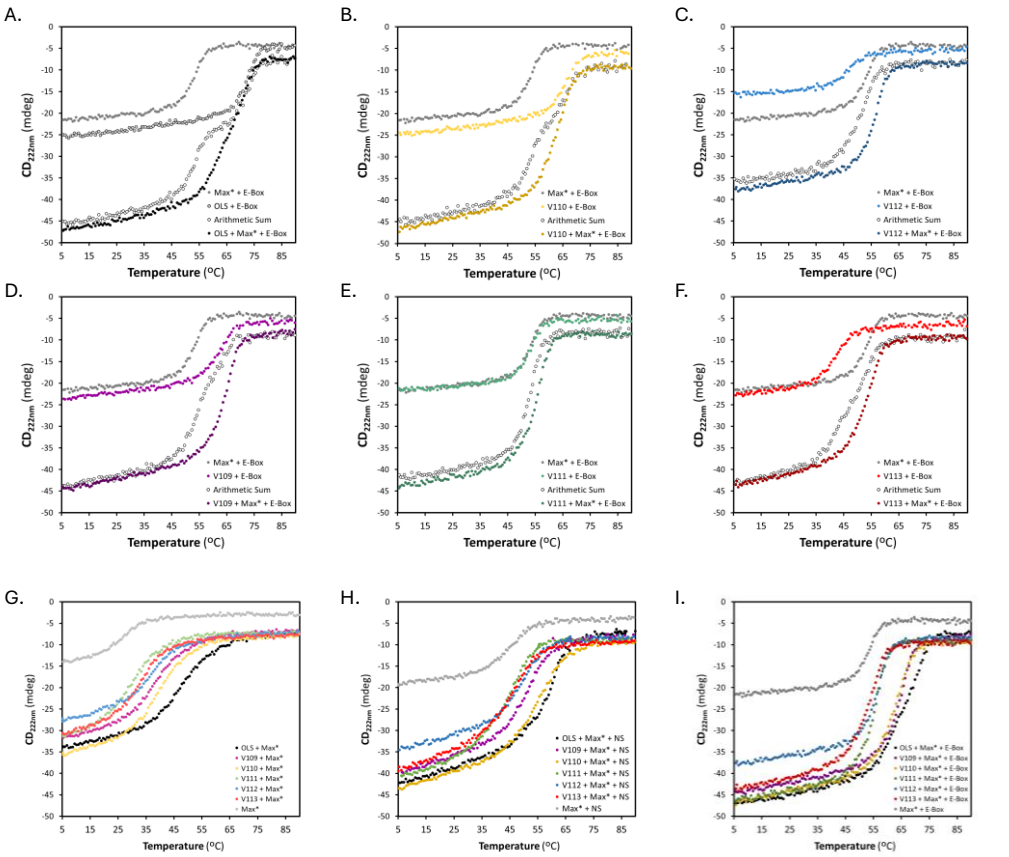


***Figure 45, A-F**: Overlay of the thermal denaturation curve of Max* (gray) with (**A**) OLS (black-circled gray dots) and the experimental mix of OLS and Max* (black), (**B**) V110 (yellow) and the experimental mix of V110 and Max* (dark yellow), (**C**) V112 (blue) and the experimental mix of V112 and Max* (navy blue), (**D**) V109 (purple) and the experimental mix of V109 and Max* (dark purple), (**E**) V111 (green) and the experimental mix of V111 and Max* (dark green), and (**F**) V113 (red) and the experimental mix of V113 and Max* (dark red), all in the presence of the E-Box DNA probe together with the arithmetic sum of the Max* and each respective variant thermal denaturation curve (black-circled white dot). **G-I**: Overlay of the thermal denaturation curves of OLS and the charge variants in the presence of an equimolar concentration of Max* (**G**) in the absence of DNA, (**H**) in the presence of the non-specific DNA probe, or (**I**) in the presence of the E-Box DNA probe. The same color code used in Figure 44A-C is used here. All data set presented in this Figure were obtained in CD1x buffer, in the presence of 8 μM of variant, 8 μM of Max* and 12 μM of DNA probe when applicable.*

A subsequent EMSA carried out in the presence of 500 nM of fluorescently labelled IRD700-E-Box DNA probe further illustrated this affinity gain for DNA. In the absence of Max*, V112 and V113 were shown to only have a marginal effect on DNA mobility, whereas V112/Max* and V113/Max* heterodimers complexed efficiently with DNA. This experiment also confirmed that all 5 other variants could bind DNA as both homodimers and heterodimers with Max* (although the results obtained with the V111 + Max* sample could also be generated by the independent binding of V111 and Max* homodimers to DNA, since both proteins/DNA complexes migrate to very similar distances).

Interestingly, despite the presence of both variant homodimers and heterodimers, resulting from 3:2 molar excess of variant in the variant + Max* samples, no homodimer-DNA band was visible in the corresponding lanes. Although in some cases this could be due to a lack of DNA-mobility shift difference between homodimers and heterodimer, this observation suggests that the variants preferentially form heterodimers binding to DNA with similar or higher affinity than the corresponding homodimers.

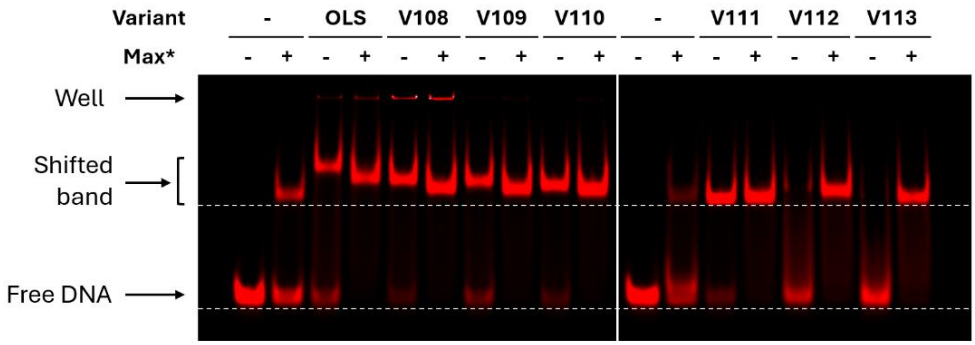
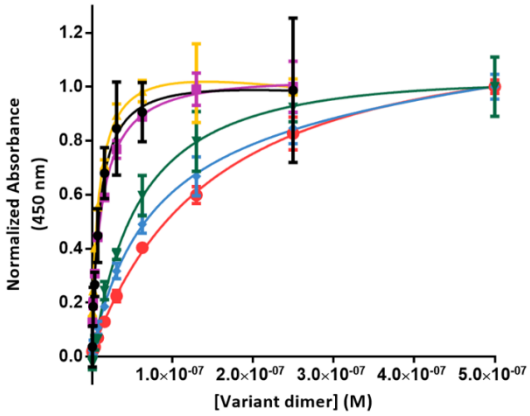


***Figure 46**, EMSA of OLS and the charge variants alone (at 3 µM) or in the presence of Max* (at 2 µM). The proteins were incubated in the presence of 1 µM of a fluorescently labeled IRD700-E-Box DNA probe and the complexes separated on a native polyacrylamide gel. Two white dashed-lines were added as visual reference for the lowest DNA shifted band position (upper line) and lowest free DNA band position (lower line). Both gels represented in this Figure were cast, loaded, run and read simultaneously.*

We have also generated some preliminary data through an enzyme-linked immunosorbent assay (ELISA) (Figure 47) and a fluorescence polarization-based assay (unshown data). These data further illustrate the loss of affinity for DNA of the variants V112 and V113.

The affinity of the variants for immobilized, biotinylated DNA was measured in the ELISA. Kds of about 10 nM were determined for the OLS, V109 and V110 homodimers. In contrast, V111, V112 and V113, which displayed the lowest Tm in presence of the E-Box by CD, also seem to display lower affinity by ELISA.

Of note, apparent differences of DNA saturation levels and important result variability were observed in these results, so that both assays are currently under further development and need optimization.



	Apparent K_d (Homodimer)
OLS	8.8 ± 3.4 nM
V109	13.7 ± 1.8 nM
V110	9.7 ± 1.5 nM
V111	58.9 ± 12.5 nM
V112	71.6 ± 12.8 nM
V113	164 ± 28.4 nM

***Figure 47**, Left: representation of the normalized absorbance at 450 nm (y-axis) in function of the Omomyc variants concentration (x-axis). In this experiment, a biotinylated E-Box DNA probe was immobilized on a streptavidin-coated plate. The plate was then incubated with decreasing concentrations of variant (11 points, 2-fold serial dilution (500 nM to 0.5 nM homodimer)). The variants bound to DNA were detected using an HRP-conjugated anti-StrepTag polyclonal antibody. In this representation, the highest concentration point of the OLS, V109 and V110 series was excluded to achieve better fitting. Results obtained with OLS are represented in black, V109 in purple, V110 in yellow, V111 in green, V112 in blue and V113 in red (error bars: SD across technical triplicate). Right: Apparent K_d obtained through each variant dataset fitting (Graphpad Prism 6, One site total binding fitting model).*

2.6. The charge variant ability to inhibit cell growth does not correlate with their ability to penetrate cells

The charge variants' ability to inhibit cancer cell proliferation in vitro was then studied in two ovarian cancer cell lines, EFO-27 and A2780, as well as in a control BZ1 mouse fibroblast cell line (Figure 48A). The three cell lines were treated with 20 µM of either Omomyc, OLS, V109, V110, V112 or V113 for 5 days. V108 and V111 were not considered in this study due to limiting solubility in cell culture medium. At the end of the 5th day of treatment, the cells fitness was assessed using an Alamar blue-based metabolic assay. Overall, the A2780 cell growth was significantly more impacted by Omomyc and variants treatment than the EFO-27 cells. Indeed, while the Omomyc treatment caused a 75% reduction in the A2780 metabolic activity, a modest 30% inhibition was observed with EFO-27 cells. OLS had a very similar, yet slightly stronger inhibition effect on the metabolic activity of both cancer lines. However, this difference may be explained by OLS aggregation observed in cell culture media after a few treatment days (clearly visible in Figure 48B), more than by a better therapeutic effect, since such aggregation could be toxic for cells and most likely affect their proliferation. In A2780 cells, V109 and V110 seemed to retain most of the Omomyc therapeutic effect, since both mini-proteins caused a 60% reduction in metabolic activity. In contrast, they had little to no effect over the EFO-27 cells proliferation. Finally, variants V112 and V113 did not have any significant effect over either cell line.

Notably, despite only being mildly affected by Omomyc and OLS treatment, the EFO-27 cells displayed significant phenotypic changes upon treatment with some mini-proteins: instead of growing as a spread monolayer, the EFO-27 cells treated with Omomyc and OLS formed compact, spheroid-like structures. This behavior might be indicative of some response to treatment, since such structures seem to be forming also in the presence of V109 (although to a lesser extent), the only charge variant that caused a reduction in metabolic activity. (Figure 48B).

As a sign of aggregation-induced toxicity, OLS treatments were observed to cause a dramatic drop in BZ1-cells proliferation, unlike any of the other treatments which did not cause any significant effect over this cell line proliferation and metabolic activity. These results suggest that this mouse fibroblast cell line could be particularly sensitive to protein precipitation.

Notably, despite the presence of a StrepTag on the C-terminus of each charge variant, no significant aggregation was observed for any of the four charge variants tested in this experiment. This finding suggests that these sequences exhibit improved solubility in cell culture media compared to the original Omomyc sequence.

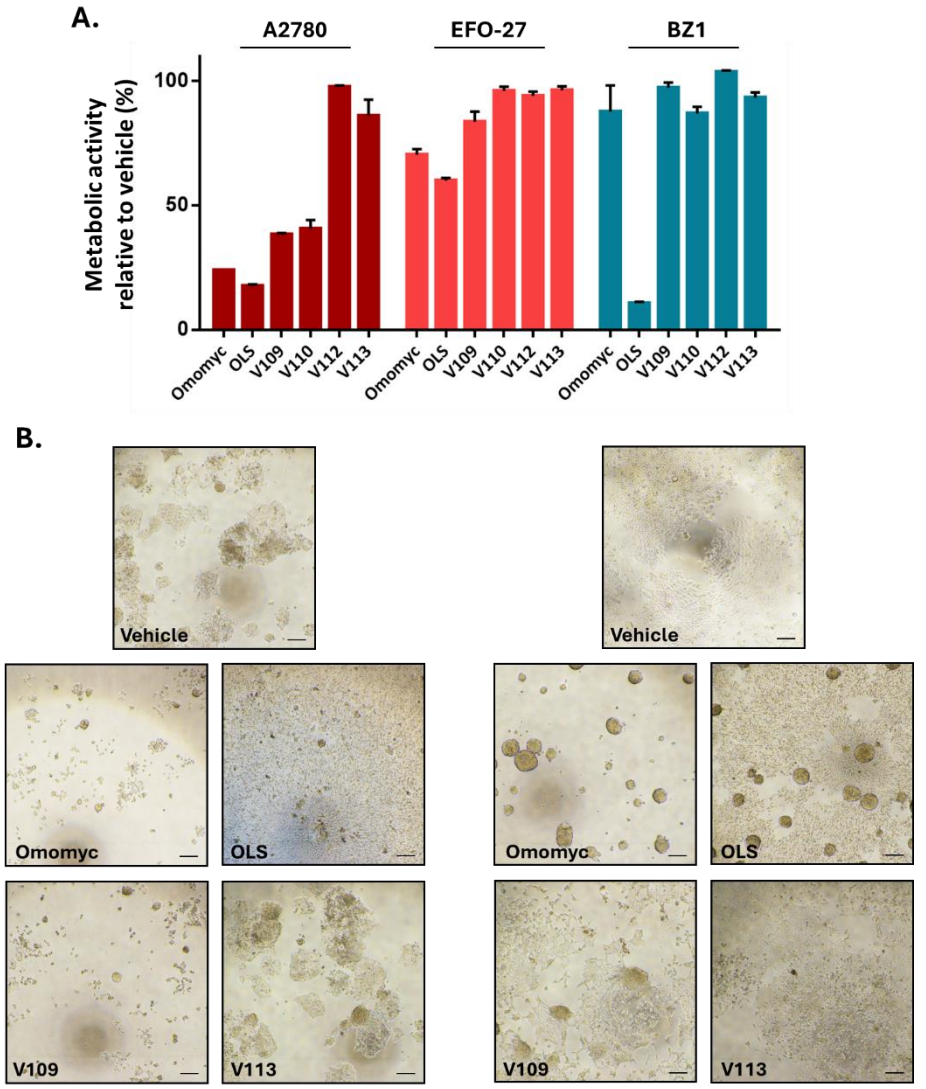
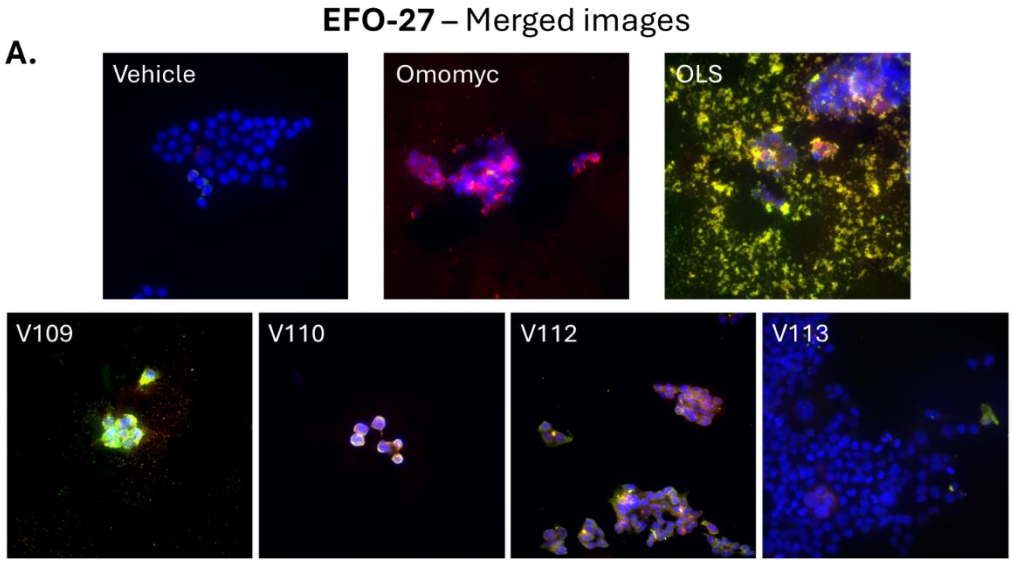


***Figure 48**, **A**, Metabolic activity of A2780, EFO-27 or BZ1 cells treated with 20 µM of Omomyc, OLS, V109, V110, V112 or V113 for 5 days, relative to results obtained with vehicle-treated cells. The metabolic activity of each cell pool was assessed by Alamar blue degradation. Error bars: SD across technical triplicate. **B**, Bright field microscopy images of A2780 and EFO-27 cells after a 5-day treatment with 20 µM of Omomyc, OLS, V109 or V113. Scale bar: 200 µm*

Given that the design of the charge variants implied a significant alteration in the distribution of basic amino acids compared to the original Omomyc sequence, we decided to test whether differences in cell penetration and/or intracellular localization could account for the loss of activity observed in the previous experiment. To this end, we performed an immunofluorescence of A2780 and EFO-27 cells treated with 20 µM of Omomyc, OLS or the variants for 24 h (Figure 49). Those results, although hardly quantitative, demonstrated that the variants V109, V110 and V112 could penetrate cells efficiently, regardless of the cell line. Only V113, the variant with the lowest net charge, did not display convincing signs of cell penetration. Furthermore, no significant differences in intracellular localization between Omomyc, OLS or the variants was observed: all of them were seen to be mainly cytoplasmic, but also co-localized with the DAPI-stained nuclei. Altogether, these results tend to disprove the hypothesis that the lack of variants activity could be due to impaired cell penetration properties, given that no clear correlation could be established between these results and the efficacy results described previously. A more quantitative assay should be developed to further confirm this conclusion.



