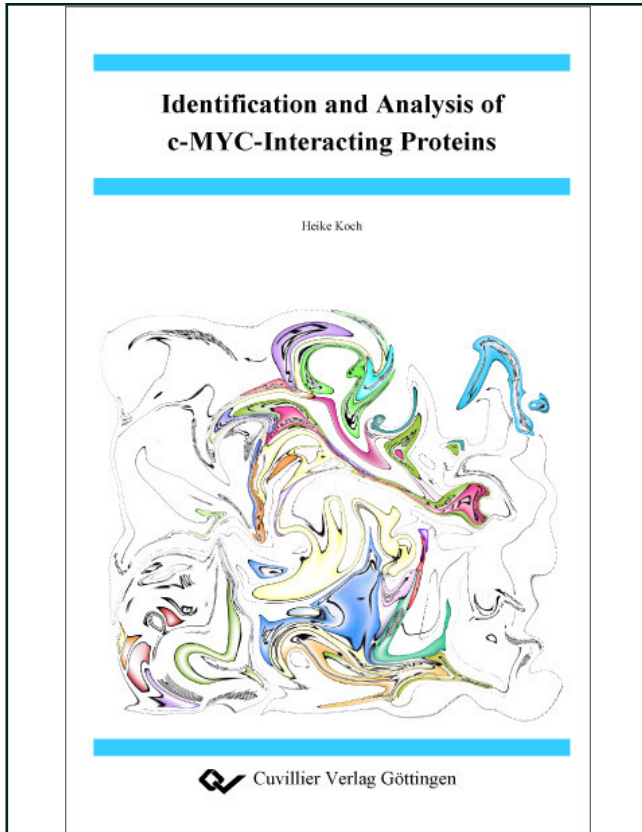




Heike Koch (Autor)
**Identification and Analysis of c-MYC-Interacting
Proteins**



<https://cuvillier.de/de/shop/publications/1826>

Copyright:
Cuvillier Verlag, Inhaberin Annette Jentsch-Cuvillier, Nonnenstieg 8, 37075 Göttingen,
Germany
Telefon: +49 (0)551 54724-0, E-Mail: info@cuvillier.de, Website: <https://cuvillier.de>

1 INTRODUCTION

1.1 Homeostasis

Life of multicellular organisms depends on a balanced equilibrium of cellular growth, division, differentiation, and death. A large variety of mechanisms are in function to orchestrate these processes and assure permanent homeostasis. On the cellular level, progression through the cell cycle has to be tightly regulated in order to coordinate the interplay of growth and genome replication steps, which is monitored by different checkpoint pathways. In normal ontogenesis, a large number of proto-oncogenes fulfills essential functions for normal cell proliferation in response to extracellular signals like growth factors and cell-to-cell contacts. Mutation of any of these proto-oncogenes leading to proliferative defects can cause developmental abnormalities, a potentially preneoplastic state, and, when accumulated, result in cancer.

The *c-MYC* (*v-MYC* avian myelocytomatosis viral oncogene homologue) gene represents a paradigm for the oncogenic switch in many human and animal cancers. In normal cells, proto-oncogenic *c-MYC* acts as one of the major regulators of proliferation indispensable for a wide spectrum of cell cycle processes. Precise control of the activity of *c-MYC* is a prerequisite to maintain cellular homeostasis as activation of *c-MYC* leads to unscheduled cell proliferation, extends the self-renewal capacity, and blocks the terminal differentiation of various cell types.

1.2 The family of *MYC* genes

The *MYC* family comprises five structurally related genes termed *c-MYC*, *B-MYC*, *L-MYC*, *N-MYC*, and *S-MYC*. *S-MYC* and *B-MYC* have only been described in rodents and are the least characterized members of the *MYC* family. *S-MYC* is an intron-less gene and its protein product is potentially tumor suppressive^{1,2}. *B-MYC*, which is almost ubiquitously expressed³, acts as a negative regulator of transcriptional activation and neoplastic transformation of *c-MYC*⁴. The main member of the *MYC* gene family is *c-MYC*, which has been originally identified as the cellular progenitor of the *v-MYC* oncogene of avian myelocytomatosis retrovirus *MC29* causing leukemias and carcinomas in chicken⁵⁻⁸. There is no evidence for *c-MYC* homologues in *Saccharomyces cerevisiae*⁹ and *Caenorhabditis elegans*¹⁰, but *c-MYC* is conserved throughout chordate evolution including man¹¹⁻¹⁴, mice¹⁵, birds^{6,7}, amphibians¹⁶, and fish¹⁷. In addition, *c-MYC* homologues were identified in sea stars¹⁸ and flies (*dMYC*^{19,20}). These data suggest that the *c-MYC* gene has evolved very