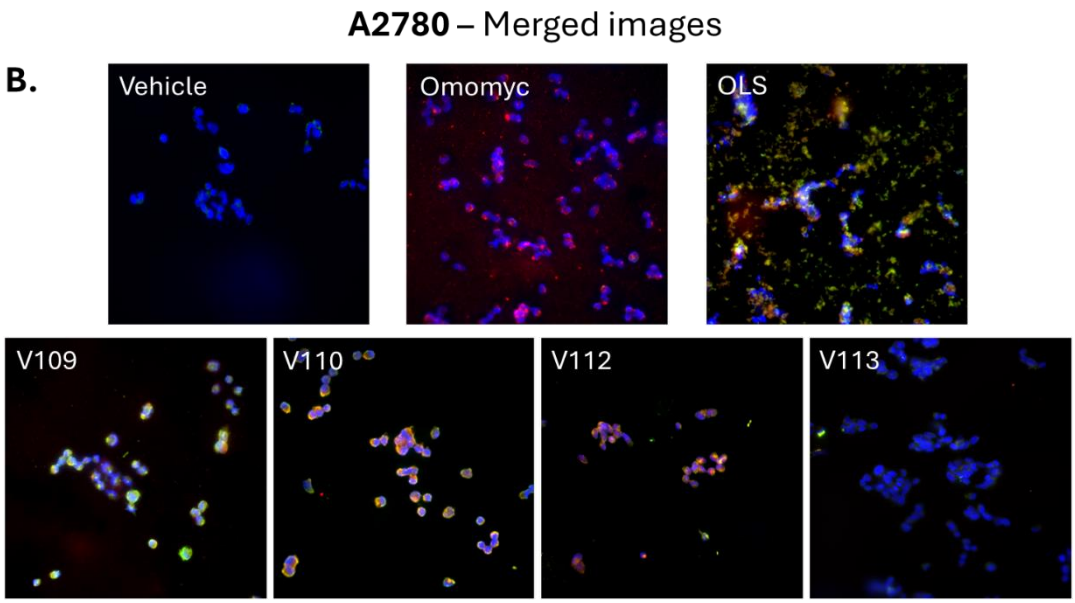


***Figure 49**, immunofluorescence of (**A**) EFO-27 and (**B**) A2780 cells treated with 20 µM of Omomyc, OLS, V109, V110, V112 or V113 for 24h. Each image was obtained through the merge of three distinct color channels: blue channel, DAPI staining; red channel, anti-Omomyc staining; green channel: anti-StrepTag staining. Images were captured by epifluorescence microscopy.*

2.7. A newly developed assay highlights the charge variants different behavior in plasma

One of the main goals of this study was to identify whether a change of net charge at physiological pH could benefit the PK properties of Omomyc. To this end, an assay capable of determining low variant concentrations in biological matrices, such as plasma or serum, needed to be developed. To have an assay compatible with a broad array of sequences, and to allow a direct comparison of the results obtained for the different Omomyc-derived variants, a competition assay was designed. In this assay, the variant quantification is based on the competitive binding of the StrepTag-displaying variants against a known amount of StrepTag-GFP fusion protein for an anti-StrepTag antibody-coated plate. After the final washing steps, the GFP fluorescence intensity is inversely proportional to the concentration of

variant originally present in the sample to be analyzed, regardless of the variant sequence. The rationale of this assay is summarized in Figure 50.

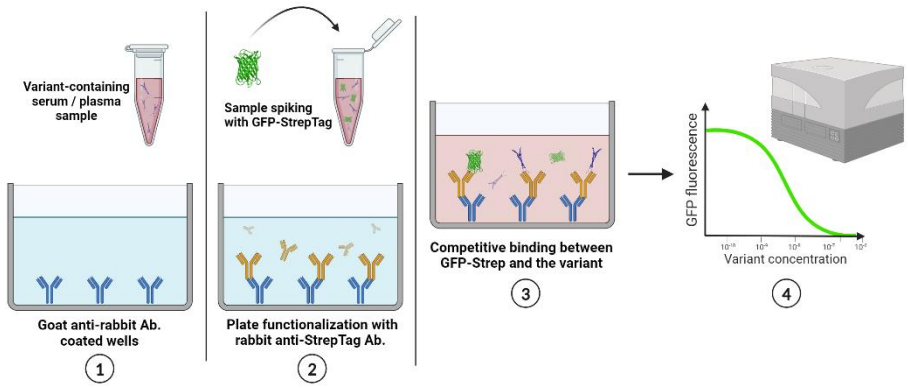


Figure 50, *Graphic representation of the ELISA protocol for Streptagged protein concentration determination. (**1**) This assay requires black 96-well plates, functionalized with anti-rabbit antibodies. (**2**) The serum / plasma sample is spiked with a known amount of GFP-StrepTag while the plate is functionalized with rabbit anti-StrepTag antibodies. (**3**) The spiked sample is then incubated onto the plate where the GFP-StrepTag and variant compete for binding. (**4**) After the final washing step, the GFP fluorescence intensity is read and is inversely proportional to the concentration of Streptagged variant originally present in the sample.*

The first step of this assay development aimed to determine the optimal concentrations of both the anti-Streptag antibody and GFP-StrepTag to be spiked into the samples of interest (Figure 51A). The highest GFP signal was obtained after functionalizing the plate with a 1:1000 antibody stock dilution and the subsequent incubation of concentrations higher than 35 nM of GFP, diluted in PBS + 0.1% polysorbate (PBST). No GFP-StrepTag plate binding was detected in the absence of anti-StrepTag antibody. In PBST, OLS was shown to compete efficiently with 12 nM, 35 nM or 100 nM of GFP-StrepTag for plate binding, in a concentration-dependent manner. IC50s of about 6 nM were determined in the three conditions. In this experiment, OLS concentrations as low as 1.4 nM were shown to cause a significant decrease in GFP fluorescence intensity and was a first confirmation of the viability of this approach for the determination of low OLS concentrations. Despite being comparable, the dataset obtained in the presence of 12 nM GFP-StrepTag showed more variability across technical triplicates and a slightly lower signal variation range. Also, as one could expect, OLS competed slightly more efficiently for the plate in the presence of 35 nM GFP than of 100 nM. 35 nM was therefore considered to be an optimal concentration of competing GFP to spike and was used to obtain all subsequent results.

After determining the basic experimental conditions, the effect of biological matrices over the assay was studied. Although the assay was designed to eventually determine the PK properties of each charge variant in mice after intravenous injections, the use of blood-derived matrices from human origin was preferred during most of the assay development work, due to easier accessibility to human blood-derived products. In the experiment, the same OLS serial dilution was prepared in PBST, human serum and human plasma. These samples were then mixed 1:1 with a 70 nM GFP-Streptag solution, prepared in PBST, to achieve a constant 35 nM final GFP concentration. Interestingly, the ability of OLS to compete for the plate decreased in serum and, to an even higher extent, in plasma. In contrast, the GFP fluorescence intensity in the absence of competitor was comparable in the three conditions, suggesting that this matrix effect is unlikely to originate from the introduction of a new competing molecule. A higher OLS lower limit of detection in serum and plasma was therefore determined as a direct consequence of this effect: the first statistically significant decrease in GFP fluorescence was observed in the presence of 12.3 nM of OLS in serum and 37.0 nM in plasma, against 1.4 nM in PBST (Figure 51C). Additional characterization steps carried out with human serum and plasma demonstrated a good reproducibility for this assay. Different incubation times of OLS or GFP-StrepTag in the blood-derived matrices between 30 min and 2 h did not have a significant and consistent effect over the assay results quality. Finally, a single freeze/thaw cycle of OLS standard curves in serum/plasma was shown to slightly diminish the proteins' capacity to compete for the plate binding, although it was only observed if the non-treated plasma or serum used to make the serial dilution was freshly processed and had not undergone a freeze / thaw cycle before.

Surprisingly, when mouse and human sera were compared, radically different results were obtained. First, the maximum GFP fluorescence intensity detected in mouse serum was about half of what was usually obtained in human serum. Importantly, this comparative experiment was carried out on a single plate excluding a plate effect. Furthermore, the OLS standard curve in mouse serum was characterized by a very low Hill slope: the differences of fluorescence between each consecutive standard point were very little, and the baseline could not be reached even in presence of the highest OLS concentration point. In other words, 10 µM of OLS could not completely displace the 35 nM of GFP-StrepTag in these conditions.

The experiment was then repeated, adding rat serum into the comparison (Figure 51D). Interestingly, the results obtained in rat serum were halfway between the ones observed in mouse and human sera, both in terms of maximum GFP fluorescence and OLS competitiveness. In contrast, when the same comparison was made between the three species plasma, nearly identical standard curves were obtained (Figure 51E). Therefore, since the results obtained in mouse plasma showed greater sensitivity than in serum, this blood fraction was selected to perform the following experiments.

Then, the ability of OLS, V109, V110, V112 and V113 to competitively bind anti-StrepTag-functionalized plates was analyzed. V108 and V111 were again excluded from this study, since their tendency to heavily precipitate in neutral pH conditions would be a source of major technical concerns. Interestingly, all the tested charge variants competed more efficiently for the plate binding than OLS in mouse plasma, although not in an equal manner (Figure 51F). As competitive inhibitors of the anti-StrepTag antibody/GFP-StrepTag interaction, OLS displayed an IC50 of 245 ± 60 nM, V109 showed an IC50 of 146 ± 31.0 nM, and V110 had an IC50 of 82 ± 12 nM. V112 and V113 were the best competitors, with apparent IC50 values of 32 ± 8 nM and 14 ± 2 nM, respectively. These significant differences in apparent IC50 values highlight the fact that this assay displays different sensitivity for the different variants and that future variant concentration interpolation can only be made with the corresponding variant standard curve.

In a last set of assay characterization experiments, we aimed at evaluating the variant separation into the different blood fractions and to determine the portion of variant spiked into whole blood that would be detectable after plasma processing. This set of experiments will be referred to as *fractionation experiments*. To emulate the conditions of an actual PK determining experiment, mouse blood was spiked with pure variant only a few seconds after collection. For each variant, the blood of three mice was spiked with three different variant concentrations, each representing a value expected at different timings during a PK time-course experiment. The highest concentration, 50 µM, would be expected at t_0 after an intra-venous injection of 50 mg/kg dose to a mouse weighing about 25 g. The medium concentration, 1 µM, would be expected to be measured a few hours later. Finally, the low concentration of 50 nM would be expected after several hours to several days following the same dose injection. The assay lower limit of detection for OLS was determined to be close to this value. After

blood processing, a part of the plasma derived from 50 µM and 1 µM-spiked blood was further diluted with an untreated mouse plasma pool (those samples will be referred to as 50 µM series, 1 µM series and 50 nM series sample; Figure 52A).

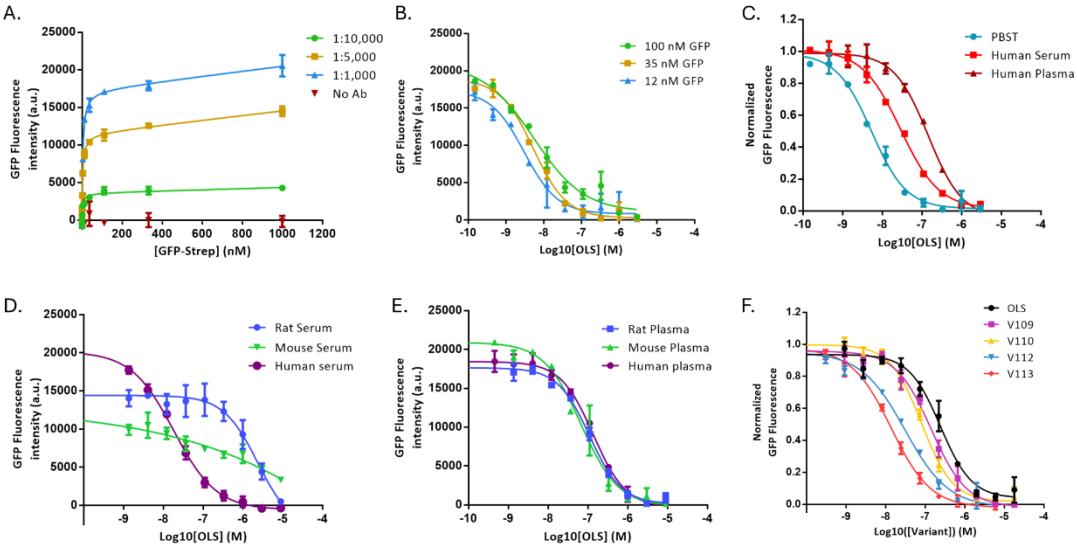


***Figure 51**, **A**, Plate saturation experiment, representation of the GFP fluorescence intensity (y-axis) in function of the GFP-StrepTag concentration plated (x-axis). The plate was first incubated with three different α-StrepTag Ab stock dilution (1:10,000, green; 1:5,000, yellow; 1:1,000, blue) and with buffer only (brown). The plate was then incubated with decreasing concentrations of GFP-StrepTag diluted in PBST (10 points, 3-fold serial dilution (1 µM to 0.05 nM)). **B**, OLS competes with the GFP-StrepTag for binding to the plate in PBST, representation of the GFP fluorescence intensity (y-axis) in function of competing OLS concentration (logarithmic values, x-axis). Three different competiting GFP-StrepTag concentrations (final concentration: 100 nM, green; 35 nM, yellow, 12 nM, blue) were spiked in an OLS standard curve (10 points, 3-fold serial dilution (3 µM to 0.15 nM). **C**, OLS is less efficient at competing for plate-binding in human serum and human plasma than in PBST. Representation of the normalized GFP fluorescence intensity (y-axis) in function of the GFP-StrepTag concentration plated (x-axis). OLS standard curves (10 points, 3-fold serial dilution (3 µM to 0.15 nM)) were prepared in PBST (blue), human serum (red) and human plasma (brown). All points were spiked with 35 nM (final concentration) of GFP-StrepTag prior to incubation on the plate. The fluorescence intensity values were normalized with values obtained for the samples containing 35 nM of GFP-StrepTag without competing OLS. **D-E**, OLS standard curves (10 points, 3-fold serial dilution (9 µM to 0.46 nM)) prepared in (D) serum or (E) plasma of rat (blue), mouse (green) and human (purple) origin. **F**, Standard curves in mouse plasma of OLS (9 µM to 0.46 nM, black), V109 (9 µM to 0.46 nM, purple), V110 (9 µM to 0.46 nM, yellow), V112 (3 µM to 0.15 nM, blue) and V113 (3 µM to 0.15 nM, red) competing with 35 nM of GFP-StrepTag. The background signal was rested in all data points presented in this panel, each point representing the average value obtained from a technical triplicate (error bars: SD across triplicate).*

To our surprise, only a small fraction of the protein spiked in whole blood could be detected in the resulting plasma (Figure 52). A significant difference between observed and expected signal was found for all variants. Considering only the data points with a fluorescence intensity falling in the linear part of each respective standard curve, OLS and V109 were found to be the variants with the lowest detected fraction of theoretical concentration, with signal corresponding on average to only 3.5% and 1.2% of the theoretical concentration, respectively. V110 and V112 gave slightly better results, with 6.2% and 10.4%, respectively. V113 was the most detected variant, since the interpolated concentration values corresponded to an average of 22.9% of the theoretical concentrations. However, overall good reproducibility was observed, and the samples of same theoretical concentration from different series were seen to give very similar results, with the exception of V110.

These observations imply that most of any variant originally present in whole blood is lost or becomes unavailable to bind the plate in the resulting plasma, further decreasing the sensitivity of the assay. As consequence, only concentrations of OLS, V109 or V110 originally present in the whole blood in the micromolar range would be detected with this assay, most likely rendering it unsuitable to monitor the variants concentration in blood a few hours after an intravenous injection.

At this point, the use of detergent was considered as an attempt to free variants potentially entrapped in complexes. The results of a short condition screening demonstrated that the presence of SDS inhibited the StrepTag/antibody union, even at low concentrations. In contrast, very high polysorbate concentrations affected neither the GFP fluorescence nor the StrepTag/antibody interaction. Therefore, in parallel to a simple repeat of the 1st fractionation experiment, a final concentration of 5% Tween®20, 50 mM sodium phosphate, 10 mM of potassium chloride and 550 mM of sodium chloride were added into the blood, 5 min after collection and spiking with pure variant in a second set of samples. The results obtained through the 2nd fractionation experiment are shown in Figure 53. As expected, the detergent-containing blood samples hemolyzed. Consequently, after centrifugation, the supernatant corresponding to plasma in a normal blood sample conserved an intense red color. Of note, the Buffy coat was not visible, and the red blood cell fraction was smaller than in the absence of detergent. In the absence of detergent, the results of

the 1st fractionation experiment were largely recapitulated. Interestingly, despite increased variability, in detergent-spiked samples we observed variant concentrations overall closer to their theoretical concentrations, particularly in the case of V109, V110 and V112. Although preliminary, these results demonstrate that there is probably enough room to improve this assay's sensitivity in order to make it viable to accurately determine the PK of virtually any StrepTag-displaying protein.

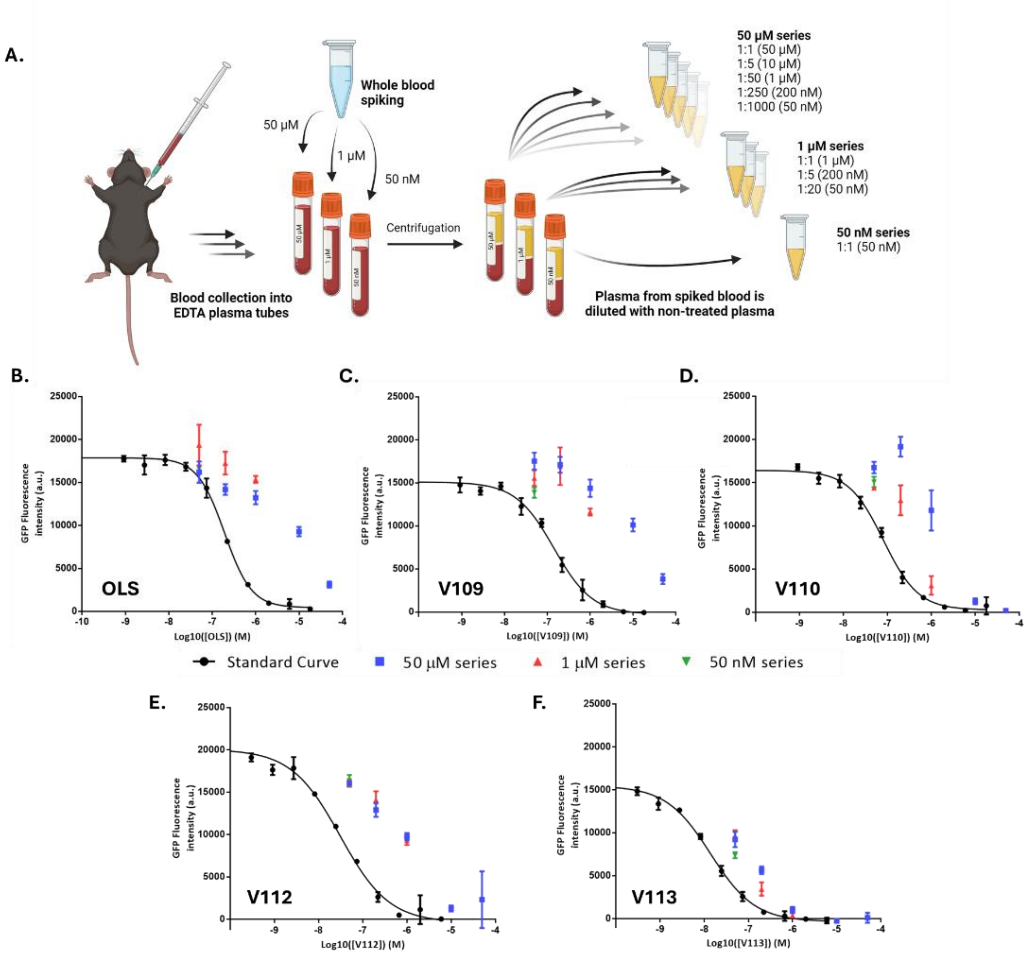


***Figure 52, A**. Schematic representation of the fractionation experiment sample generation. For each variant, the blood of three mice was collected intracardiacally and placed in three corresponding, EDTA-containing tubes. After 30 sec, the three blood samples were spiked with 50 µM, 1 µM or 50 nM of variant, final concentration. Each tube is then processed, and the plasma fraction saved in a distinct tube. The plasma originating from the 50 µM-spiked blood is then divided into five tubes, where it was either left undiluted, or diluted 5, 50, 250 or 1000-fold with an untreated mouse plasma pool. The plasma originating from the 1 µM-spiked blood was split into three tubes where it was either left undiluted or diluted 5 or 20-fold. The plasma originating from the 50 nM-spiked blood was*

*left undiluted. **B-F**, Results of the fractionation experiment obtained with (**B**) OLS, (**C**) V109, (**D**) V110, (**E**) V112 and (**F**) V113. In each Figure, the results obtained with samples from the 50 µM series are represented in blue, from the 1 µM series in red and from the 50 nM series in green. The GFP fluorescence resulting from each of those samples was plotted at their respective theoretical variant concentration after dilution in untreated plasma. Each respective standard curve is represented in black. Each data point represents the average GFP fluorescence intensity obtained from a technical triplicate (error bars: SD across technical triplicate).*

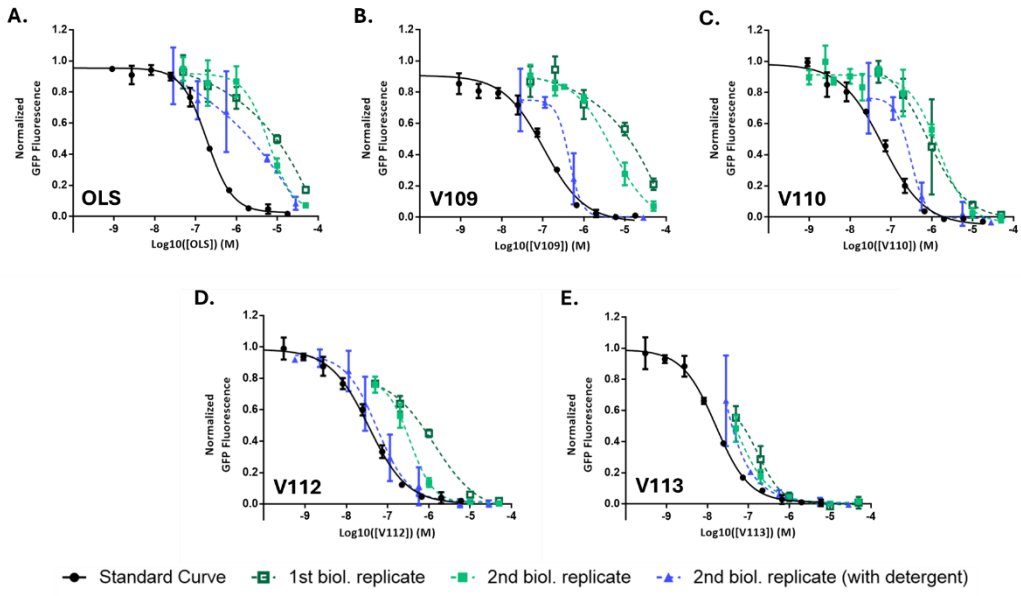


***Figure 53**, Overlay of the results obtained in the experiment 1st replicate (dark green, empty squares), and 2nd replicate in the absence (green squares) or the presence of detergent (blue triangles) with (**A**) OLS, (**B**) V109, (**C**) V110, (**D**) V112 and, (**E**) V113. Each respective standard is represented in black. In this Figure, each data point represents the mean fluorescence value obtained with every sample of equal theoretical variant concentration, regardless of its original sample series (50 µM, 1 µM or 50 nM). Error bars: SD across replicate.*

V. Discussion

It is common knowledge that an ideal drug candidate should possess high affinity and specificity for its target(s), as well as appropriate tissue exposure to exert its activity. As a consequence, the characteristics and early steps of a drug development process are profoundly linked to the nature of both the molecular target and the drug candidate model. Due to the strong relationship between structure and activity in proteins, during the development of protein-based therapeutics, careful consideration must be given to structural aspects, as these are likely to influence the protein's affinity and specificity for its target. Interestingly, while most proteins are required to be correctly folded prior to exerting their biological activity, the b-HLH-LZ proteins tend to fold at the same time as they exert it. The capacity of b-HLH-LZ to properly fold in the presence of their binding partner originates from protein-protein or protein-DNA sequences complementarity and is therefore indicative of affinity and specificity for their binding partner, making structural studies particularly relevant at early stages of a b-HLH-LZ-based MYC inhibitor screening/development process.

Moreover, while most published studies focus exclusively on target binding efficiency, adequate PK or tissue exposure turns out to be as important. In fact, a recent review analyzing the clinical trial data of the past decade estimated that up to 15% of clinical trial failure originated from the drug's poor PK profiles[255]. However, in the case of MYC, effective exposure depends not only on a favorable PK but also on the drug's ability to penetrate cells and localize within the nucleus. Indeed, unlike small molecules that can passively cross the cell membrane due to their small size and rather hydrophobic nature, proteins typically require active mechanisms to penetrate cells. Since the discovery of the first cell penetrating peptides, penetration mechanistic studies have demonstrated that protein cell penetration is achieved through endocytosis, followed by endosomal escape or by direct translocation to the plasma membrane[256]. Based on all these notions, early discovery processes of protein-based MYC inhibitors should therefore be primarily focused on estimating the candidate's binding efficiency to MYC, MAX and E-Box DNA, as well as their cell penetration properties.

This was the general rationale of this thesis project.

Section 1: Modulation of the cell penetration activity and nuclear localization of b-HLH-LZ proteins using short-stretch mutation and cell penetration tags

Intrinsic cell penetration is a fundamental aspect of the Omomyc mini-protein activity, and its discovery led our group and the scientific community to finally consider Omomyc as a viable cancer drug candidate, through its MYC-inhibiting effect. A few years of development later, a historical landmark was achieved by making Omomyc the first direct MYC protein inhibitor to reach clinical phases, more than twenty years after its first design.

The attention generated by the pre-clinical demonstration of the safety and anti-tumoral Omomyc activity undoubtedly contributed to the recent emergence of b-HLH(Z) proteins as a new class of therapeutic candidates for the inhibition of hard-to-reach transcription factors. Incidentally, several publications related to this topic use Omomyc as a benchmark to evaluate the efficacy of new molecules, many being directly Omomyc-based. Furthermore, among those studies, some suggest that the fusion of Omomyc to a CPP or a NLS sequence could be a strategy to further strengthen Omomyc's activity[190,196].

Following a similar design strategy, we decided to fuse three well-characterized short CPP sequences, namely SynB3, Penetratin and R9, to the N-terminus of Omomyc, Max* and 5 Omomyc-derived sequences presenting a truncated LZ domain and/or short stretch substitutions with short sequences with a renown NLS or CPP activity. Moreover, we added a HisTag to the N-terminus of all these fusion proteins for purification and detection.

From the initial phases of this project, the use of HisTag has been most valuable to streamline the purification of this panel composed of 23 different constructs. Although all of them display a high content in basic amino acids and would be expected to have a significant affinity for a CEX resin, the use of HisTag affinity chromatography resins was particularly convenient, due to its high tolerance to high concentrations of chaotropic agents, such as urea, which can interfere with protein binding to CEX resins. Indeed, we considered the use of urea at the early stages of this project because it allowed for the development of an extraction protocol that would be efficient and common for all variants. Results proved us right, since a

significant amount of protein could be purified for almost all the variants using a single extraction and purification protocol, which required very little troubleshooting to be developed. Indeed, urea proved to be a valuable ally during both the variants extraction step, and the purification step elution phase, where it helped maintaining an acceptable solubility for some variants in the seemingly unfavorable neutral pH conditions required for HisTag affinity chromatography. Exception cases were observed with V93 (HisTag-SynB3-OmoT3) and V94 (HisTag-Penetratin-OmoT3) that were curiously under-expressed by all the tested colonies of transformed *E. coli* and therefore could not be purified with high yields. However, these results did not completely surprise us since past attempts to purify other Omomyc truncated LZ variants were not successful, notably due to poor expression levels. Therefore, we could hypothesize that Omomyc-derived variants with large LZ truncations exhibit different expression and stability behaviors in E. coli cells, such as changes in oligomeric state, interactions with cellular components, and subcellular distribution. These changes may affect the stability of the variants and, as a result, hinder their intracellular accumulation. From that perspective, V95 (HisTag-R9-OmoT3) becomes an exception, since it was seen to be well-expressed by *E. coli,* in contrast to the other two OmoT3-based variants. V95, however, later displayed very poor solubility in the absence of urea, even in acidic conditions, which were observed to be the most favorable for all the MYC or MAX b-HLH-LZ-derived constructs studied by our group. Altogether, these observations may provide clues as to why the truncated variants with improved activity described in Brown et al.[193] have not yet been the subject of a follow-up publication.

In fact, V95's lack of solubility is illustrative of the rest of the R9 variants sub-group, which showed a tendency to have a lower solubility than their respective SynB3 and Penetratin counterparts, with the exception of V88 (or H8-R9-Omo), that displayed comparable solubility as V84 and V86 in the few conditions where they were tested together. This difference remains rather mysterious, since all three CPP tested have a high content in basic residues and a similar size. In this context, though, one could imagine that the monotonous sequence of the R9 tag might be involved in stronger non-specific interactions with contaminant traces present in the protein pool (such as nucleic acids or phospholipids), further destabilizing their structure fragilized by their respective short stretch mutations and leading to aggregation events during concentration attempts.

The execution of CD experiments provided valuable insights into the complex structural behavior of this variant panel:

First of all, through the characterization of the H8-Omo control variant we could demonstrate that the addition of a HisTag on the Omomyc N-terminus does not significantly impact its secondary structure or its ability to bind Max*. As a matter of fact, both Omomyc and H8-Omo showed an almost perfectly over-lapping far-UV CD spectra and thermal denaturation curves, even represented as mean residual ellipticity (which is normalized to the number of amino acids of each protein sequence), suggesting that the HisTag may adopt a certain α-helical content and/or induce it in Omomyc's BR.

Similarly, none of the 3 CPPs were seen to significantly impact Omomyc structure or its ability to bind Max*. Supporting this idea, the pairs of variants with the same Omomyc moiety displayed very similar CD spectra and thermal denaturation. Interestingly, within each pair, the SynB3 variants displayed a slightly superior α-helix content compared to their respective Penetratin-displaying counterpart, suggesting that the SynB3 tag has a higher propensity to fold into α -helices than Penetratin. Another plausible explanation for this observation could be a systematic deviation during the variants stock concentration determination by spectrophotometry at 280 nm. This deviation could originate from the fact that the Penetratin sequence comprises two Trp while the SynB3 only displays a single Tyr and no Trp (like the Omomyc sequence).

Compared to the fusion of Omomyc to CPP tags, the short stretch mutations in the Omomyc's loop or C-terminus, creating the OmoCrot, OmoEGFR, OmoNLS2 and OmoNLS4 sequences, caused more profound changes in the variants structure. Indeed, each of them displayed a significantly lower α-helix content than H8-Omo at RT. These results suggest that mutations located at the loop or C-terminus of the Omomyc sequence can significantly impact its structural integrity. Additionally, the OmoCrot and OmoNLS4 variants have significantly lower Tm. The OmoCrot-derived sequences, presenting both a mutated loop and a truncated LZ, were affected the most. Their monotonous, non-cooperative transition strongly suggests that these mini-proteins exist primarily in a monomeric state.

In contrast, although displaying a lower Tm in comparison to H8-Omo, V105 (HisTag-SynB3-OmoNLS4) presented a relatively cooperative transition.

This could be due to the fact that the NLS4 insertion substituted a part of Omomyc's LZ domain, and its high content in basic residues may create destabilizing electrostatic repulsions at the dimer interface. However, the presence of a glycine at the beginning of the NLS4 stretch mutation might allow enough structural flexibility to avoid forcing a close proximity of the basic patch in each monomer, therefore allowing a partial LZ formation, resulting in an overall lower α-helix content and Tm, but not a loss in transition cooperativity, which would support an intact HLH domains interaction.

Similarly, an explanation for the destabilization of the V99 and V100 structures (the OmoEGFR variants) may also reside in the composition of the EGFR-targeting sequence (DSAQFWLHH). This sequence of 9 amino acids contains 1 Ala, 1 Phe, 1 Trp, 1 Leu, and 2 His residues, which would give it a predominantly hydrophobic nature, particularly at neutral pH (or near neutral pH, as in the case of the CD experiments), where the His residues are likely to be mostly deprotonated. Since this tag is at a position relatively exposed to the solvent, its hydrophobic side chains have no putative hydrophobic pocket to "hide" from the aqueous environment, and could form aberrant but favorable structures, which in turn would destabilize the rest of the structure and/or lead to aggregation. In fact, V99 and V100 were the variants with the lowest residual ellipticity at high temperatures compared to H8-Omo, which could be simply explained by a tendency to precipitate in the CD experimental conditions, reducing the protein soluble concentration and therefore the signal acquired. Following this hypothesis, Omo-EGFR variants would therefore have an overall structure that is very similar to the one of V84 or V86 (their parent sequence) but with a higher tendency to aggregate at RT, which would explain their thermal denaturation curve similarity in terms of Tm and cooperativity, as well as their differences in signal intensity.

Finally, rather than a higher tendency to aggregate, the observed OmoNLS2 folding destabilization is more likely to originate from an intramolecular destabilization of their respective HLH domain. Indeed, the MYC loop central region, mainly constituted by polar and acidic residues (ELENNE) was replaced by a highly basic sequence (KKKRKV) in those variants. The first Glu residue of the original loop replaced sequence is shown to form an attractive electrostatic interaction with a Lys residue in the Helix-2 domain (cf. 5I50 and 5I4Z PDB entries) which probably contributes to the

stabilization of the HLH domain folding. Its substitution with another Lys in the Omo-NLS2 sequence therefore changes the nature of this electrostatic interaction to a repulsive one, destabilizing the HLH folding. As a consequence, whereas the apparent lack of cooperativity from the OmoEGFR-derived variants during thermally-induced structural transitions could be explained by a higher tendency to aggregate, the lack of cooperativity observed in V102 and V103 transitions is more likely to reflect their monomeric nature.

For what refers to the results obtained in the Max*-binding CD experiments, they are in line with the hypotheses previously made to explain each variant's folding destabilization. Indeed, the OmoCrot and OmoNLS2 variants, with a mutated loop most likely affecting their respective HLH domain stability and therefore the stability of their dimeric state, also seemed to be impaired in their ability to bind the MAX b-HLH-LZ domain. Indeed, no clear signs of Max* binding were observed for the OmoCrot variants and only partial Max* heterodimerization occurred with OmoNLS2 variants, even if present in 2:1 molar excess and at micromolar concentrations. These findings support the idea that a correctly folding MYC HLH is particularly important for efficient binding of the MAX b-HLH-LZ, and that LZ complementarities would not be sufficient by itself to stabilize the putative heterodimer.

In contrast, the three C-terminus-mutated variants display an efficient Max* binding. The Tm of the V99/Max* and V100/Max* heterodimers are increased of about 20 ºC compared to Max*'s Tm, similarly to what was observed when Max* was placed in the presence of Omomyc or H8-Omo. Although notable, V105 heterodimerization with Max* resulted in a less pronounced structural stabilization, both in terms of α-helix content at low temperatures and Tm. This can be probably explained by the fact that the Max* LZ cannot fold completely while interacting with V105, due to the OmoNLS4 C-terminal stretch mutation. Altogether, these results indicate that the variants V99 and V100 may be the only variants of this series presenting a mutated Omomyc moiety with the potential to bind MAX at physiological temperatures.

While seemingly not causing any major change in the rest of the variants structure, both the HisTag and the CPP tag present at each construct N-terminus were observed to cause dramatic impairment in specific DNA binding. Although we had originally hypothesized that the CPP tag would be

the main responsible of the complex structural events observed in our initial DNA binding experiments, the characterization of H8-Omo proved us wrong. Indeed, in the usual CD experiment buffer conditions (50 mM potassium phosphate, 50 mM KCl, 2 mM TCEP, pH 6.7), the introduction of DNA probes in H8-Omo led to rapid sample precipitation suggesting that the HisTag alone may bind non-specifically to DNA, leading to the precipitation of the whole complex. These non-specific interactions are most likely of electrostatic nature, since elevated salt concentrations were found to rescue specific H8-Omo-E-Box interaction. Yet, since V84 and V86 were still showing signs of aggregation followed by refolding in these elevated salts condition, we could conclude that both SynB3 and Penetratin also perturbed the variants specific binding to DNA, perhaps synergistically with the HisTag, the basic nature of these CPP being undoubtedly at the source of this observation. The refolding observed at around 35 ºC for both V84 and V86 in the presence of DNA suggests that the CPP/HisTag interactions with DNA are relatively weak in comparison to Omomyc-DNA specific interactions, but sufficient to potentially impair the initial contact of Omomyc's BR with DNA, that would precede an induced fit process leading to the folding stabilization of this domain. Interestingly, temperatures below 30 ºC were also found to stabilize these non-specific interactions, while incubation at 37 ºC for 1h allowed the E-Box DNA probe to interact specifically with the Omomyc BR. These findings suggest that these non-specific DNA interactions are enthalpically favorable but do not particularly benefit from a strong entropic contribution. The two CPPs of opposed monomer could interact with two distant DNA sequence backbone creating a structural constraint that would lead to an overall destabilization of the homodimer structure. In contrast, specific DNA binding leading to a structurally tighter interaction of the Omomyc variant BR would benefit from a strong entropic contribution, naturally favored by an increase in temperature.

While in these newly defined conditions the presence of the canonical E-box sequence has a stabilizing effect on V84 and V86 structures, we found that the presence of a non-specific DNA sequence had the opposite effect, inducing the aggregation of those variants as well as a slightly lower Tm than in the absence of DNA. One can imagine that, when sequence specific induced fit DNA binding events are not allowed, the Omomyc BR artificially prolongated by the presence of the SynB3 and Penetratin sequences

remains stuck in an unnatural conformation, destabilizing the adjacent HLH and LZ by extension.

In contrast to their respective parent peptides, the four loop-mutated variants were found to heavily precipitate in the presence of DNA, regardless of the sequence. This phenomenon can probably be explained by the fact that they have lost the ability to homodimerize efficiently, rendering specific DNA binding impossible. Moreover, the introduction of basic residues in both the OmoCrot and the OmoNLS2 loop sequence further strengthens the chances of these constructs interacting non-specifically with DNA.

Finally, while the OmoEGFR- and OmoNLS4-derived variants were also found to precipitate in the presence of the non-specific DNA probe, hints of partial refolding and specific binding were observed in the presence of the E-box DNA probe, as a sign that some structural integrity was conserved in those variants, despite being made more fragile by their respective LZ mutation.

Although we could determine conditions limiting the non-specific interaction of the HisTag and CPP tags to DNA, the evaluation of the actual contribution of each alternative Omomyc moiety to DNA binding events remains challenging. To that end, the five Omomyc-derived sequences should ideally be studied without any tag. Moreover, alternative techniques should be used to complement this structural study solely relying on CD data. Indeed, although powerful, CD methods do not allow for distinguishing between signals from different molecular populations or identifying with certainty the contributions of individual protein domains to the overall signal. From that perspective, the characterization of these Omomyc variants by NMR could be valuable to complete this dataset, although the obtention of samples with sufficient concentration could be particularly challenging for some variants.

Concomitantly to the characterization by CD, we studied the cell penetration of the variant panel first by IF and then through a flow cytometry-based assay. To ensure a certain comparability, we took full advantage of the HisTag displayed by each construct by using anti-HisTag antibodies.

Anti-HisTag IF of A549 NSCLC cells treated with 10 µM of each variant was informative at several levels. First, we could confirm that all the tested variants conserved the ability to penetrate cells, and mainly localized within the cytoplasm. Again, the SynB3 and Penetratin-fused variants were found

to yield very similar results. In these experiments, OmoEGFR, OmoNLS2 and OmoNLS4-derived variants seemed slightly more intra-cellular than V84 or V86, whereas the OmoCrot variants (V96 and V97) displayed poorer cell penetration properties than their parent peptide and were mainly detected as concentrated dots within the cells, strongly suggesting that these variants remained trapped within endosomes, after 4 hours of treatment.