early and is likely to fulfill functions essential for cellular and organismal physiology. The *c-MYC*, *L-MYC*, and *N-MYC* genes share extensive homology in their coding regions²¹ and consist of three exons with the major open reading frame residing in the second and third exon²². The corresponding proteins resemble each other in the ability to interact with MAX (MYC-associated factor X)²³⁻²⁶, to bind DNA when dimerized with MAX²⁷⁻³⁰, and to directly transactivate or repress gene expression³¹. They exhibit malignant co-transformation of rat embryo fibroblasts with mutant RAS³²⁻³⁴ and generate tumors when over-expressed in transgenic mice³⁵. Yet, the oncogenic potential of the three proteins markedly differs^{26,36} and they show distinct expression patterns with respect to cell type and developmental stage³⁷⁻³⁹. These differences in activity and expression together with evolutionary conservation of these three family members over large phylogenetic distances⁴⁰ indicate that the functional redundancy of c-MYC, N-MYC, and L-MYC is limited and that they serve distinct physiological roles. Whereas L-MYC-negative mice lack a phenotype⁴¹, homozygous null mutations of *c-MYC* or *N-MYC* are embryonically lethal at 9.5-10.5 days⁴² or 10.5-12.5 days^{39,43} after gestation, respectively. c-MYC-deficient embryos generally display retarded development with abnormalities of heart, pericardium, and neural tube, and a delay or failure in turning of the embryo⁴². Consequences of N-MYC-deficiency include developmental defects of the central and peripheral nervous system, limb buds, heart, mesonephros, lung, liver, and gut^{39,43,44} despite compensatory up-regulation of *c-MYC* in the neuroepithelium³⁹. Furthermore, *N-MYC* knockout embryos show extensive hepatic apoptosis and a marked reduction in the number of hematopoietic cells in the liver⁴⁵. These data confirm the inability of c-MYC and N-MYC to functionally substitute for each other in the midgestation embryo. Nevertheless, knock-in of the *N-MYC* gene into the *c-MYC* locus preserving most of the untranslated region of *c-MYC* allows development of almost normal mice⁴⁶. Different physiological roles of c-MYC and N-MYC therefore most likely relate primarily to differences in their transcriptional regulation rather than to distinct biochemical properties of the two proteins.

1.3 The structure of the c-MYC protein and c-MYC isoforms

The *c-MYC* mRNA is transcribed from four distinct promoters of different strength. The main promoters P1 and P2 contribute 10-25 % or 75-90 % of the cytoplasmic *c-MYC* mRNA, respectively, and are located within the first exon of the *c-MYC* locus⁴⁷. Two additional promoters P0 and P3 upstream of the first and the third exon play an inferior role in normal *c-MYC* transcription⁴⁷ (Figure 1).

INTRODUCTION

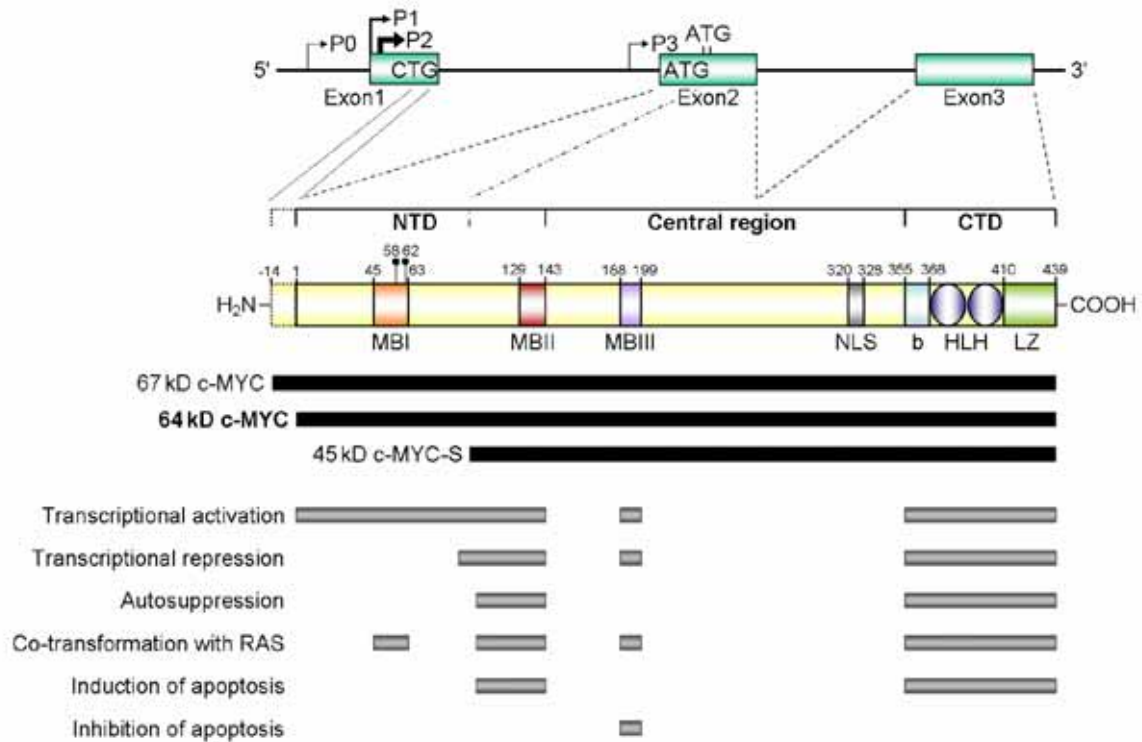


Figure 1. Schematic representation of the *c-MYC* gene, domain architecture of the *c-MYC* protein, and functional activity of regions of the *c-MYC* protein (adapted from ⁴⁷ and ⁵¹).

The *c-MYC* gene consists of three exons and is transcribed from four promoters indicated as P0, P1, P2, and P3. Four alternative translation initiation sites are shown as the codons of the initial amino acid.

The *c-MYC* protein comprises an N-terminal domain (NTD), a central region, and a C-terminal domain (CTD). The NTD harbors the two major phosphorylation sites Thr 58 and Ser 62 of *c-MYC* and the conserved MYC boxes (MBI and MBII), which are essential for the transcriptional activity of *c-MYC*. MYC box III (MBIII) and the NLS (nuclear localization signal) are located within the central region. The CTD consists of the basic (b) HLH (helix-loop-helix) region and a LZ (leucine zipper) motif for dimerization with MAX and DNA-binding of *c-MYC*/MAX-heterodimers. The specific functions of regions of the *c-MYC* protein are depicted.

The *c-MYC* protein represents a nuclear phosphoprotein with three naturally occurring isoforms, which are produced by translation initiation at different codons ^{48,49}. The major 64 kD isoform is initiated from a canonical AUG start codon within the second exon while the longer 67 kD isoform results from translation initiation upstream of the AUG at a CUG codon located in the first exon. The 67 kD isoform with 14 additional N-terminal amino acids is less abundant and differs from the 64 kD isoform in the binding specificity of promoter regions ⁵⁰. The third isoform *c-MYC-S* with a protein mass of 45 kD arises from alternative translation initiated at two closely spaced downstream AUG codons when the pre-initiation scanning complex bypasses the first AUG codon by a leaky ribosomal scanning mechanism ⁴⁹. In the following text, "*c-MYC*" refers to the major 64 kD isoform unless indicated otherwise.

The c-MYC protein is tripartite in organization (Figure 1). Notable structural features of the protein include a globular amino (N)-terminal region (AA 1-143), an unstructured partially acidic middle area (AA 144-354), and an α -helical carboxy (C)-terminal region (AA 355-439). The N- and C-terminal domains and a small region within the central part of c-MYC are crucial for all its biological functions whereas so far no essential activity has been assigned to large regions in the central portion of the protein.

1.3.1 The amino-terminal domain (NTD) of c-MYC

The N-terminus of the c-MYC protein represents its transactivation domain containing a glutamine-rich region (AA 1-41), a proline-rich region (AA 42-103), and an acidic region (AA 104-143)⁵². Two conserved amino acid stretches unique to the MYC family are located within the N-terminus of c-MYC and have been termed MYC box I (MBI) and MYC box II (MBII). Both regions display very high homology between c-MYC proteins of different species as well as between distinct MYC family members.