Moreover, we could not observe clear enhancement of the nuclear localization for any of the variants with a mutated Omomyc moiety. On the contrary, the precipitation patches visible in all the slides with variant-treated cells further supported our initial impression that the Omomyc variants of this set displayed limited solubility at neutral pH. Precipitation patches were particularly evident in OmoEGFR and OmoNLS2 variants-treated samples, although we cannot exclude that the precipitated variants in other slides might just have been more efficiently washed away throughout the various washing steps of the immunostaining protocol.

Although some variants seemed to display slightly superior cell penetration properties than others, a quantitative evaluation of their respective cell transduction ability was lacking. That is why we decided to use a flow cytometry-based assay.

Interestingly, in this assay, the HisTag of H8-Omo was shown to enhance the cell penetration of Omomyc by itself. Although HisTags are not particularly renowned for their cell penetration properties, this observation is not entirely surprising, since a study demonstrated that His residues placed in the vicinity of a natural Lys-rich cell penetrating peptide could enhance the efficiency of its membrane translocation into cells[257]. In addition, other studies suggested that the protonation of the imidazole group in His side chain in endosome acidic conditions could lead to more efficient endosomal escape[258,259]. Finally, Histidine-rich motifs were identified as nuclear localization signal in several human proteins[260]. Our experience further supported the latter point, in the development phases of this assay, we had observed that using a harsh permeabilization buffer enabling the labeling of both cytosolic and nuclear proteins would result in much higher HisTag staining background signal than using a milder buffer, which mainly limited HisTag detection to the cytosolic fraction. Of note, since little HisTag nuclear staining was observed by IF in the non-treated conditions, we could conclude that the cell permeabilization step in that assay might have not allowed the efficient staining of intra-nuclear variants.

In any case, our flow-cytometry assay confirmed the intracellular presence of all variants after only 15 min incubation. As described in previous characterization of Omomyc[160] and other CPPs[222,256], the variants' cell transduction ability rely on ATP-dependent mechanisms, abrogated by incubations at 4 ºC. This study also demonstrated that OmoEGFR variants had significantly better cell penetration properties than H8-Omo and accumulated more efficiently into A549 cells after a 15 min treatment, although this difference was less marked after 1h-incubations. This difference of accumulation kinetics observed between H8-Omo and the variants V84 and V99 suggests that the fusion to SynB3 might change the mini-proteins entrance route to the cells. Alternatively, the difference could be explained by V84 and V99 altered solubility: during the first minute of treatment, the SynB3 tag would allow more efficient cell transduction, but after some minutes, the soluble fractions of V84 and V99 could have been significantly reduced due to precipitation, explaining why we only observed a marginal intracellular accumulation increase between the 15 min and 1h time points.

Considering that both V99 and V100 displayed significantly improved cell penetration, further investigation of the entry mechanisms of the OmoEGFR variants should be considered. EGFR expression levels in A549 cells are typically considered as intermediate[261]. Repeating this study with cell lines expressing high and low levels of EGFR would help understanding whether the improvement in OmoEGFR-variants cell penetration is actually mediated through the interaction with EGFR itself. Moreover, these experiments could be repeated in the presence of different endocytosis pathways inhibitors, to study whether the fusion of Omomyc to a CPP tag could tilt its preference to a particular penetration pathway.

Overall, this short study demonstrated that the cell penetration of Omomyc could indeed be modulated by its fusion to cell penetration and appropriate short stretch mutations. Nevertheless, this enhancement came at the cost of structural stability and solubility. From a therapeutic point of view, improving the cell transduction kinetics of Omomyc-based proteins could be a double-edge sword. On the one hand, improving its cell penetration efficiency could increase its accumulation in the tumor cells and therefore a higher anti-tumoral activity. On the other hand, it could enhance overall non-tumor tissue absorption and possibly even lead to faster clearance, rather than tumor accumulation. Unfortunately, we know very little about

the “phylomer” peptide used as a fusion protein with Omomyc by Wang and colleagues: establishing a link between higher efficacy and a particular cell penetration path could be highly valuable.

It should be stated that assays testing cell penetration, and the cell localization of a molecule may be most valuable when one needs to understand a dramatic change of activity, good or bad, from one molecule to another. Yet, since the b-HLH-LZ structure and functions are highly intertwined, slight changes in cell penetration might not necessarily translate into a change of activity and enhanced cell penetration could come at the cost of target binding efficiency or stability within the cells. Illustratively, neither IF nor flow cytometry give information about the activity of a variant once found in the cell, where it could remain trapped in non-intended complexes or could get partially degraded, still allowing its detection but abrogating its biological activity. Therefore, as a means to an end, cell penetration is probably not the best early criterion for the determination of the potential of a drug candidate. Instead replacing cell penetration assays with cell-based functional reporting assay may be the best alternative for b-HLH-LZ-based drug screening, making the development of such assays highly desirable. For instance, one can imagine the development of a cancer cell line with a bicistronic expression of MYC and MAX constructs each fused to a fluorescent protein which would allow to obtain a FRET signal upon their heterodimerization. The evolution of this signal could be monitored upon cell treatment with Omomyc-derived variants. Therefore, a decrease in FRET signal would be indicative of the disruption of the MYC-MAX heterodimers. Alternatively, the expression of a reporting gene (i.e. coding for a luciferase or a fluorescent protein) could be placed under the control of a minimal promoter located downstream of E-Box sequences. In that case, the efficient invasion of E-Boxes by Omomyc-derived proteins would reduce the expression of the reporter gene. A decrease in signal in both strategies would be indicative of an intracellular target engagement as well as efficient cell penetration.

As a tool for purification, the use of HisTag probably spared a significant amount of troubleshooting for the obtention of pure stocks for multiple constructs. In IF and flow cytometry-based assays, it enabled an equivalent detection of all the variants without having to conjugate them individually with fluorescent labels. This point is rather important, since labeling cost and variable labeling efficiency and solubility can become serious pitfalls in

the development of such assays. However, we demonstrated that fusing a HisTag to the Omomyc N-terminus could greatly impair its capacity to bind DNA in a sequence-specific manner. Moreover, unshown data suggested that fusing a HisTag to the Max* C-terminus led to a significant drop in this construct's solubility compared to the rather soluble Max* mini-protein. Therefore, the use of HisTag was judged to be incompatible with the discovery of b-HLH-LZ-based MYC inhibitors deriving from Omomyc's modes of action.

Section 2: Impact of the net charge on the Omomyc-derived mini-proteins mode of action.

Perhaps due to lesser academic interest, the solubility of therapeutic protein candidates is not often mentioned, let alone published. Yet, appropriate solubility is of major importance during the development of a drug candidate, especially for those that cannot be administered orally, as is the case for most therapeutic proteins. Based on our group's experience with MYC and MAX b-HLH-LZ-derived molecules, the solubility of b-HLH-LZ can be rather puzzling and sometimes counterintuitive. For instance, the b-HLH-LZ domain of MYC is notoriously challenging to obtain recombinantly and to work with, primarily because of its high tendency to aggregate, whereas Omomyc, whose sequence shares more than 95% identity with the MYC b-HLH-LZ domain, exhibits a much greater solubility profile and stability, likely due to its ability to homodimerize. Also, Omomyc was found to have the best solubility in acetate buffers, yet by CD analysis, we observed a more folded structure in phosphate buffers. Thus, while conformational stability seems important, it does not directly correlate with colloidal stability.

The charges at a protein's surface are well-known to be important factors in their interaction with surrounding molecules and are consequently crucial for their solubility. Although the significant role of basic residues in Omomyc's folding and DNA-binding capacity is well understood, their impact on Omomyc's solubility profile is yet to be studied. Moreover, a neutral net charge at physiological pH has been shown to correlate with a longer half-life of monoclonal antibodies (mAbs) *in vivo*, as discussed in section I.3.3.3. Because of its basic residue-rich sequence giving it an estimated net charge of +5.2 at pH 7.4, we hypothesized that the Omomyc mini-protein could be a good protein model to study whether net charge modulation could impact the PK of a class of protein other than antibodies. We therefore designed a total of six Omomyc-derived charge variants with the primary goal of reducing their net charge and understanding the impact of basic residue mutations on Omomyc's mode of action.

To study these new sequences, the use of StrepTag was explored as an alternative to HisTag using three Omomyc control variants displaying a StrepTag at either their N-terminus (SLO and SO) or C-terminus (OLS). A Gly-Ser linker (GGGS) was also inserted between the StrepTag and the Omomyc

sequence in SLO and OLS, with the idea that its flexibility would allow the two parts of the constructs to function independently. All three constructs were rapidly purified using a urea-based extraction protocol and StrepTag affinity chromatography, which showed similar performance to HisTag affinity chromatography, though perhaps with a more limited binding capacity. Moreover, all three control constructs demonstrated the ability to bind the Max* protein as well as the canonical E-box DNA sequence. However, in complex with DNA, the OLS variant exhibited slightly superior thermal stability compared to the other two constructs, suggesting that the StrepTag placed at Omomyc's N-terminus may interfere with DNA binding, albeit to a much lesser extent than the HisTag, as observed with H8-Omo. Interestingly, the presence of a Gly-Ser linker at the N-terminus of Omomyc was found to destabilize its structure: the SLO variant displayed a significantly higher random coil content compared to the other two constructs. Additionally, all three constructs retained the ability to penetrate cells, although with slightly lower efficiency than Omomyc. Altogether, these results suggest that any of the three tag positions could be a viable option for further studies, since none of them critically affected Omomyc's mode of action. The model of OLS was chosen and applied to the charge variants to avoid any potential interference during DNA binding events.

Then, the six Omomyc charge variants were successfully expressed in *E. coli* and purified, without further optimization of the protocols established for the obtention of the StrepTag control variants.

Before completing the purification process, we used the denatured pool of purified variants to perform a rapid formulation assay, evaluating the viability of a panel of five simple buffers encompassing a broad range of pHs and different buffering species. The solubility of a given protein under specific conditions can be roughly assessed by the shape of its elution peak during a buffer exchange experiment. A smooth transfer of a protein to a new buffer is typically indicative of good solubility under the new buffer conditions, which is reflected in a narrow and symmetrical elution peak. In contrast, limited solubility in the new conditions is characterized by a tailing elution peak, caused by a series of partial protein precipitation in the new buffer, and subsequent resolubilization by the advancing former buffer salts front within the desalting column.

Overall, the results obtained with the Omomyc charge variants in this assay demonstrated that positive charge substitutions can significantly affect the solubility profile of Omomyc-based proteins. The variants with a negative net charge at pH 7.4, specifically V111, V112, and V113, displayed improved solubility at basic pH compared to the other variants, including OLS. In fact, the variant with the lowest pI (V113) exhibited poorer solubility in acidic buffers than in neutral or alkaline conditions, regardless of the buffering species or salt concentration. Notably, this particular variant is the first b-HLH-LZ protein we have studied that displayed such a solubility profile. In contrast, V108 showed very poor solubility in any buffer that was not acidic: the two interfacial basic residues substituted in the V108 sequence likely cause electrostatic repulsion, which may contribute to Omomyc's colloidal stability and/or support the LZ structural stability through electrostatic interactions. The solubility profile of this variant thus illustrates the important stabilizing role of these residues in the Omomyc sequence. For simplicity, and because V108 displayed a solubility too limiting to be studied in most of the subsequent assays, any following general mention of the charge variant panel will not refer to V108, unless specified otherwise. Intriguingly, the variant V111, which shares the same mutations as V108 along with other similar substitutions, did not display such limited solubility at neutral pH (judging by the results obtained in this assay and subsequent CD characterization) and actually demonstrated improved solubility at basic pH compared to OLS. The case of V111 highlights the limitations of this assay: although it displayed symmetrical elution peaks in neutral pH buffers, normally indicative of good solubility, the corresponding collected fractions became cloudy after a few minutes, which is a clear signs of protein precipitation. This phenomenon was not observed at basic or alkaline pH suggesting that it is induced by the formation of aberrant salt bridges leading to protein aggregation events.

Originally conceived and designed to facilitate the selection of optimal storage and experimental conditions for studying this variant panel, the results obtained from this assay demonstrated a certain predictive value based on observations made during their subsequent characterization. However, this predictive value was more limited when experiments were conducted in more complex matrices such as cell culture media or blood-derived samples. Indeed, this formulation assay only gives information on solubility at "t_0" and RT, in simple buffers. Notably, in IF experiments, OLS was seen to display a lower solubility than V109, V110 or V112 in cell culture

media although the results obtained in this formulation assay with neutral buffers suggested the contrary. Performing this experiment with buffers that are more physiologically relevant, such as bicarbonate or HEPES-based buffers, might help in that regard. Also, in the current setup of this assay, the elution fractions corresponding to protein successfully buffer-exchanged are rather underexploited. For instance, these fractions could be incubated at 37 ºC and characterized by dynamic light scattering to follow their respective polydispersity and detect the formation of aggregates over time. Other parameters affecting protein solubility and stability in physiological contexts, such as degradation or interaction with other proteins, would however require a distinct setup to be studied.

As expected, substitutions of basic amino acids significantly impacted the Omomyc structure, as demonstrated by CD. All the charge variants showed a decrease in α-helix content at 5 ºC, as well as lower thermal stability. In general, the substitution of basic residues in the LZ of the variants likely caused destabilization by eliminating stabilizing attractive electrostatic interactions. Moreover, the introduction of acidic residues in some variants (as it is the case in V109, V111, V112 and V113) probably led to the formation of electrostatic repulsions instead. Illustratively, although the same residues were mutated to obtain both V110 and V111 sequences, V110 displayed higher α-helix content and Tm than V111, in which basic amino acids were substituted with acidic ones. The MAX-like mutation present in helix-2 of V112 and V113 (K49D) could be one of the main reasons for these variants' apparent behavior as monomers (judging by their monotonous transition to a denatured state), since this Lys residue is observed in Omomyc homodimer crystal structures to interact with the Glu 21 residue of the opposite monomer, thereby enhancing dimer stability. Notably, this study demonstrated that the highly cooperative structure of Omomyc can be easily destabilized by a few ill-advised point mutations.

Interestingly, despite showing signs of structural alteration, all the variants were observed to heterodimerize with Max* as well as full-length MAX. By CD, the thermal stability of Max* increased when paired with each of the charge variants, except for V111, where heterodimerization with Max* did not significantly affect the secondary structure of either protein, making this result somewhat ambiguous. The complementarity, and by extension the affinity, of each charge variant for the Max* sequence was likely affected by their respective charge substitutions, as the Tm determined for any of the

variant/Max* heterodimers was significantly lower than that obtained for the OLS/Max* heterodimer. In fact, the Tm of the heterodimers formed by Max* with V109, V110, or V111 is very similar to that of their respective homodimers. In contrast, the binding of Max* to V112 and V113 had a stabilizing effect on both dimer moieties, further supporting the hypothesis that these two charge variants are primarily monomeric under these experimental conditions, in the absence of Max*.

Binding to full-length MAX was demonstrated for all the charge variants, including V108 and V111, in a co-IP experiment performed in MYC-ER-expressing NHDF cell lysate. Notably, both the p21 and p22 MAX isoforms were co-precipitated by StrepTactin XT-functionalized beads in the presence of any charge variant. Interestingly, the interaction of the charge variants with full-length MYC was also demonstrated in the same experiment, although it appeared more clearly for OLS, V109, and V110, suggesting that these proteins might be the only ones having conserved significant affinity for MYC. Two MYC-related protein bands were systematically detected in the NHDF cell lysate. The band corresponding to the highest molecular weight (about 90 kDa) is most likely related to the MYC-ER construct exogenously expressed by this cell line. The origin of the second band, corresponding to a lower molecular weight (about 42 kDa), is more mysterious: the most common MYC isoforms have a theoretical molecular weight of about 50 kDa but typically migrate to positions corresponding to 65 kDa. Thus, the band observed in this assay could correspond to an N-terminally truncated MYC isoform or degradation product, which would have conserved its b-HLH-LZ domain and therefore its ability to interact with Omomyc-derived variants. Despite using a large theoretical bead excess, a significant amount of each variant was still found in the depleted lysate fractions, particularly in lysates spiked with 50 µM of Omomyc variant, which could indicate that the extent of MYC and MAX binding is underestimated in this assay.

Through this Co-IP assay, we could confirm the ability of each variant to bind both full length MYC and MAX in much more complex conditions than the ones used in CD experiments. Now that the basic protocol parameters are established, performing this assay with lysate obtained from treated cells would be a natural step forward, since it could allow to determine whether these variants can penetrate cells with sufficient efficiency to bind their target molecules.

We next focused our attention on the charge variants' ability to bind DNA and to discriminate cognate sequences from a non-specific one. First, all the charge variants displayed a higher α-helix content and a higher thermal stability in the presence of DNA, and this stabilizing effect was particularly pronounced in the presence of the canonical E-Box, clearly suggesting that these variants conserved the ability to bind DNA in a sequence specific manner. These observations also suggest that interactions to a non-specific DNA sequence would not allow the complete folding of the MYC BR, in contrast to the canonical E-Box sequence, since in that case, DNA binding would mainly rely on the electrostatic interactions established between basic residues side chain and the DNA phosphodiester backbone. Of note, no clear difference in DNA binding was observed between the variants displaying a mutation in their BR (namely, V110 to V113) compared to V109, which conserved the same BR as Omomyc/MYC. Interestingly, this Arg residue is not conserved in MAX (substituted by an Ala), MNT (substituted by a Glu) or the MXD protein family (substituted by a Ser), but the presence of a basic residue at this position is conserved in both MYCL and MYCN. Looking at the crystal structure of the Omomyc dimer bound to DNA, this Arg residue side chain does not seem to interact with DNA, unlike all the other basic residues of the BR, suggesting that it could be involved in early contacts with DNA but would not contribute to stabilizing an established interaction. By CD and EMSA, the charge variants were also seen to interact with DNA as heterodimers with Max*. In fact, the combined results of these assays indicated that each variant preferentially bound DNA as a heterodimer, in the presence of Max*.

Interestingly, the variants V112 and V113 exhibited marginal DNA binding when assessed by EMSA. However, their affinity for DNA as heterodimers with Max* is notably enhanced, as indicated by a significantly increased melting temperature (Tm) observed in CD experiments and a more intense shifted DNA band in EMSA. The discrepancy between the CD and EMSA results regarding the DNA binding ability of V112 and V113 in the absence of Max* can likely be attributed to differences in protein and DNA concentrations used in the assays. The more transient nature of their DNA binding as monomers may lead to their detachment from DNA during the gel running phase of the EMSA assay, which would explain the DNA smear visible above the free DNA band. Overall, the results obtained with this pair of charge variant illustrate the apparent correlation between binding affinity for DNA and Tm in the presence of DNA determined by CD. In fact, the

existence of such correlation was further supported by preliminary results obtained by ELISA: the DNA binding Kd determined for each variant followed the same trend as their respective Tm in the presence of DNA.

Following the assessment of the charge variants' ability to bind DNA, MYC, and MAX, we tested their capacity to inhibit cell proliferation with two ovarian cancer cell lines (A2780 and EFO-27) as well as with a control murine fibroblast cell line, BZ1. Due to solubility issues, variants V108 and V111 were excluded from this study. To our surprise, OLS also displayed a certain tendency to precipitate in cell culture media, while Omomyc remained soluble in the same conditions. Yet, in contrast to V111 that was observed to precipitate rapidly after its dilution into cell culture media, OLS aggregation became apparent only after a few days of treatment, suggesting that their respective paths to aggregation differ in nature. Indeed, on the one hand, the slow OLS aggregation could be induced by partial degradation, chemical modifications and/or the formation of low affinity complexes with FBS proteins. On the other hand, the rapid V111 precipitation in culture media probably originates from the same structural defects as the ones hypothesized earlier to explain its loss of solubility in neutral pH phosphate- or Trizma®-based buffers. Notably, the proliferation of control BZ1 cells was largely unaffected by Omomyc treatment and was only inhibited by OLS treatment. This suggests that BZ1 cells are particularly sensitive to protein precipitation, potentially due to their tendency to grow in a spread-out manner, requiring a large surface contact area with the culture plate, impaired in the presence of protein precipitates.