MBI (AA 45-63) exhibits only minor transactivation qualities and is primarily involved in the regulation of c-MYC. It undergoes cell cycle-regulated phosphorylation with Thr 58 and Ser 62 representing the major phosphorylation sites, which fulfill opposing roles⁵³. Phosphorylation of Ser 62 by ERK coincides with RAS-induced stabilization of c-MYC⁵⁴. In contrast, phosphorylation of Thr 58 by GSK-3 β , which is dependent on the prior phosphorylation of Ser 62^{53,55}, targets c-MYC for ubiquitin-mediated proteasomal destruction^{54,56}. For degradation of dually phosphorylated c-MYC, Ser 62 is dephosphorylated by PP2A and phospho-Thr 58 thereby becomes a substrate for the FBXW7 E3 ligase⁵⁷. Binding of FBXW7 as a subunit of the SCF-complex to c-MYC stimulates its degradation via the 26S proteasome pathway. The protein stability of c-MYC is highly controlled and increased by RAS because RAS not only activates the RAF-MAPK-ERK pathway leading to phosphorylation of Ser 62 but concurrently inhibits GSK-3 β through PI3K preventing subsequent phosphorylation of Thr 58. The tight connection between RAS and c-MYC is reflected in their ability to transform cells when co-expressed. Mutation of Thr 58 strongly enhances whereas mutation of Ser 62 reduces c-MYC co-transformation with oncogenic *RAS*⁵⁵.

The hydrophobic MBII (AA 129-143) represents an interaction motif not constitutively exposed, which is bound by a large number of proteins^{58,59}. MBII is indispensable for the majority of the biological activities of c-MYC including transcriptional regulation^{60,61}, cell

transformation⁶², the block of differentiation⁶³, apoptosis¹², and c-MYC auto-regulation⁶⁴. Most importantly, this region is required for recruitment of TRRAP (transformation/transcription domain associated protein) and HAT (histone acetyltransferase) activity to enhancer box (E-box) elements during transcriptional activation⁶⁵.

1.3.2 The central region of c-MYC

The central region (AA144-354) is the least characterized part of the c-MYC protein. It harbors the NLS (nuclear localization signal) spanning amino acids 320 to 328⁶⁶ and the conserved MYC box III (MBIII)^{40,67}. MBIII (AA 188-199) is essential for full transactivation and transrepression of many c-MYC target genes⁶⁸ and contributes to proteasomal degradation of c-MYC⁶⁷. Furthermore, MBIII positively regulates the transformation potential of c-MYC *in vitro* and *in vivo* by attenuation of the induction of apoptosis⁶⁸. Recently, a fourth domain conserved in c-, N-, and L-MYC from fish to humans was identified and termed MYC box IV (MBIV)⁶⁹. The function of MBIV of c-MYC (AA 294-314) has not been determined but deletion of the corresponding region from mouse N-MYC leads to defects in MYC-induced apoptosis, transformation, and to a limited extend in c-MYC-mediated transactivation and transrepression⁶⁹.

1.3.3 The carboxy-terminal domain (CTD) of c-MYC

The C-terminus of c-MYC consists of a basic helix-loop-helix leucine zipper (bHLH-LZ) domain, which is a hallmark of several transcription factors⁷⁰. The basic region (AA 355-367) is followed by a HLH motif (AA 368-410) that covers two amphipathic helices separated by a loop. Residues from the bHLH region can contact the major groove of DNA and allow for interaction with other bHLH proteins like the c-MYC dimerization partner MAX. The LZ (AA 410-439) region of c-MYC with a heptad repeat of leucine residues constitutes an amphipathic α -helix and represents another well described domain for protein-protein interactions. The bHLH-LZ of c-MYC is necessary for DNA-binding^{27,71,72}, c-MYC induced transformation⁶², inhibition of cell differentiation⁶³, c-MYC auto-suppression⁶⁴, and the induction of apoptosis^{73,74}. In contrast to MBI, MBII, and MBIII, which all contain signals for rapid turnover of c-MYC^{67,75}, a so far unknown region within the C-terminus stabilizes the c-MYC protein⁷⁶.