Both ovarian cancer cell lines responded to Omomyc or OLS treatments, but did not do so to the same extent. The A2780 cells proliferation was significantly reduced by a 20 µM treatment of Omomyc, OLS, V109 and V110. In contrast, Omomyc and OLS only had a modest 30% inhibitory effect over the EFO-27 metabolic activity and treatment with any of the charge variants yielded little to no inhibition of this cell line proliferation. Intriguingly, these cells seemed to respond to the treatments inhibiting cell proliferation by a significant change of phenotype: the cells treated with Omomyc and OLS (and to a lesser extent, with V109) grew in spheroid-like structures, in contrast to their usual growth as a spread monolayer. One can hypothesize that the formation of such compact structures might be a stress-response of the cells, which helps them to limit the diffusion of the mini-proteins present in the medium, therefore limiting their therapeutic

activity. However, a subsequent IF assay conducted with the two cancer cell lines in similar conditions allowed to reject the hypothesis that the difference of inhibitory effect observed across cell lines and treatments is linked to a difference in mini-protein cell penetration. Indeed, with the exception of V113, all the variants were observed to have conserved the ability to penetrate cells, regardless of the cell line. Moreover, all the cell penetrating charge variants, OLS and Omomyc displayed a similar cell localization in both cell lines. Explaining the difference of response to treatment between the A2780 and EFO-27 cell lines could therefore require a long and complex answer; MYC and MAX protein levels, mutational profile, or addiction levels to the MYC transcriptional activity could all be factors explaining this difference.

To summarize, we could demonstrate through a series of binding assays that each of the Omomyc charge variants conserved the ability to bind MAX, MYC and DNA in a sequence specific manner, despite having a destabilized structure. In particular, V109 and V110 were found to retain an affinity for DNA similar to the one of OLS, and the capacity to bind MYC and MAX efficiently. Moreover, these two variants demonstrated to retain the ability to penetrate cells by IF and seemed to be more soluble than OLS in cell culture media. Yet, most of the data obtained through the assays employed were mainly qualitative and subtle changes in affinity for MYC, MAX or DNA may have been missed and could explain the significant decreases in activity observed for both variants compared to OLS and Omomyc, particularly in inhibiting the growth of EFO-27 cells. Therefore, the further development of assays capable of quantitatively assessing the interaction with the Omomyc target molecules in conditions as physiologically relevant as possible would be highly desirable. In that regard, fluorescence polarization-based (FP) assays could be particularly relevant since molecular interactions are monitored in solution (in contrast to assays like EMSAs or ELISAs) and do not require the direct detection of the tested protein. In fact, the validity of this approach has been already demonstrated as several FP-based protein-DNA binding assays have been independently developed over the years, including with Omomyc, and other b-HLH-LZ constructs[12,160,262]. In a drug screening process, the possibility of adapting FP-based assay to a 384- and 1536-well plate format also constitutes a major advantage of these methods over lower-throughput assays.

Determining the PK profile of therapeutic molecules early in the drug discovery process can be critical for evaluating their viability as drug candidates. The PK profile provides essential information on how a drug is absorbed, distributed, metabolized, and excreted within the body. Understanding these parameters helps predict clinically important parameters such as the molecule's bioavailability, optimal dosage, and potential toxicity. Many pre-clinical PK studies rely on radiolabeled or fluorescently labeled versions of the molecule of interest giving them a distinct signal that can be easily detected and quantified. Yet the labeling process of recombinant proteins with rather hydrophobic small molecules can be troublesome, costly and can affect the protein's properties and distribution[263–265]. Importantly, differences in labeling efficiency as well as in residual concentration of non-conjugated label can limit the comparison of the results obtained for different molecules. Moreover, the detection of a fluorescent or radioactivity signal would not account for the integrity of the molecule of interest. Radioactive or fluorescent labeling could therefore represent an important pitfall during the development of a PK study. Alternatively, mass-spectrometry-based PK determination methods do not rely on any labeling and can provide valuable information on the detected protein sequence integrity. However, the development of such methods can be long and costly, requiring specific expertise and devices. Furthermore, the detection of distinct variants of a molecule can require specific development and potential optimization. In contrast, competition binding assays enable to determine the concentration of a molecule of interest by measuring its ability to compete with a known labeled ligand for binding to a common target. Therefore, the concentration of two distinct molecules can be determined through the same assay, as long as they can bind to the same target. Since a StrepTag was fused to each charge variant sequence, and its ability to interact with binding partners is likely to be largely independent of the fused sequence, we utilized this common feature to develop an assay capable of determining the charge variant concentration in complex biological samples. Through initial development steps, we could validate our experimental approach as well as the chosen commercial material (anti-StrepTag antibody, functionalized 96-well plate, GFP-StrepTag protein stock) and determine an OLS lower limit of detection (LLOD) in the low nanomolar range, in PBST. In these conditions, the employed anti-StrepTag antibody affinity is probably the main limiting factor, and therefore largely determines the observed LLOD. A determination of the wells binding

capacity through a GFP-StrepTag titration could help to confirm this hypothesis.

Although initially designed to enable PK studies in murine models, the scarcity of mouse blood-derived matrices motivated the use of blood samples derived from other species for several key assay development steps. Rather surprisingly, the assay did not show similar performances when done with mouse and rat sera compared to when OLS was spiked in human serum. Indeed, both maximal GFP intensity and the ability of OLS to compete for the plate were diminished in rat serum and even more evidently in mouse serum. In contrast, when the assay's performances were compared with OLS spiked in each of the three species plasma, no significant difference was observed. Therefore, the difference across the serum of the three species could be caused by molecules generated or activated during the blood clotting process and would interfere with the assay by causing the degradation of the GFP-Strep and potentially OLS, and/or by inhibiting their interaction with the anti-StrepTag antibodies in a direct or indirect manner. One could even imagine that the partial degradation of the GFP-StrepTag leads to the formation of a StrepTag-containing digestion product efficiently competing with both full length GFP-Strep and OLS for the plate. Overall, these results indicated that mouse plasma interfered significantly less on the results obtained through this assay compared to mouse serum. Moreover, human or rat plasma could be an acceptable substitute for mouse plasma for assay development.

Importantly, whereas OLS and the charge variants equally competed with GFP-StrepTag for binding to the functionalized plate when diluted in PBST (unshown data), displaying an IC50 of about 5 nM, this assay had a lower sensitivity for their detection when they were spiked in plasma. The loss of sensitivity was however variant-dependent, with OLS displaying the worst IC50 in plasma, 17.5-fold higher than the value determined for V113, the variant for which the assay displayed the best sensitivity in these conditions, and about 50-fold higher than the OLS IC50 value in PBST. Importantly, the maximum GFP fluorescence intensity (in the absence of competitor) was not impacted by its dilution in plasma, as compared to the results obtained in PBST, suggesting that the observed matrix effect is mediated through an alteration of the variants capacity to bind the plate. Moreover, changing the variants incubation time in plasma before adding the sample to the plate did not significantly and consistently change the results, disproving the

hypothesis of a potential variant digestion or precipitation. Altogether, these results allowed us to safely conclude that the observed plasma matrix effect is most likely due to interactions between some of its components and the Omomyc-derived variants.

The possible implications of these observations could be of importance to better understand the different dynamics of these variants in plasma: if one imagines a simple model where OLS complexed to plasma macromolecules cannot interact with the functionalized plate, these results would then suggest that only 2% of the OLS mini-protein retain a free form when spiked in plasma (in reality, some complexed OLS form could be expected to retain a certain ability to bind the plate, making this number even lower). Following this hypothesis and considering the fact that the extent of the observed loss of sensitivity seems correlated to the variants pI (Figure 54), we could assume that the putative complexes formation is stabilized through attractive electrostatic interactions mediated by the Omomyc basic residues.

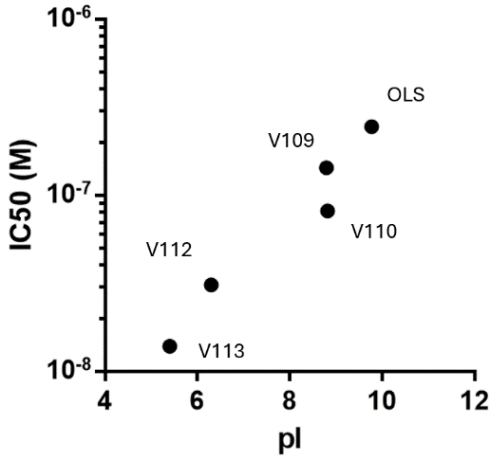


***Figure 54**, Omomyc-derived variants GFP fluorescence intensity IC50 value in mouse plasma (y-axis, logarithmic scale) in function of their respective pI (x-axis).*

To gain a better understanding of the variants' detectability in conditions close to the ones of an actual PK study, we then proceeded to analyze mouse plasma samples obtained after processing whole blood samples spiked with known concentrations of Omomyc variants. The results of this new study showed that the difference in the variants detectability observed in plasma were exacerbated when they were spiked in whole blood. Indeed, OLS became detectable through this assay only when micromolar

concentrations were added to whole blood samples, while the further loss of signal was more modest for the charge variants, especially for V113. Of course, the whole blood composition is more complex than plasma, meaning that there are more opportunities for the variants to form complexes with its different components. For example, interactions with blood cells could explain their partial clearance from the plasma fraction after processing. In these conditions, the LLOD for OLS would however be too limiting for PK characterization, since, based on previous experiments, we would not expect to detect micromolar concentrations for more than a few hours after intravenous injections at therapeutically effective doses. In those conditions, this assay is however informative on the state of the variants in blood at a given moment, which is relevant since one could expect them to have a higher propensity to perform their biological function and transduce into cancer cells in a free form, rather than in complex with large blood molecules. Nevertheless, transient complex formation could also be beneficial for the activity of Omomyc and its derivatives *in vivo* by reducing their renal clearance and limiting their degradation by circulating proteases. The significant difference in results obtained through this assay with the different charge variants is likely to illustrate a change of behavior in blood, seemingly well correlated with their respective net charge in these conditions. Determining whether this change of behavior could contribute to improve their anti-tumoral activity or not would be highly valuable to better characterize the PK/PD properties of b-HLH-LZ proteins, helping the design of Omomyc-derived variants with improved activity.

Although causing a pronounced hemolytic effect on the samples, the variants detection was overall improved by the presence of salts and detergent, particularly at the highest concentration points, further supporting the hypothesis that the variants inability to bind the plate is, at least partially, caused by the formation of transient complexes with blood components.

Overall, this assay proved to have the potential to quantify nanomolar concentration of StrepTag displaying proteins in complex matrices like plasma. The scarcity of mouse blood-derived matrices was a major limitation during the development of this assay and an adaptation of the existing protocol to a 384-well plate format would help reducing the amount of plasma needed in each assay, accelerating troubleshooting, and would allow to further increase the results reliability, allowing the simultaneous

analysis of more technical replicates with the same starting sample volume. In the future, we will focus our efforts on optimizing the variants conditions extractions, with the aim to abrogate their interaction with whole blood components and perform a PK study which will hopefully help us to understand the relationship between net charge and PK profile. A subsequent adaptation of this assay to cell lysate or lysed tissue samples would be also highly desirable.

Although the work presented in this section may have not led to the discovery of an Omomyc variant with overall improved properties, it contributed to develop methods and set the bases for systematic evaluation of future designs. First, the use of StrepTag was key for rapid and straightforward purification of a panel of variants but was also at the center of the design of key assays allowing to evaluate the variants capacity to bind full length MYC and MAX and penetrate cells, as well as to determine their concentrations and study their dynamics in complex biological matrices. Second, we determined a method to quickly evaluate protein solubility in various conditions which allowed to identify a defective sequence at very early stages of the project. Third, the variants' secondary structure and ability to bind Omomyc target molecules was assessed through a panel of assays which underlined a certain correlation between structural stability and DNA binding efficiency. Forth, contrasted efficacy results highlighted the need for the development of quantitative assays to assess the variant binding efficiency to Omomyc's target as well as the need to correlate these values with *in vitro*/*in vivo* efficacy data. Fifth, we established the bases of an assay with the potential to determine the concentration of protein fused to a StrepTag and this assay highlighted that net charge modulation could significantly impact the variants dynamics in blood-derived matrices.

Yet, due to the intricate mode of action of Omomyc, largely resulting from the overwhelming complexity of MYC biological activity, there is still much that needs to be understood. The precise determination of the conditions required for efficient MYC inhibition, the definition of acceptance criteria and the systematic characterization of Omomyc-derived proteins through quantitative assays are likely to be key concepts that will enable the design of a potent second generation of MYC inhibitors.

VI. Conclusions

Section 1

1. Fused to the N-terminus, the HisTag enables easy purification and detection of the Omomyc variants.
2. The fusion of SynB3 and Penetratin to Omomyc variants enhances their cell penetration activity upon short treatments.
3. When fused to the Omomyc variants N-terminus, both the HisTag and the CPPs impair the variants' ability to bind DNA in a sequence specific manner and limit their evaluation.
4. Short stretch mutations in Omomyc's loop domain impair its structural stability and ability to bind Max*.
5. The fusion of Omomyc to an EGFR-targeting sequence improves its cell penetration properties.

Section 2

6. Fused to the C-terminus, the StrepTag enables efficient purification and detection of the Omomyc charge variants and the development of meaningful comparative assays, while having an acceptable impact on their biological properties.
7. Charge mutation is an effective strategy to modulate the solubility profile of Omomyc-based mini-proteins.
8. Although having a destabilized structure, the charge variants conserve the ability to bind MYC, MAX and DNA in a sequence specific manner.
9. The charge variant ability to inhibit cell growth does not correlate with their ability to penetrate cells.
10. Basic residues substitution impacts the variants interactions with plasma and whole blood components.

VII. References

1. Roussel M, Saule S, Lagrou C, Rommens C, Beug H, Graf T, et al. Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. Nature. 1979;281(5731):452–5.
2. Sheiness D, Bishop JM. DNA and RNA from Uninfected Vertebrate Cells Contain Nucleotide Sequences Related to the Putative Transforming Gene of Avian Myelocytomatosis Virus. J Virol. 1979;31(2):514–21.
3. Hann SR, Eisenman RN. Proteins Encoded by the Human c-myc Oncogene : Differential Expression in Neoplastic Cells. Mol Cell Biol. 1984;4(11):2486–97.
4. Amati B, Dalton S, Brooks MW, Littlewood TD, Evan GI, Land H. Transcriptional activation by the human c-Myc oncoprotein in yeast requires interaction with Max. Nature. 1992;359(6394):423–6.
5. Amati B, Brooks MW, Levy N, Littlewood TD, Evan GI, Land H. Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell. 1993;72(2):233–45.
6. Blackwood EM, Eisenman RN. Max: A Helix-Loop-Helix Zipper Protein That Forms a Sequence-Specific DNA-Binding Complex with Myc. Science (80-). 1991;251(4998):1211–7.
7. Davis AC, Wims M, Spotts GD, Hann SR, Bradley A. A null c-myc mutation causes lethality before 10.5 days of gestation in homozygotes and reduced fertility in heterozygous female mice. Genes Dev. 1993;7(4):671–82.
8. Brooks TA, Hurley LH. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. Nat Rev Cancer. 2009;9(12):849–61.
9. Eilers M, Eisenman RN. Myc's broad reach. Genes Dev. 2008;22(20):2755–66.
10. Macek P, Cliff MJ, Embrey KJ, Holdgate GA, Nissink JWM, Panova S, et al. Myc phosphorylation in its basic helix–loop–helix region destabilizes transient α-helical structures, disrupting Max and DNA binding. J Biol Chem [Internet]. 2018 Jun;293(24):9301–10. Available

from:
https://linkinghub.elsevier.com/retrieve/pii/S0021925820389031

11. Beaulieu ME, McDuff FO, Frappier V, Montagne M, Naud JF, Lavigne P. New structural determinants for c-Myc specific heterodimerization with Max and development of a novel homodimeric c-Myc b-HLH-LZ. J Mol Recognit. 2012;25(7):414–26.

12. Maltais L, Montagne M, Bédard M, Tremblay C, Soucek L, Lavigne P. Biophysical characterization of the b-HLH-LZ of δ Max, an alternatively spliced isoform of Max found in tumor cells: Towards the validation of a tumor suppressor role for the Max homodimers. PLoS One. 2017;12(3):1–17.

13. McDuff FO, Naud JF, Montagne M, Sauvé S, Lavigne P. The Max homodimeric b-HLH-LZ significantly interferes with the specific heterodimerization between the c-Myc and Max b-HLH-LZ in absence of DNA: A quantitative analysis. J Mol Recognit. 2009;22(4):261–9.

14. Beaulieu M-E, Castillo F, Soucek L. Structural and Biophysical Insights into the Function of the Intrinsically Disordered Myc Oncoprotein. Cells [Internet]. 2020 Apr 22;9(4):1038. Available from: https://www.mdpi.com/2073-4409/9/4/1038

15. Sauvé S, Tremblay L, Lavigne P. The NMR solution structure of a mutant of the max b/HLH/LZ free of DNA: Insights into the specific and reversible DNA binding mechanism of dimeric transcription factors. J Mol Biol. 2004;342(3):813–32.

16. Nair SK, Burley SK. X-Ray Structures of Myc-Max and Mad-Max Recognizing DNA. Cell. 2003;112(2):193–205.

17. Fisher F, Goding CR. Single amino acid substitutions alter helix-loop-helix protein specificity for bases flanking the core CANNTG motif. EMBO J. 1992;11(11):4103–9.

18. Pellanda P, Dalsass M, Filipuzzi M, Loffreda A, Verrecchia A, Cano VC, et al. Integrated requirement of non-specific and sequence-specific DNA binding in MYC-driven transcription. EMBO. 2020;1–17.

19. Naud JF, Gagnon F, Wellinger R, Chabot B, Lavigne P. Improving the thermodynamic stability of the leucine zipper of Max increases the stability of its b-HLH-LZ:E-box complex. J Mol Biol.

2003;326(5):1577–95.

20. Sears R, Leone G, DeGregori J, Nevins JR. Ras enhances Myc protein stability. Mol Cell. 1999;3(2):169–79.

21. Sears R, Nuckolls F, Haura E, Taya Y, Tamai K, Nevins JR. Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. Genes Dev. 2000;14(19):2501–14.

22. Helander S, Montecchio M, Pilstål R, Su Y, Kuruvilla J, Elvén M, et al. Pre-Anchoring of Pin1 to Unphosphorylated c-Myc in a Fuzzy Complex Regulates c-Myc Activity. Structure. 2015;23(12):2267–79.

23. Richards MW, Burgess SG, Poon E, Carstensen A, Eilers M, Chesler L, et al. Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. Proc Natl Acad Sci U S A. 2016;113(48):13726–31.

24. Schütz S, Bergsdorf C, Hänni-Holzinger S, Lingel A, Renatus M, Gossert AD, et al. Intrinsically Disordered Regions in the Transcription Factor MYC:MAX Modulate DNA Binding via Intramolecular Interactions. Biochemistry. 2023;

25. Conacci-Sorrell M, McFerrin L, Eisenman RN. An overview of MYC and its interactome. Cold Spring Harb Perspect Med. 2014;4(1).