All statements concerning the function of distinct domains of the c-MYC protein have to be re-considered with respect to the features of the short c-MYC-S isoform, which lacks the first 100 amino acids of c-MYC including MBI. c-MYC-S retains dimerization with MAX as well as DNA-binding activity⁴⁹ but shows strongly reduced transactivation capabilities⁷⁷ and behaves as a dominant-negative inhibitor of transactivation by full-length c-MYC⁴⁹. Despite these expected defects, c-MYC-S is still able to stimulate proliferation, transformation, apoptosis, and transcriptional repression and, surprisingly, rescues the slow growth phenotype (60 hours population doubling time compared to 20 hours for parental cells) of RAT1a fibroblasts with a homozygous deletion of *c-MYC*⁷⁸. Moreover, c-MYC-S displays the same half-life as c-MYC and is also degraded by the ubiquitin-proteasomal pathway despite the absence of MBI⁷⁹. Therefore, the majority of the functions of the c-MYC protein and even its degradation seem to be primarily determined by the region from MBII to the C-terminus of c-MYC.

1.4 The function of c-MYC as a transcription factor

The proto-oncogene *c-MYC* regulates a wide array of processes such as cellular proliferation, growth and competition, replicative potential, differentiation, stem cell self-renewal, angiogenesis, transformation, and apoptosis (reviewed in^{21,74,80-82}). The physiological effects of c-MYC are closely tied to its ability to both activate and repress genes by directly modulating their transcription⁵¹. A model for the function of c-MYC as an activator and repressor of transcription and the activity of antagonistic factors is depicted in Figure 2.

Amati *et al.* were the first to demonstrate a direct role for c-MYC as a transcription factor using a yeast transactivation assay⁸⁴. Consistently, activation of transcription by c-MYC over-expression in mammalian cells could be shown³¹ and the critical determinant of c-MYC-mediated transactivation was identified as the MAX protein, which associates with c-MYC *in vitro*^{23,85} and *in vivo*^{24,86}. c-MYC heterodimerizes with MAX in yeast, mice, and humans^{84,85,87}. Dimer formation not only is a prerequisite for c-MYC-induced transcriptional activation^{34,84} and repression^{88,89}, but also for oncogenic activity of c-MYC⁸⁷, which is mainly attributed to secondary effects due to changes in the expression of c-MYC target genes. Negative and positive regulation of transcription by c-MYC likely represent mechanistically distinct processes since some mutant alleles of c-MYC specifically affect either the capacity of c-MYC to activate or to repress genes^{13,90}.

INTRODUCTION

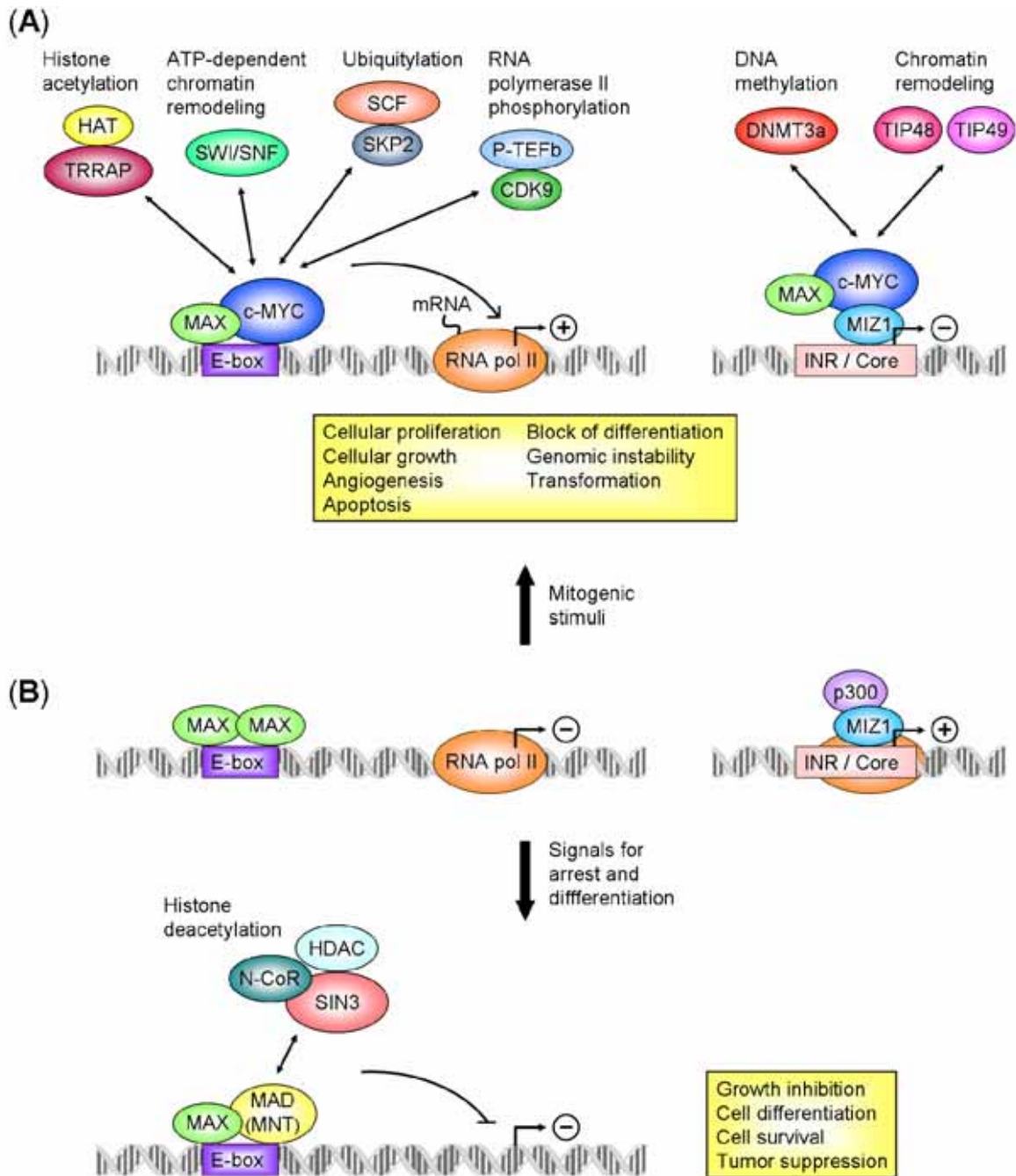


Figure 2. Model for c-MYC-mediated transcriptional regulation and antagonistic activity of MAD (adapted from ⁸¹ and ⁸³).

c-MYC/MAX-heterodimers control different sets of target genes by separate mechanisms and are antagonized by MAX/MAX- and MAD/MAX-dimers.

(A) In actively proliferating cells, E-box sequences are occupied by c-MYC/MAX-complexes, which transactivate the promoters of c-MYC target genes through recruitment of proteins involved in histone acetylation, chromatin remodeling, ubiquitylation, and phosphorylation. In parallel, transcription of cell cycle inhibitory genes is reduced by binding of c-MYC/MAX-complexes to transcription factors like MIZ1 thereby displacing the co-activator p300 from MIZ1 and inhibiting transcription from core and INR-containing promoter regions.

(B) Upon signals for growth arrest or differentiation, MAX/MAX- or MAD/MAX-complexes preferentially bind to E-box sequences and prevent c-MYC-mediated activation of pro-proliferative genes. In addition, MAD/MAX-complexes actively repress transcription through binding of the adaptor protein SIN3 associated with the co-repressor N-CoR and histone deacetylase complexes. In parallel, negative regulation of INR-containing and core promoters by c-MYC is reduced by release of c-MYC/MAX-complexes from MIZ1.

INTRODUCTION

The development of techniques for detecting differential gene expression such as SAGE, DNA microarray, and subtractive hybridization has led to the discovery of a vast collection of c-MYC target genes during the last years. According to expression studies and the mapping of genomic binding sites of MYC *in vivo*, c-MYC and dMYC are estimated to regulate about 10-15 % of all genes in the human and the *Drosophila* genome⁹¹⁻⁹³. The majority of the described target genes are activated while about 10-25 % of the genes are subject to negative regulation by c-MYC⁹⁴. The wide spectrum of target genes allows c-MYC to influence multiple biochemical pathways within the cell such as glucose metabolism, mitochondrial homeostasis, amino acid metabolism, translation initiation, DNA repair, and iron metabolism^{59,83,95}. To date, none of its known transcriptional targets is able to completely substitute for c-MYC in *c-MYC*-deficient rat fibroblasts⁹⁶ indicating that c-MYC is unique in its ability to influence manifold pathways.

Table 1 gives an overview of selected *bona fide* c-MYC target genes and depicts the functional relevance of these genes.

Table 1. Selected c-MYC target genes and their type of regulation by c-MYC.