26. Wei Y, Resetca D, Li Z, Johansson-Åkhe I, Ahlner A, Helander S, et al. Multiple direct interactions of TBP with the MYC oncoprotein. Nat Struct Mol Biol [Internet]. 2019;26(11):1035–43. Available from: http://dx.doi.org/10.1038/s41594-019-0321-z

27. Chacón Simon S, Wang F, Thomas LR, Phan J, Zhao B, Olejniczak ET, et al. Discovery of WD Repeat-Containing Protein 5 (WDR5)-MYC inhibitors using fragment-based methods and structure-based design. J Med Chem. 2020;63(8):4315–33.

28. Kurland JF, Tansey WP. Myc-mediated transcriptional repression by recruitment of histone deacetylase. Cancer Res. 2008;68(10):3624–9.

29. Mark A. Gregory, Stephen R Hann. c-Myc proteolysis by the ubiquitin-proteasome pathway:stabilization of c-Myc in Burkitt's lymphoma cells. Mol Cell Biol. 2000;20(7):2423–35.

30. Thomas LR, Foshage AM, Weissmiller AM, Popay TM, Grieb BC, Qualls SJ, et al. Interaction of MYC with host cell factor-1 is

mediated by the evolutionarily conserved Myc box IV motif. Oncogene [Internet]. 2016 Jul 7;35(27):3613–8. Available from: https://www.nature.com/articles/onc2015416

31. Cowling VH, Chandriani S, Whitfield ML, Cole MD. A Conserved Myc Protein Domain, MBIV, Regulates DNA Binding, Apoptosis, Transformation, and G 2 Arrest . Mol Cell Biol. 2006;26(11):4226–39.

32. Stone J, de Lange T, Ramsay G, Jakobovits E, Bishop JM, Varmus H, et al. Definition of Regions in Human c- myc that are Involved in Transformation and Nuclear Localization . Mol Cell Biol. 1987;7(5):1697–709.

33. Rosales T, Nie Z, Kapoor V, Casellas R, Knutson JR, Levens D. Partition of Myc into immobile vs. Mobile complexes within nuclei. Sci Rep. 2013;3.

34. Arabi A, Rustum C, Hallberg E, Wright APH. Accumulation of c-Myc and proteasomes at the nucleoli of cells containing elevated c-Myc protein levels. J Cell Sci. 2003;116(9):1707–17.

35. Rizzuti B, Iovanna JL, Neira JL. Deciphering the Binding of the Nuclear Localization Sequence of Myc Protein to the Nuclear Carrier Importin α3. Int J Mol Sci. 2022;23(23).

36. Ogawa H, Ishiguro KI, Gaubatz S, Livingston DM, Nakatani Y. A complex with chromatin modifiers that occupies E2f- and Myc-responsive genes in G0 cells. Science (80-). 2002;296(5570):1132–6.

37. Bédard M, Maltais L, Montagne M, Lavigne P. Miz-1 and Max compete to engage c-Myc: implication for the mechanism of inhibition of c-Myc transcriptional activity by Miz-1. Proteins Struct Funct Bioinforma [Internet]. 2017 Feb;85(2):199–206. Available from: https://onlinelibrary.wiley.com/doi/10.1002/prot.25214

38. Ascanelli C, Dahir R, Wilson CH. Manipulating Myc for reparative regeneration. Front Cell Dev Biol. 2024;12(March):1–18.

39. Kress TR, Sabò A, Amati B. MYC: Connecting selective transcriptional control to global RNA production. Nat Rev Cancer. 2015;15(10):593–607.

40. Bernabé-Rubio M, Ali S, Bhosale PG, Goss G, Mobasseri SA, Tapia-Rojo R, et al. Myc-dependent dedifferentiation of Gata6+ epidermal cells resembles reversal of terminal differentiation. Nat Cell Biol.

2023;25(10):1426–38.

41. Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell. 2006;126(4):663–76.

42. Varlakhanova N, Cotterman R, Bradnam K, Korf I, Knoepfler PS. Myc and Miz-1 have coordinate genomic functions including targeting Hox genes in human embryonic stem cells. Epigenetics and Chromatin. 2011;4(1):1–14.

43. Smith KN, Singh AM, Dalton S. Myc Represses Primitive Endoderm Differentiation in Pluripotent Stem Cells. Cell Stem Cell [Internet]. 2010 Sep;7(3):343–54. Available from: doi:10.1016/j.stem.2010.06.023

44. Chappell J, Dalton S. Roles for MYC in the establishment and maintenance of pluripotency. Cold Spring Harb Perspect Med. 2013;3(12):1–9.

45. Chappell J, Sun Y, Singh A, Dalton S. MYC/MAX control ERK signaling and pluripotency by regulation of dual-specificity phosphatases 2 and 7. Genes Dev. 2013;27(7):725–33.

46. Watt FM, Frye M, Benitah SA. MYC in mammalian epidermis: How can an oncogene stimulate differentiation? Nat Rev Cancer. 2008;8(4):316.

47. Sumi T, Tsuneyoshi N, Nakatsuji N, Suemori H. Apoptosis and differentiation of human embryonic stem cells induced by sustained activation of c-Myc. Oncogene. 2007;26(38):5564–76.

48. Thompson NL, Mead JE, Braun L, Goyette M, Shank PR, Fausto N. Sequential protooncogene expression during rat liver regeneration. Cancer Res. 1986 Jun;46(6):3111–7.

49. Bretones G, Delgado MD, León J. Myc and cell cycle control. Biochim Biophys Acta - Gene Regul Mech [Internet]. 2015;1849(5):506–16. Available from: http://dx.doi.org/10.1016/j.bbagrm.2014.03.013

50. García-Gutiérrez L, Bretones G, Molina E, Arechaga I, Symonds C, Acosta JC, et al. Myc stimulates cell cycle progression through the activation of Cdk1 and phosphorylation of p27. Sci Rep. 2019;9(1):1–17.

51. Mateyak MK, Obaya AJ, Sedivy JM. c-Myc Regulates Cyclin D-Cdk4 and -Cdk6 Activity but Affects Cell Cycle Progression at Multiple Independent Points. Mol Cell Biol. 1999;19(7):4672–83.

52. Steiner P, Philipp A, Lukas J, Godden-Kent D, Pagano M, Mittnacht S, et al. Identification of a Myc-dependent step during the formation of active G1 cyclin-cdk complexes. EMBO J. 1995;14(19):4814–26.

53. Pérez-Roger I, Solomon DLC, Sewing A, Land H. Myc activation of cyclin E/Cdk2 kinase involves induction of cyclin E gene transcription and inhibition of p27(Kip1) binding to newly formed complexes. Oncogene. 1997;14(20):2373–81.

54. Dyson N. The regulation of E2F by pRB-family proteins. Genes Dev. 1998;12(15):2245–62.

55. Nevins JR. The Rb/E2F pathway and cancer. Hum Mol Genet. 2001;10(7):699–703.

56. Wu S, Cetinkaya C, Munoz-Alonso MJ, Von Der Lehr N, Bahram F, Beuger V, et al. Myc represses differentiation-induced p21CIP1 expression via Miz-1-dependent interaction with the p21 core promoter. Oncogene. 2003;22(3):351–60.

57. Staller P, Peukert K, Kiermaier A, Seoane J, Lukas J, Karsunky H, et al. Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol. 2001;3(4):392–9.

58. Seoane J, Le H Van, Massagué J. Myc suppression of the p21Cip1 Cdk inhibitor influences the outcome of the p53 response to DNA damage. Nature. 2002;419(6908):729–34.

59. Abbastabar M, Kheyrollah M, Azizian K, Bagherlou N, Tehrani SS, Maniati M, et al. Multiple functions of p27 in cell cycle, apoptosis, epigenetic modification and transcriptional regulation for the control of cell growth: A double-edged sword protein. DNA Repair (Amst) [Internet]. 2018;69:63–72. Available from: https://doi.org/10.1016/j.dnarep.2018.07.008

60. Wong P-P, Miranda F, Chan K V., Berlato C, Hurst HC, Scibetta AG. Histone Demethylase KDM5B Collaborates with TFAP2C and Myc To Repress the Cell Cycle Inhibitor p21 cip (CDKN1A) . Mol Cell Biol. 2012;32(9):1633–44.

61. Jung P, Hermeking H. The c-MYC-AP4-p21 cascade. Cell Cycle. 2009;8(7):982–9.

62. Li HL, Dong LL, Jin MJ, Li QY, Wang X, Jia MQ, et al. A Review of the Regulatory Mechanisms of N-Myc on Cell Cycle. Molecules. 2023;28(3).

63. Sur S, Agrawal DK. Phosphatases and kinases regulating CDC25 activity in the cell cycle: clinical implications of CDC25 overexpression and potential treatment strategies. Mol Cell Biochem [Internet]. 2016 May 2;416(1–2):33–46. Available from: http://link.springer.com/10.1007/s11010-016-2693-2

64. Yap CS, Peterson AL, Castellani G, Sedivy JM, Neretti N. Kinetic profiling of the c-Myc transcriptome and bioinformatic analysis of repressed gene promoters. Cell Cycle. 2011;10(13):2184–96.

65. Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ, Ren B. A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proc Natl Acad Sci U S A. 2003;100(14):8164–9.

66. Campbell KJ, White RJ. MYC regulation of cell growth through control of transcription by RNA polymerases I and III. Cold Spring Harb Perspect Med. 2014;4(5):1–12.

67. Van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. Nat Rev Cancer. 2010;10(4):301–9.

68. Hsieh AL, Walton ZE, Altman BJ, Stine ZE, Dang C V. MYC and metabolism on the path to cancer. Semin Cell Dev Biol [Internet]. 2015;43:11–21. Available from: http://dx.doi.org/10.1016/j.semcdb.2015.08.003

69. Dong Y, Tu R, Liu H, Qing G. Regulation of cancer cell metabolism: oncogenic MYC in the driver's seat. Signal Transduct Target Ther [Internet]. 2020;5(1). Available from: http://dx.doi.org/10.1038/s41392-020-00235-2

70. Bettess MD, Dubois N, Murphy MJ, Dubey C, Roger C, Robine S, et al. c-Myc Is Required for the Formation of Intestinal Crypts but Dispensable for Homeostasis of the Adult Intestinal Epithelium. Mol Cell Biol. 2005;25(17):7868–78.

71. Konsavage WM, Jin G, Yochum GS. The Myc 3' Wnt-Responsive Element Regulates Homeostasis and Regeneration in the Mouse Intestinal Tract . Mol Cell Biol. 2012;32(19):3891–902.

72. Wang N, Yang T, Li J, Lei M, Shi J, Qiu W, et al. The expression and

role of c-Myc in mouse hair follicle morphogenesis and cycling. Acta Histochem [Internet]. 2012;114(3):199–206. Available from: http://dx.doi.org/10.1016/j.acthis.2011.04.009

73. Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, et al. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nat Cell Biol [Internet]. 2010 Mar 21;12(3):247–56. Available from: https://www.nature.com/articles/ncb2024

74. Venkateswaran N, Conacci-Sorrell M. MYC leads the way. Small GTPases. 2020;11(2):86–94.

75. Chavez MG, Buhr CA, Petrie WK, Wandinger-Ness A, Kusewitt DF, Hudson LG. Differential downregulation of E-Cadherin and desmoglein by epidermal growth factor. Dermatol Res Pract. 2012;2012.

76. Baudino TA, McKay C, Pendeville-Samain H, Nilsson JA, Maclean KH, White EL, et al. c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression. Genes Dev. 2002;16(19):2530–43.

77. Sawai S, Shimono A, Wakamatsu Y, Palmes C, Hanaoka K. Defects of embryonic organogenesis resulting from targeted disruption of the N-myc gene in the mouse. Development. 1993;117:1445–55.

78. Kokai E, Voss F, Fleischer F, Kempe S, Marinkovic D, Wolburg H, et al. Myc regulates embryonic vascular permeability and remodeling. Circ Res. 2009;104(10):1151–9.

79. Pelengaris S, Khan M. The many faces of c-MYC. Arch Biochem Biophys. 2003;416(2):129–36.

80. Dang C V. MYC on the path to cancer. Cell [Internet]. 2012;149(1):22–35. Available from: http://dx.doi.org/10.1016/j.cell.2012.03.003

81. Chen S, Su L, Qiu J, Xiao N, Lin J, Tan JH, et al. Mechanistic studies for the role of cellular nucleic-acid-binding protein (CNBP) in regulation of c-myc transcription. Biochim Biophys Acta - Gen Subj [Internet]. 2013;1830(10):4769–77. Available from: http://dx.doi.org/10.1016/j.bbagen.2013.06.007

82. He L, Liu J, Collins I, Sanford S, O'Connell B, Benham CJ, et al. Loss of FBP function arrests cellular proliferation and extinguishes c-myc expression. EMBO J. 2000;19(5):1034–44.

83. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, Da Costa LT, et al. Identification of c-MYC as a target of the APC pathway. Science (80-). 1998;281(5382):1509–12.

84. Sabò A, Amati B. BRD4 and MYC—clarifying regulatory specificity. Science (80-) [Internet]. 2018 May 18;360(6390):713–4. Available from: https://www.science.org/doi/10.1126/science.aat6664

85. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell [Internet]. 2011;146(6):904–17. Available from: http://dx.doi.org/10.1016/j.cell.2011.08.017

86. Coudé MM, Braun T, Berrou J, Dupont M, Bertrand S, Masse A, et al. BET inhibitor OTX015 targets BRD2 and BRD4 and decreases c-MYC in acute leukemia cells. Oncotarget. 2015;6(19):17698–712.

87. Wittig B, Dorbic T, Rich A. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. Proc Natl Acad Sci U S A. 1991;88(6):2259–63.

88. Brooks TA, Hurley LH. Targeting MYC Expression through G-Quadruplexes. Genes and Cancer. 2010;1(6):641–9.

89. Spiniello M, Steinbrink MI, Cesnik AJ, Miller RM, Scalf M, Shortreed MR, et al. Comprehensive in vivo identification of the c-Myc mRNA protein interactome using HyPR-MS. Rna. 2019;25(10):1337–52.

90. Leeds P, Kren BT, Boylan JM, Betz NA, Steer CJ, Gruppuso PA, et al. Developmental regulation of CRD-BP, an RNA-binding protein that stabilizes c-myc mRNA in vitro. Oncogene. 1997;14(11):1279–86.

91. Lemm I, Ross J. Regulation of c- myc mRNA Decay by Translational Pausing in a Coding Region Instability Determinant . Mol Cell Biol. 2002;22(12):3959–69.

92. Weidensdorfer D, Stöhr N, Baude A, Lederer M, Köhn M, Schierhorn A, et al. Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. Rna. 2009;15(1):104–15.

93. Liu L, Ouyang M, Rao JN, Zou T, Xiao L, Chung HK, et al. Competition between RNA-binding proteins CELF1 and HuR modulates MYC translation and intestinal epithelium renewal. Mol Biol Cell. 2015;26(10):1797–810.

94. Sampson VB, Rong NH, Han J, Yang Q, Aris V, Soteropoulos P, et al.

MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. Cancer Res. 2007;67(20):9762–70.

95. Cannell IG, Kong YW, Johnston SJ, Chen ML, Collins HM, Dobbyn HC, et al. p38 MAPK/MK2-mediated induction of miR-34c following DNA damage prevents Myc-dependent DNA replication. Proc Natl Acad Sci U S A. 2010;107(12):5375–80.

96. Kress TR, Cannell IG, Brenkman AB, Samans B, Gaestel M, Roepman P, et al. The MK5/PRAK Kinase and Myc Form a Negative Feedback Loop that Is Disrupted during Colorectal Tumorigenesis. Mol Cell [Internet]. 2011;41(4):445–57. Available from: http://dx.doi.org/10.1016/j.molcel.2011.01.023

97. Liao JM, Zhou X, Gatignol A, Lu H. Ribosomal proteins L5 and L11 co-operatively inactivate c-Myc via RNA-induced silencing complex. Oncogene. 2013;33(41):4916–23.

98. Zhou X, Hao Q, Liao JM, Liao P, Lu H. Ribosomal protein s14 negatively regulates c-Myc activity. J Biol Chem [Internet]. 2013;288(30):21793–801. Available from: http://dx.doi.org/10.1074/jbc.M112.445122

99. Dani C, Blanchard JM, Piechaczyk M, El Sabouty S, Marty L, Jeanteur P. Extreme instability of myc mRNA in normal and transformed human cells. Proc Natl Acad Sci U S A. 1984;81(22 I):7046–50.

100. Wang X, Cunningham M, Zhang X, Tokarz S, Laraway B, Troxell M, et al. Phosphorylation Regulates c-Myc's Oncogenic Activity in the Mammary Gland. Cancer Res [Internet]. 2011 Feb 1;71(3):925–36. Available from: https://aacrjournals.org/cancerres/article/71/3/925/575277/Phosphorylation-Regulates-c-Myc-s-Oncogenic

101. Farrell AS, Sears RC. MYC degradation. Cold Spring Harb Perspect Med. 2014;4(3):1–16.

102. Welcker M, Wang B, Rusnac DV, Hussaini Y, Swanger J, Zheng N, et al. Two diphosphorylated degrons control c-Myc degradation by the Fbw7 tumor suppressor. Sci Adv. 2022;8(4):1–12.

103. Kim SY, Herbst A, Tworkowski KA, Salghetti SE, Tansey WP. Skp2 regulates Myc protein stability and activity. Mol Cell. 2003;11(5):1177–88.

104. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature [Internet]. 2010 Feb;463(7283):899–905. Available from: https://www.nature.com/articles/nature08822

105. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet [Internet]. 2013 Oct 26;45(10):1127–33. Available from: https://www.nature.com/articles/ng.2762

106. Kalkat M, De Melo J, Hickman KA, Lourenco C, Redel C, Resetca D, et al. MYC deregulation in primary human cancers. Genes (Basel). 2017;8(6):2–30.

107. Schaub FX, Dhankani V, Berger AC, Trivedi M, Richardson AB, Shaw R, et al. Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. Cell Syst. 2018;6(3):282-300.e2.

108. Murphy DJ, Junttila MR, Pouyet L, Karnezis A, Shchors K, Bui DA, et al. Distinct Thresholds Govern Myc's Biological Output In Vivo. Cancer Cell [Internet]. 2008 Dec;14(6):447–57. Available from: http://dx.doi.org/10.1016/j.ccr.2008.10.018

109. Blain SW, Massagué J. Different sensitivity of the transforming growth factor-β cell cycle arrest pathway to c-Myc and MDM-2. J Biol Chem. 2000;275(41):32066–70.

110. Zhao K, Wang Q, Wang Y, Huang K, Yang C, Li Y, et al. EGFR/c-myc axis regulates TGFβ/Hippo/Notch pathway via epigenetic silencing miR-524 in gliomas. Cancer Lett [Internet]. 2017;406:12–21. Available from: http://dx.doi.org/10.1016/j.canlet.2017.07.022

111. Fernandez PC, Frank SR, Wang L, Schroeder M, Liu S, Greene J, et al. Genomic targets of the human c-Myc protein. Genes Dev. 2003;17(9):1115–29.

112. Allevato M, Bolotin E, Grossman M, Mane-Padros D, Sladek FM, Martinez E. Sequence-specific DNA binding by MYC/MAX to low-affinity non-E-box motifs. PLoS One. 2017;12(7):1–20.

113. Lin CY, Lovén J, Rahl PB, Paranal RM, Burge CB, Bradner JE, et al. Transcriptional amplification in tumor cells with elevated c-Myc. Cell. 2012;151(1):56–67.

114. Sarosiek K. Myc heals all (tumor) wounds. Sci Transl Med [Internet].

2017 Dec 13;9(420). Available from: https://www.science.org/doi/10.1126/scitranslmed.aar4434

115. Wolfer A, Ramaswamy S. MYC and Metastasis. Cancer Res [Internet]. 2011 Mar 15;71(6):2034–7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf

116. Liu H, Radisky DC, Yang D, Xu R, Radisky ES, Bissell MJ, et al. MYC suppresses cancer metastasis by direct transcriptional silencing of αv and β3 integrin subunits. Nat Cell Biol [Internet]. 2012 Jun 13;14(6):567–74. Available from: https://www.nature.com/articles/ncb2491

117. Duffy MJ, O'Grady S, Tang M, Crown J. MYC as a target for cancer treatment. Cancer Treat Rev [Internet]. 2021 Mar;94(5):102154. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0305737221000025

118. Lee HY, Cha J, Kim SK, Park JH, Song KH, Kim P, et al. C-MYC drives breast cancer metastasis to the brain, but promotes synthetic lethality with TRAIL. Mol Cancer Res. 2019;17(2):544–54.

119. Massó-Vallés D, Beaulieu ME, Jauset T, Giuntini F, Zacarías-Fluck MF, Foradada L, et al. MYC Inhibition Halts Metastatic Breast Cancer Progression by Blocking Growth, Invasion, and Seeding. Cancer Res Commun. 2022;2(2):110–30.

120. Evan GI, Wyllie AH, Gilbert CS, Littlewood TD, Land H, Brooks M, et al. Induction of apoptosis in fibroblasts by c-myc protein. Cell. 1992;69(1):119–28.

121. Hermeking H, Eick D. Mediation of c-Myc-Induced Apoptosis by p53. Science (80-) [Internet]. 1994 Sep 30;265(5181):2091–3. Available from: https://www.science.org/doi/10.1126/science.8091232

122. Eischen CM, Roussel MF, Korsmeyer SJ, Cleveland JL. Bax Loss Impairs Myc-Induced Apoptosis and Circumvents the Selection of p53 Mutations during Myc-Mediated Lymphomagenesis. Mol Cell Biol. 2001;21(22):7653–62.

123. Eischen CM, Weber JD, Roussel MF, Sherr CJ, Cleveland JL. Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. Genes Dev. 1999;13(20):2658–69.

124. Alt JR, Greiner TC, Cleveland JL, Eischen CM. Mdm2 haplo-

insufficiency profoundly inhibits Myc-induced lymphomagenesis. EMBO J. 2003;22(6):1442–50.

125. Schuster C, Berger A, Hoelzl MA, Putz EM, Frenzel A, Simma O, et al. The cooperating mutation or "second hit" determines the immunologic visibility toward MYC-induced murine lymphomas. Blood. 2011;118(17):4635–45.

126. Kim J, Gao P, Liu Y-C, Semenza GL, Dang C V. Hypoxia-Inducible Factor 1 and Dysregulated c-Myc Cooperatively Induce Vascular Endothelial Growth Factor and Metabolic Switches Hexokinase 2 and Pyruvate Dehydrogenase Kinase 1. Mol Cell Biol. 2007;27(21):7381–93.

127. Huang LE. Carrot and stick: HIF-α engages c-Myc in hypoxic adaptation. Cell Death Differ. 2008;15(4):672–7.

128. Dang C V. The Interplay Between MYC and HIF in the Warburg Effect. In: Cancer [Internet]. 2008. p. 35–53. Available from: http://link.springer.com/10.1007/2789_2008_088

129. Casey SC, Baylot V, Felsher DW. The MYC oncogene is a global regulator of the immune response. Blood. 2018;131(18):2007–15.

130. Li J, Dong T, Wu Z, Zhu D, Gu H. The effects of MYC on tumor immunity and immunotherapy. Cell Death Discov. 2023;9(1):1–10.

131. Kortlever RM, Sodir NM, Wilson CH, Burkhart DL, Pellegrinet L, Brown Swigart L, et al. Myc Cooperates with Ras by Programming Inflammation and Immune Suppression. Cell [Internet]. 2017;171(6):1301-1315.e14. Available from: http://dx.doi.org/10.1016/j.cell.2017.11.013

132. Casey SC, Tong L, Li Y, Do R, Walz S, Fitzgerald KN, et al. MYC regulates the antitumor immune response through CD47 and PD-L1. Science (80-) [Internet]. 2016 Apr 8;352(6282):227–31. Available from: https://www.science.org/doi/10.1126/science.aaf7984

133. Atsaves V, Tsesmetzis N, Chioureas D, Kis L, Leventaki V, Drakos E, et al. PD-L1 is commonly expressed and transcriptionally regulated by STAT3 and MYC in ALK-negative anaplastic large-cell lymphoma. Leukemia [Internet]. 2017 Jul 27;31(7):1633–7. Available from: http://dx.doi.org/10.1038/leu.2017.103

134. Casacuberta-Serra S, González-Larreategui Í, Capitán-Leo D, Soucek L. MYC and KRAS cooperation: from historical challenges to

therapeutic opportunities in cancer. Signal Transduct Target Ther [Internet]. 2024 Aug 21;9(1):205. Available from: http://dx.doi.org/10.1038/s41392-024-01907-z

135. Greten FR, Grivennikov SI. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. Immunity [Internet]. 2019;51(1):27–41. Available from: https://doi.org/10.1016/j.immuni.2019.06.025

136. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell [Internet]. 2011;144(5):646–74. Available from: http://dx.doi.org/10.1016/j.cell.2011.02.013

137. Martínez-Martín S, Soucek L. MYC inhibitors in multiple myeloma. Cancer Drug Resist. 2021;4(4):842–65.

138. Donati G, Amati B. MYC and therapy resistance in cancer: risks and opportunities. Mol Oncol. 2022;16(21):3828–54.

139. Cyberski TF, Singh A, Korzinkin M, Mishra V, Pun F, Shen L, et al. Acquired resistance to immunotherapy and chemoradiation in MYC amplified head and neck cancer. npj Precis Oncol. 2024;8(1):1–10.

140. Elbadawy M, Usui T, Yamawaki H, Sasaki K. Emerging roles of C-myc in cancer stem cell-related signaling and resistance to cancer chemotherapy: A potential therapeutic target against colorectal cancer. Int J Mol Sci. 2019;20(9).

141. D'avola A, Kluckova K, Finch AJ, Riches JC. Spotlight on New Therapeutic Opportunities for MYC-Driven Cancers. Onco Targets Ther. 2023;16:371–83.

142. Poortinga G, Quinn LM, Hannan RD. Targeting RNA polymerase i to treat MYC-driven cancer. Oncogene [Internet]. 2015;34(4):403–12. Available from: http://dx.doi.org/10.1038/onc.2014.13

143. Papadimitropoulou A, Makri M, Zoidis G. MYC the oncogene from hell: Novel opportunities for cancer therapy. Eur J Med Chem [Internet]. 2024;267(October 2023):116194. Available from: https://doi.org/10.1016/j.ejmech.2024.116194

144. Fletcher S, Prochownik E V. Small-molecule inhibitors of the Myc oncoprotein. Biochim Biophys Acta - Gene Regul Mech [Internet]. 2015;1849(5):525–43. Available from: http://dx.doi.org/10.1016/j.bbagrm.2014.03.005

145. Weber LI, Hartl M. Strategies to target the cancer driver MYC in tumor cells. Front Oncol. 2023;13(March):1–22.

146. Whitfield JR, Beaulieu ME, Soucek L. Strategies to inhibit Myc and their clinical applicability. Front Cell Dev Biol. 2017;5(FEB):1–13.

147. Llombart V, Mansour MR. Therapeutic targeting of “undruggable” MYC. eBioMedicine [Internet]. 2022;75:103756. Available from: https://doi.org/10.1016/j.ebiom.2021.103756

148. Berg T, Cohen SB, Desharnais J, Sonderegger C, Maslyar DJ, Goldberg J, et al. Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. Proc Natl Acad Sci U S A. 2002;99(6):3830–5.

149. Shi J, Stover JS, Whitby LR, Vogt PK, Boger DL. Small molecule inhibitors of Myc/Max dimerization and Myc-induced cell transformation. Bioorganic Med Chem Lett [Internet]. 2009;19(21):6038–41. Available from: http://dx.doi.org/10.1016/j.bmcl.2009.09.044

150. Xu Y, Shi J, Yamamoto N, Moss JA, Vogt PK, Janda KD. A credit-card library approach for disrupting protein-protein interactions. Bioorganic Med Chem. 2006;14(8):2660–73.

151. Yap JL, Wang H, Hu A, Chauhan J, Jung K-Y, Gharavi RB, et al. Pharmacophore identification of c-Myc inhibitor 10074-G5. Bioorg Med Chem Lett [Internet]. 2013;23:370–4. Available from: https://doi.org/10.1016/j.bmcl.2012.10.013

152. Han H, Jain AD, Truica MI, Izquierdo-Ferrer J, Anker JF, Lysy B, et al. Small-Molecule MYC Inhibitors Suppress Tumor Growth and Enhance Immunotherapy. Cancer Cell [Internet]. 2019;36(5):483-497.e15. Available from: https://doi.org/10.1016/j.ccell.2019.10.001

153. Webb MS, Tortora N, Cremese M, Kozlowska H, Blaquiere M, Devine D V., et al. Toxicity and toxicokinetics of a phosphorothioate oligonucleotide against the c-myc oncogene in cynomolgus monkeys. Antisense Nucleic Acid Drug Dev. 2001;11(3):155–63.

154. Sekhon HS, London CA, Sekhon M, Iversen PL, Devi GR. c-MYC antisense phosphosphorodiamidate morpholino oligomer inhibits lung metastasis in a murine tumor model. Lung Cancer. 2008;60(3):347–54.

155. Tolcher AW, Papadopoulos KP, Patnaik A, Rasco DW, Martinez D,

Wood DL, et al. Safety and activity of DCR-MYC, a first-in-class Dicer-substrate small interfering RNA (DsiRNA) targeting MYC, in a phase I study in patients with advanced solid tumors. J Clin Oncol [Internet]. 2015 May 20;33(15_suppl):11006. Available from: https://doi.org/10.1200/jco.2015.33.15_suppl.11006

156. Zhang Y, Peng L, Mumper RJ, Huang L. Combinational delivery of c-myc siRNA and nucleoside analogs in a single, synthetic nanocarrier for targeted cancer therapy. Biomaterials [Internet]. 2013;34(33):8459–68. Available from: http://dx.doi.org/10.1016/j.biomaterials.2013.07.050

157. Bidwell GL, Perkins E, Raucher D. A thermally targeted c-Myc inhibitory polypeptide inhibits breast tumor growth. Cancer Lett [Internet]. 2012;319(2):136–43. Available from: http://dx.doi.org/10.1016/j.canlet.2011.12.042

158. Verzele D, Madder A. Synthetic progress in cMyc-Max oncoprotein miniaturization: Semi-online monitoring gives solid-phase access to hydrophobic b(-HLH-)ZIP peptidosteroid tweezers. European J Org Chem. 2013;(4):673–87.

159. Montagne M, Beaudoin N, Fortin D, Lavoie CL, Klinck R, Lavigne P. The max b-HLH-LZ can transduce into cells and inhibit c-Myc transcriptional activities. PLoS One. 2012;7(2):2–10.

160. Beaulieu ME, Jauset T, Massó-Vallés D, Martínez-Martín S, Rahl P, Maltais L, et al. Intrinsic cell-penetrating activity propels omomyc from proof of concept to viable anti-myc therapy. Sci Transl Med. 2019;11(484):1–14.

161. Li Z, Huang Y, Hung TI, Sun J, Aispuro D, Chen B, et al. MYC-Targeting Inhibitors Generated from a Stereodiversified Bicyclic Peptide Library. J Am Chem Soc. 2024;

162. Mukhopadhyay A, Hanold LE, Thayele Purayil H, Gisemba SA, Senadheera SN, Aldrich J V. Macrocyclic peptides decrease c-Myc protein levels and reduce prostate cancer cell growth. Cancer Biol Ther. 2017;18(8):571–83.

163. Brockmann M, Poon E, Berry T, Carstensen A, Deubzer HE, Rycak L, et al. Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma. Cancer Cell [Internet]. 2013;24(1):75–89. Available from: http://dx.doi.org/10.1016/j.ccr.2013.05.005

164. Follis AV, Hammoudeh DI, Wang H, Prochownik E V., Metallo SJ. Structural Rationale for the Coupled Binding and Unfolding of the c-Myc Oncoprotein by Small Molecules. Chem Biol [Internet]. 2008;15(11):1149–55. Available from: http://dx.doi.org/10.1016/j.chembiol.2008.09.011

165. Jung KY, Wang H, Teriete P, Yap JL, Chen L, Lanning ME, et al. Perturbation of the c-Myc-Max Protein-Protein Interaction via Synthetic α-Helix Mimetics. J Med Chem. 2015;58(7):3002–24.

166. Holmes AG, Parker JB, Sagar V, Truica MI, Soni PN, Han H, et al. A MYC inhibitor selectively alters the MYC and MAX cistromes and modulates the epigenomic landscape to regulate target gene expression. Sci Adv. 2022;8(17):1–17.

167. Tang M, O'Grady S, Crown J, Duffy MJ. MYC as a therapeutic target for the treatment of triple-negative breast cancer: preclinical investigations with the novel MYC inhibitor, MYCi975. Breast Cancer Res Treat [Internet]. 2022;195(2):105–15. Available from: https://doi.org/10.1007/s10549-022-06673-6

168. Structural N. MYC phase separation selectively modulates the transcriptome. 2024;

169. Hatton KS, Mahon K, Chin L, Chiu F, Lee H, Peng D, et al. Expression and Activity of L-Myc in Normal Mouse Development. Mol Cell Biol. 1996;16(4):1794–804.

170. Soucek L, Whitfield J, Martins CP, Finch AJ, Murphy DJ, Sodir NM, et al. Modelling Myc inhibition as a cancer therapy. Nature. 2008;455(7213):679–83.

171. Soucek L, Helmer-Citterich M, Sacco A, Jucker R, Cesareni G, Nasi S. Design and properties of a Myc derivative that efficiently homodimerizes. Oncogene. 1998;17(19):2463–72.

172. Jung LA, Gebhardt A, Koelmel W, Ade CP, Walz S, Kuper J, et al. OmoMYC blunts promoter invasion by oncogenic MYC to inhibit gene expression characteristic of MYC-dependent tumors. Oncogene. 2016;36(14):1911–24.

173. Savino M, Annibali D, Carucci N, Favuzzi E, Cole MD, Evan GI, et al. The action mechanism of the Myc inhibitor termed Omomyc may give clues on how to target Myc for cancer therapy. PLoS One. 2011;6(7).

174. Annibali D, Whitfield JR, Favuzzi E, Jauset T, Serrano E, Cuartas I, et al. Myc inhibition is effective against glioma and reveals a role for Myc in proficient mitosis. Nat Commun. 2014;5:1–11.

175. Fiorentino FP, Tokgün E, Solé-Sánchez S, Giampaolo S, Tokgün O, Jauset T, et al. Growth suppression by MYC inhibition in small cell lung cancer cells with TP53 and RB1 inactivation. Oncotarget. 2016;7(21):31014–28.

176. Zacarías-Fluck MF, Massó-Vallés D, Giuntini F, González-Larreategui Í, Kaur J, Casacuberta-Serra S, et al. Reducing MYC's transcriptional footprint unveils a good prognostic gene signature in melanoma. Genes Dev. 2023;37(7–8):303–20.

177. Soucek L, Whitfield JR, Sodir NM, Massó-Vallés D, Serrano E, Karnezis AN, et al. Inhibition of Myc family proteins eradicates KRas-driven lung cancer in mice. Genes Dev. 2013;27(5):504–13.

178. von Eyss B, Jaenicke LA, Kortlever RM, Royla N, Wiese KE, Letschert S, et al. A MYC-Driven Change in Mitochondrial Dynamics Limits YAP/TAZ Function in Mammary Epithelial Cells and Breast Cancer. Cancer Cell. 2015;28(6):743–57.

179. Galardi S, Savino M, Scagnoli F, Pellegatta S, Pisati F, Zambelli F, et al. Resetting cancer stem cell regulatory nodes upon MYC inhibition . EMBO Rep. 2016;17(12):1872–89.

180. Eyss V, Eilers M. Addicted to Myc — but why? Genes Dev. 2011;(25):895–7.

181. The United States Food and Drug Administration. FDA Approval Brings First Gene Therapy to the United States. Online Publ. 2017;9(3):17–8.

182. Noguchi H, Bonner-weir S, Wei F, Matsushita M, Matsumoto S, Neurod B. BETA2/NeuroD Protein Can Be Transduced Into Cells Due to an Arginine- and Lysine-Rich Sequence. Diabetes. 2005;54(October).

183. Demma MJ, Mapelli C, Sun A, Bodea S, Ruprecht B, Javaid S, et al. Omomyc Reveals New Mechanisms To Inhibit the MYC Oncogene. Mol Cell Biol. 2019;39(22):1–27.

184. Beaulieu ME, Martínez-Martín S, Kaur J, Castillo Cano V, Massó-Vallés D, Foradada Felip L, et al. Pharmacokinetic Analysis of Omomyc Shows Lasting Structural Integrity and Long Terminal Half-

Life in Tumor Tissue. Cancers (Basel). 2023;15(3).

185. Garralda E, Beaulieu ME, Moreno V, Casacuberta-Serra S, Martínez-Martín S, Foradada L, et al. MYC targeting by OMO-103 in solid tumors: a phase 1 trial. Nat Med. 2024;

186. Xu J, Chen G, De Jong AT, Shahravan SH, Shin JA. Max-E47, a designed minimalist protein that targets the E-box DNA site in vivo and in vitro. J Am Chem Soc. 2009;131(22):7839–48.

187. Lustig LC, Dingar D, Tu WB, Lourenco C, Kalkat M, Inamoto I, et al. Inhibiting MYC binding to the E-box DNA motif by ME47 decreases tumour xenograft growth. Oncogene. 2017;36(49):6830–7.

188. Popa SC, Inamoto I, Thuronyi BW, Shin JA. Phage-assisted continuous evolution (PACE): A guide focused on evolving protein−DNA interactions. ACS Omega. 2020;5(42):26957–66.

189. Inamoto I, Sheoran I, Popa SC, Hussain M, Shin JA. Combining Rational Design and Continuous Evolution on Minimalist Proteins That Target the E-box DNA Site. ACS Chem Biol. 2021;16(1):35–44.

190. Wang E, Sorolla A, Cunningham PT, Bogdawa HM, Beck S, Golden E, et al. Tumor penetrating peptides inhibiting MYC as a potent targeted therapeutic strategy for triple-negative breast cancers. Oncogene [Internet]. 2019;38(1):140–50. Available from: http://dx.doi.org/10.1038/s41388-018-0421-y

191. PYC Therapeutics. A new generation of RNA Therapeutics [Internet]. 2023. Available from: https://pyctx.com/our-science/

192. Demma MJ, Hohn MJ, Sun A, Mapelli C, Hall B, Walji A, et al. Inhibition of Myc transcriptional activity by a mini-protein based upon Mxd1. FEBS Lett. 2020;594(10):1467–76.