Process	Gene	Regulation	Process	Gene	Regulation
<i>Auto-regulation</i>	<i>c-MYC</i>	Down ⁶⁴	<i>Cell growth</i>	<i>GADD45</i>	Down ¹⁰⁹
<i>Cell cycle</i>	<i>CDC25A</i>	Up ⁹⁷		<i>GAS1</i>	Down ¹¹⁰
	<i>CDK4</i>	Up ⁹⁸	<i>Cell adhesion</i>	<i>Collagen</i>	Down ¹¹¹
	<i>p15^{INK4B}</i>	Down ⁹⁹		<i>α3 integrin</i>	Down ¹¹²
	<i>p21^{CIP1}</i>	Down ^{100,101}		<i>Thrombospondin-1</i>	Down ¹¹²
	<i>p27^{KIP1}</i>	Down ¹⁰²	<i>Ribosome biogenesis</i>	<i>BN51</i>	Up ¹¹³
	<i>PDGFβ-R</i>	Down ¹⁰³		<i>Nucleolin</i>	Up ¹¹³
<i>DNA synthesis</i>	<i>CAD</i>	Up ^{104,105}	<i>Metabolism</i>	<i>ODC</i>	Up ^{114,115}
<i>Immortalization</i>	<i>hTERT</i>	Up ^{106,107}		<i>H-Ferritin</i>	Down ¹¹⁶
<i>Protein synthesis</i>	<i>EIF2A</i>	Up ¹⁰⁸		<i>IRP2</i>	Up ¹¹⁶
	<i>EIF4E</i>	Up ¹⁰⁸		<i>LDH-A</i>	Up ¹¹⁷

1.4.1 Activation of transcription

c-MYC/MAX-complexes function as sequence-specific positive transcription factors when bound to the hexameric E-box sequence CA(C/T)(G/A)(T/C)G within RNA polymerase II promoter regions²⁷ or the first intron of the target gene¹¹⁸. E-boxes flanked by a 5' cytosine and a 3' guanine are preferentially occupied by c-MYC/MAX-complexes¹¹⁹, but also direct

INTRODUCTION

c-MYC target genes without an E-box sequence exist⁸³. In any case, transcription from E-box and non-E-box promoters is enhanced through the transcriptional activation domain of c-MYC^{52,84} since MAX is transcriptionally silent¹²⁰. c-MYC recruits transcriptional co-activators to c-MYC/MAX-complexes such as BAF53A¹²¹, TRRAP¹²², and INI1 (hSNF5)¹²³ (Figure 2A). These three proteins are parts of both shared and distinct multiprotein complexes. BAF53A, as an integral component of the SWI/SNF-like complex, interacts with the ATP-dependent helicases TIP48 and TIP49⁸¹, which have both been described as binding partners and functional co-factors for c-MYC¹²⁴. TRRAP mainly associates with HAT-complexes containing either GCN5¹²⁵ or TIP60 in combination with TIP48 and TIP49^{126,127}. These enzyme complexes facilitate transcription by ATP-dependent chromatin remodeling and histone acetylation, which reduces the affinity of nucleosomes for DNA⁶⁵. Both processes alleviate the access of transcription factors or the movement of transcription complexes along the chromatin.

c-MYC furthermore drives transcriptional elongation through interaction with subcomponents of the P-TEFb kinase leading to C-terminal phosphorylation of RNA polymerase II and its clearance from promoter regions¹²⁸. For transcriptional activation of certain target genes, ubiquitylation of c-MYC seems mandatory and is achieved by direct recruitment of the E3 ligase SKP2 of the SCF-complex to c-MYC during G₁- to S-phase transition^{129,130}.

In case intracellular c-MYC protein levels are limited, E-boxes are occupied by MAX/MAX-homodimers antagonizing c-MYC function^{24,31,84} (Figure 2B). MAX/MAX-complexes are formed with lower affinity, exhibit lower sequence specificity for binding to DNA, and do not transactivate E-box promoters^{84,119,120,131}. Other factors like MAD, MNT, and MXI family proteins also compete with c-MYC for association with MAX especially during growth arrest and differentiation¹³²⁻¹³⁵. In heterodimeric complexes with MAX they actively repress transcription from E-box promoters by direct interaction with the co-repressor SIN3, which in turn recruits the transcriptional co-repressor N-CoR¹³⁶ and histone deacetylases HDAC1 and HDAC2^{81,137} (Figure 2B). HDAC1 and HDAC2 remove acetylated tails on core histones leading to nucleosomal condensation and transcriptional silencing¹³⁸. The function of MAD/MAX-complexes is not limited to antagonizing c-MYC transactivation as they also regulate genes distinct from c-MYC target genes and their DNA-binding specificity only partially overlaps with that of c-MYC/MAX-complexes¹³⁹.