193. Brown ZZ, Mapelli C, Farasat I, Shoultz A V., Johnson SA, Orvieto F, et al. Multiple Synthetic Routes to the Mini-Protein Omomyc and Coiled-Coil Domain Truncations. J Org Chem. 2020;85(3):1466–75.

194. Mijalis AJ, Thomas DA, Simon MD, Adamo A, Beaumont R, Jensen KF, et al. A fully automated flow-based approach for accelerated peptide synthesis. Nat Chem Biol [Internet]. 2017;13(5):464–6. Available from: http://dx.doi.org/10.1038/nchembio.2318

195. Jbara M, Pomplun S, Schissel CK, Hawken SW, Boija A, Klein I, et al. Engineering Bioactive Dimeric Transcription Factor Analogs via

Palladium Rebound Reagents. J Am Chem Soc. 2021;143(30):11788–98.

196. Pomplun S, Jbara M, Schissel CK, Wilson Hawken S, Boija A, Li C, et al. Parallel Automated Flow Synthesis of Covalent Protein Complexes That Can Inhibit MYC-Driven Transcription. ACS Cent Sci. 2021;7(8):1408–18.

197. Speltz TE, Qiao Z, Swenson CS, Shangguan X, Coukos JS, Lee CW, et al. Targeting MYC with modular synthetic transcriptional repressors derived from bHLH DNA-binding domains. Nat Biotechnol. 2022;

198. Rasmussen JH. Synthetic peptide API manufacturing: A mini review of current perspectives for peptide manufacturing. Bioorganic Med Chem [Internet]. 2018;26(10):2914–8. Available from: https://doi.org/10.1016/j.bmc.2018.01.018

199. Ellenbroek BD, Kahler JP, Evers SR, Pomplun SJ. Synthetic Peptides: Promising Modalities for the Targeting of Disease-Related Nucleic Acids. Angew Chemie - Int Ed. 2024;63(20).

200. Naud JF, McDuff FO, Sauvé S, Montagne M, Webb BA, Smith SP, et al. Structural and thermodynamical characterization of the complete p21 gene product of Max. Biochemistry. 2005;44(38):12746–58.

201. Mc Duff F-O. Expression, Purification et Caractérisation Structurale et Thermodynamique des Complexes Formés de la Reconnaissance Moléculaire des Domaines b-HLH-LZ des Protéines Humaines c-MYC et MAX. Université de Sherbrooke; 2008.

202. Schütz S, Bergsdorf C, Goretzki B, Lingel A, Renatus M, Gossert AD, et al. The disordered MAX N-terminus modulates DNA binding of the transcription factor MYC:MAX. J Mol Biol. 2022;(September):167833.

203. Deprey K, Becker L, Kritzer J, Plückthun A. Trapped! A Critical Evaluation of Methods for Measuring Total Cellular Uptake versus Cytosolic Localization. Bioconjug Chem [Internet]. 2019 Apr 17;30(4):1006–27. Available from: https://pubs.acs.org/doi/10.1021/acs.bioconjchem.9b00112

204. Liu J, Afshar S. In vitro assays: Friends or foes of cell-penetrating peptides. Int J Mol Sci. 2020;21(13):1–20.

205. Uhlén M, Nilsson B, Guss B, Lindberg M, Gatenbeck S, Philipson L. Gene fusion vectors based on the gene for staphylococcal protein A.

Gene. 1983;23(3):369–78.

206. Mahmoodi S, Pourhassan-Moghaddam M, Wood DW, Majdi H, Zarghami N. Current affinity approaches for purification of recombinant proteins. Cogent Biol. 2019;5(1):1665406.

207. Pina AS, Lowe CR, Roque ACA. Challenges and opportunities in the purification of recombinant tagged proteins. Biotechnol Adv. 2014;32(2):366–81.

208. Hochuli E, Bannwarth W, Dobeli H, Gentzi R, Stuber D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology. 1988;6(11):1321–5.

209. Schmidt TGM, Koepke J, Frank R, Skerra A. Molecular interaction between the strep-lag affinity peptide and its cognate target, streptavidin. J Mol Biol. 1996;255(5):753–66.

210. Korndörfer IP, Skerra A. Improved affinity of engineered streptavidin for the Strep -tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site . Protein Sci. 2002;11(4):883–93.

211. Yeliseev A, Zoubak L, Schmidt TGM. Application of Strep-Tactin XT for affinity purification of Twin-Strep-tagged CB 2 , a G protein-coupled cannabinoid receptor. Protein Expr Purif [Internet]. 2017 Mar;131(1):109–18. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1046592816303321

212. Magdeldin S, Moser A. Affinity Chromatography: Principles and Applications. In: Magdeldin S, editor. Affinity Chromatography [Internet]. Rijeka: IntechOpen; 2012. Available from: https://doi.org/10.5772/39087

213. Borrelli A, Tornesello AL, Tornesello ML, Buonaguro FM. Cell penetrating peptides as molecular carriers for anti-cancer agents. Molecules. 2018;23(2).

214. Zou L-L, Ma J-L, Wang T, Yang T-B, Liu C-B. Cell-Penetrating Peptide-Mediated Therapeutic Molecule Delivery into the Central Nervous System. Curr Neuropharmacol. 2013;11(2):197–208.

215. Langel Ü. Cell-penetrating peptides and transportan. Vol. 13, Pharmaceutics. 2021. 1–31 p.

216. Derakhshankhah H, Jafari S. Cell penetrating peptides: A concise review with emphasis on biomedical applications. Biomed

Pharmacother [Internet]. 2018;108(June):1090–6. Available from: https://doi.org/10.1016/j.biopha.2018.09.097

217. Guidotti G, Brambilla L, Rossi D. Cell-Penetrating Peptides: From Basic Research to Clinics. Trends Pharmacol Sci [Internet]. 2017;38(4):406–24. Available from: http://dx.doi.org/10.1016/j.tips.2017.01.003

218. Ramaker K, Henkel M, Krause T, Röckendorf N, Frey A. Cell penetrating peptides: A comparative transport analysis for 474 sequence motifs. Drug Deliv [Internet]. 2018;25(1):928–37. Available from: https://doi.org/10.1080/10717544.2018.1458921

219. Futaki S. Membrane-permeable arginine-rich peptides and the translocation mechanisms. Adv Drug Deliv Rev. 2005;57(4 SPEC.ISS.):547–58.

220. Aumelas A, Mangoni M, Roumestand C, Chiche L, Despaux E, Grassy G, et al. Synthesis and solution structure of the antimicrobial peptide protegrin-1. Eur J Biochem. 1996;237(3):575–83.

221. Temsamani J, Rees AR, Scherrmann JM. Vector-mediated drug delivery to the brain. Expert Opin Biol Ther. 2001;1(5):773–82.

222. Drin G, Cottin S, Blanc E, Rees AR, Temsamani J. Studies on the internalization mechanism of cationic cell-penetrating peptides. J Biol Chem [Internet]. 2003;278(33):31192–201. Available from: http://dx.doi.org/10.1074/jbc.M303938200

223. Rousselle C, Clair P, Lefauconnier J-M, Kaczorek M, Scherrmann J-M, Temsamani J. New Advances in the Transport of Doxorubicin through the Blood-Brain Barrier by a Peptide Vector-Mediated Strategy. Mol Pharmacol [Internet]. 2000 Apr 1;57(4):679–86. Available from: http://molpharm.aspetjournals.org/lookup/doi/10.1124/mol.57.4.679

224. Rousselle C, Smirnova M, Clair P, Lefauconnier JM, Chavanieu A, Calas B, et al. Enhanced delivery of doxorubicin into the brain via a peptide-vector-mediated strategy: Saturation kinetics and specificity. J Pharmacol Exp Ther. 2001;296(1):124–31.

225. Rousselle C, Clair P, Temsamani J, Scherrmann JM. Improved brain delivery of benzylpenicillin with a peptide-vector-mediated strategy. J Drug Target. 2002;10(4):309–15.

226. Liu H, Zhang W, Ma L, Fan L, Gao F, Ni J, et al. The improved blood-brain barrier permeability of endomorphin-1 using the cell-penetrating peptide synB3 with three different linkages. Int J Pharm [Internet]. 2014;476(1):1–8. Available from: http://dx.doi.org/10.1016/j.ijpharm.2014.08.045

227. Zuo L, Li W, Shi J, Su Y, Shuai H, Yu X. SynB3 Conjugated QBP1 Passes Blood-Brain Barrier Models and Inhibits PolyQ Protein Aggregation. Protein Pept Lett [Internet]. 2022 Jan;29(1):110–20. Available from: https://www.eurekaselect.com/199229/article

228. Derossi D, Joliot AH, Chassaing G, Prochiantz A. The third helix of the Antennapedia homeodomain translocates through biological membranes. J Biol Chem. 1994;269(14):10444–50.

229. Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. Cell. 1988;55(6):1189–93.

230. Davidson TJ, Harel S, Arboleda VA, Prunell GF, Shelanski ML, Greene LA, et al. Highly efficient small interfering RNA delivery to primary mammalian neurons induces microRNA-like effects before mRNA degradation. J Neurosci. 2004;24(45):10040–6.

231. Han K, Jeon MJ, Kim KA, Park J, Choi SY. Efficient intracellular delivery of GFP by homeodomains of Drosophila Fushi-tarazu and Engrailed proteins. Mol Cells [Internet]. 2000;10(6):728–32. Available from: https://doi.org/10.1007/S1016-8478(23)17542-X

232. Sayers EJ, Cleal K, Eissa NG, Watson P, Jones AT. Distal phenylalanine modification for enhancing cellular delivery of fluorophores, proteins and quantum dots by cell penetrating peptides. J Control Release [Internet]. 2014;195:55–62. Available from: http://dx.doi.org/10.1016/j.jconrel.2014.07.055

233. Lee J, Jung E, Park J, Park D. Transdermal delivery of interferon-γ (IFN-γ) mediated by penetratin, a cell-permeable peptide. Biotechnol Appl Biochem. 2005;42(2):169.

234. Mitchell DJ, Steinman L, Kim DT, Fathman CG, Rothbard JB. Polyarginine enters cells more efficiently than other polycationic homopolymers. J Pept Res. 2000;56(5):318–25.

235. Fuchs SM, Raines RT. Polyarginine as a multifunctional fusion tag. Protein Sci. 2005;14(6):1538–44.

236. Hu H, Wang J, Wang H, Tan T, Li J, Wang Z, et al. Cell-penetrating

peptide-based nanovehicles potentiate lymph metastasis targeting and deep penetration for anti-metastasis therapy. Theranostics. 2018;8(13):3597–610.

237. Morishita M, Kamei N, Ehara J, Isowa K, Takayama K. A novel approach using functional peptides for efficient intestinal absorption of insulin. J Control Release. 2007;118(2):177–84.

238. Bendifallah N, Rasmussen FW, Zachar V, Ebbesen P, Nielsen PE, Koppelhus U. Evaluation of cell-penetrating peptides (CPPs) as vehicles for intracellular delivery of antisense peptide nucleic acid (PNA). Bioconjug Chem. 2006;17(3):750–8.

239. Liu S, Verma A, Kettenberger H, Richter WF, Shah DK. Effect of variable domain charge on in vitro and in vivo disposition of monoclonal antibodies. MAbs [Internet]. 2021;13(1). Available from: https://doi.org/10.1080/19420862.2021.1993769

240. Li B, Tesar D, Boswell CA, Cahaya HS, Wong A, Zhang J, et al. Framework selection can influence pharmacokinetics of a humanized therapeutic antibody through differences in molecule charge. MAbs. 2014;6(5):1255–64.

241. Boswell CA, Tesar DB, Mukhyala K, Theil F, Fielder PJ, Khawli LA. Effects of Charge on Antibody Tissue Distribution and Pharmacokinetics. 2010;2153–63.

242. Chung S, Tian J, Tan Z, Chen J, Lee J, Borys M, et al. Industrial bioprocessing perspectives on managing therapeutic protein charge variant profiles. 2018;(December 2017):1646–65.

243. Kraft TE, Richter WF, Emrich T, Knaupp A, Schuster M, Wolfert A, et al. Heparin chromatography as an in vitro predictor for antibody clearance rate through pinocytosis. MAbs [Internet]. 2020 Jan 1;12(1). Available from: https://doi.org/10.1080/19420862.2019.1683432

244. Rosal Fontana M, Molano i Flores M, Espejo Ruiz C. REGLAMENTO INTERNO CEEA [Internet]. 2022. Available from: https://vhir.vallhebron.com/sites/default/files/2022-06/en_gb_reglament-intern-ceea.pdf

245. Soucek L, Beaulieu M-E. METHODS AND COMPOSITIONS FOR THE TREATMENT OF CANCER Pub. No.: US 2016/0122415 A1. Vol. 1. Spain; 2013. p. 1–5.

246. Kerkis A, Kerkis I, Rádis-Baptista G, Oliveira EB, Vianna-Morgante AM, Pereira L V., et al. Crotamine is a novel cell-penetrating protein from the venom of rattlesnake Crotalus durissus terrificus. FASEB J. 2004;18(12):1407–9.

247. Kerkis I, Hayashi MAF, Prieto Da Silva ARB, Pereira A, De Sá Júnior PL, Zaharenko AJ, et al. State of the art in the studies on crotamine, a cell penetrating peptide from South American rattlesnake. Biomed Res Int. 2014;2014.

248. Rádis-Baptista G, De La Torre BG, Andreu D. A novel cell-penetrating peptide sequence derived by structural minimization of a snake toxin exhibits preferential nucleolar localization. J Med Chem. 2008;51(22):7041–4.

249. Radis-Baptista G, Kerkis I. Crotamine, a Small Basic Polypeptide Myotoxin from Rattlesnake Venom with Cell-Penetrating Properties. Curr Pharm Des. 2012;17(38):4351–61.

250. Gan BK, Yong CY, Ho KL, Omar AR, Alitheen NB, Tan WS. Targeted Delivery of Cell Penetrating Peptide Virus-like Nanoparticles to Skin Cancer Cells. Sci Rep [Internet]. 2018;8(1):1–13. Available from: http://dx.doi.org/10.1038/s41598-018-26749-y

251. Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. Cell. 1984;39(3 PART 2):499–509.

252. Ragin AD, Morgan RA, Chmielewski J. Cellular import mediated by nuclear localization signal peptide sequences. Chem Biol. 2002;9(8):943–8.

253. Ballin JD, Prevas JP, Ross CR, Toth EA, Wilson GM, Record MT. Contributions of the Histidine Side Chain and the N-Terminal α-Amino Group to the Binding Thermodynamics of Oligopeptides to Nucleic Acids as a Function of pH. Biochemistry [Internet]. 2010 Mar 9;49(9):2018–30. Available from: https://pubs.acs.org/doi/10.1021/bi902027z

254. Pogostin BH, Malmendal A, Londergan CH, Åkerfeldt KS. PKA determination of a histidine residue in a short peptide using Raman spectroscopy. Molecules. 2019;24(3).

255. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B [Internet].

2022;12(7):3049–62. Available from: https://doi.org/10.1016/j.apsb.2022.02.002

256. Sahni A, Ritchey JL, Qian Z, Pei D. Cell-Penetrating Peptides Translocate across the Plasma Membrane by Inducing Vesicle Budding and Collapse. J Am Chem Soc [Internet]. 2024 Sep 2;15(9):2485–92. Available from: https://pubs.acs.org/doi/10.1021/acschembio.0c00478

257. Lee HJ, Huang YW, Chiou SH, Aronstam RS. Polyhistidine facilitates direct membrane translocation of cell-penetrating peptides into cells. Sci Rep. 2019;9(1):1–11.

258. Zhao Y, Jiang H, Yu J, Wang L, Du J. Engineered Histidine-Rich Peptides Enhance Endosomal Escape for Antibody-Targeted Intracellular Delivery of Functional Proteins. Angew Chemie - Int Ed. 2023;62(38).

259. López-Laguna H, Cubarsi R, Unzueta U, Mangues R, Vázquez E, Villaverde A. Endosomal escape of protein nanoparticles engineered through humanized histidine-rich peptides. Sci China Mater. 2020;63(4):644–53.

260. Salichs E, Ledda A, Mularoni L, Albà MM, de la Luna S. Genome-Wide Analysis of Histidine Repeats Reveals Their Role in the Localization of Human Proteins to the Nuclear Speckles Compartment. Katsanis N, editor. PLoS Genet [Internet]. 2009 Mar 6;5(3):e1000397. Available from: https://dx.plos.org/10.1371/journal.pgen.1000397

261. Deng C, Xiong J, Gu X, Chen X, Wu S, Wang Z, et al. Novel recombinant immunotoxin of EGFR specific nanobody fused with cucurmosin, construction and antitumor efficiency in vitro. Oncotarget. 2017;8(24):38568–80.

262. Hu J, Banerjee A, Goss DJ. Assembly of b / HLH / z Proteins c-Myc , Max , and Mad1 with Cognate DNA : Importance of Protein - Protein and Protein - DNA Interactions. Biochemistry. 2005;44:11855–63.

263. Simonsen JB, Kromann EB. Pitfalls and opportunities in quantitative fluorescence-based nanomedicine studies – A commentary. J Control Release [Internet]. 2021;335(May):660–7. Available from: https://doi.org/10.1016/j.jconrel.2021.05.041

264. Yin L, Wang W, Wang S, Zhang F, Zhang S. Biosensors and

Bioelectronics How does fl uorescent labeling affect the binding kinetics of proteins with intact cells ? Biosens Bioelectron [Internet]. 2015;66:412–6. Available from: http://dx.doi.org/10.1016/j.bios.2014.11.036

265. Jadavi S, Dante S, Civiero L, Sandre M, Bubacco L, Tosatto L, et al. Fluorescence labeling methods influence the aggregation process of α-syn in vitro differently. Nanoscale. 2023;15:8270–7.

Acknowledgement

First and foremost, I would like to express my deepest gratitude to Laura and Marie-Eve. Thank you for your most inspiring hunger for success, for your teachings, but also for your permanent trust and benevolence as well as for allowing me to be in the best possible conditions throughout this journey. Conducting this work under your guidance felt like an absolute privilege and will remain for me a great source of pride.

None of this work would have been possible without the constant support, moral and technical, of the amazing team members (current and past) of Peptomyc and of VHIO's Model of Cancer Therapies group. A thousand thanks to every single one of you, who made me feel like part of a new family since day 1. Working alongside you guys was a blast and undoubtedly the best remedy against bad days.

I am particularly grateful to Virginia, thank you for all your teachings throughout these years which gave me all the tools I needed to complete this work. Your kindness and humility are a lesson that I had the privilege to attend.

I am also most grateful to my fellow PhD students, Fabio and Íñigo, for all the support and companionship that have made this journey not only more manageable but truly enjoyable. Making friends like you guys would be sufficient for me to say that this experience was worth living.

I'd like to give Silvia and Sandra my special thanks for always having the right advice at the right time. Thank you for your most valuable guidance, and support as well as for your friendship. One can only wish to have the chance to work alongside people as pleasant and competent as you are.

J'aimerais aussi remercier mes parents pour leur soutien permanent et inconditionnel, ainsi que pour m'avoir enseigné la valeur de la persévérance et de l'exigence. Merci de m'avoir donné les clés pour réussir et d'être encore aujourd'hui des exemples pour moi. Merci également à mes deux sœurs Aurélie et Julia pour me soutenir et me bousculer si besoin, votre exemple m'a toujours poussé à donner le meilleur de moi-même.

Finalment, gràcies a tu, Cris, pel teu amor i suport, així com per l'alegria, la felicitat i l'energia que em portes cada dia. Compartir la teva vida dona sentit a la meva.