302
Current Topics in Microbiology and Immunology

Editors

R.W. Compans, Atlanta/Georgia
M.D. Cooper, Birmingham/Alabama
T. Honjo, Kyoto · H. Koprowski, Philadelphia/Pennsylvania
F. Melchers, Basel · M.B.A. Oldstone, La Jolla/California
S. Olsnes, Oslo · M. Potter, Bethesda/Maryland
P.K. Vogt, La Jolla/California · H. Wagner, Munich

R. N. Eisenman (Ed.)

The Myc/Max/Mad Transcription Factor Network

With 28 Figures and 3 Tables

Robert N. Eisenman, Prof. Dr., Ph.D.

Division of Basic Sciences
Fred Hutchinson Cancer Research Center
P.O. Box 19024
Seattle, Washington 98108-1024
USA

e-mail: eisenman@fhcrc.org

Cover Illustration: The cover figure depicts a Myc-Max protein heterodimer interacting with its CACGTG binding site in DNA, superimposed over an image of DAPI-stained cell nuclei. The large cell nuclei are located within a Drosophila larval salivary gland ectopically expressing Drosophila Myc (dMyc) while the small nuclei are from fat body cells lacking ectopic dMyc (from Pierce SB, Yost C, Britton JS, Loo LW, Flynn EM, Edgar BA, Eisenman RN (2004) dMyc is required for larval growth and endoreplication in Drosophila. Development 131: 2317-2327). See chapter by Nair and Burley, within, for review of structural studies on Myc-Max dimers.

Library of Congress Catalog Number 72-152360

ISSN 0070-217X
ISBN-10 3-540-23968-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-23968-0 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September, 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com
© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
Product liability: The publisher cannot guarantee the accuracy of any information about dosage and application contained in this book. In every individual case the user must check such information by consulting the relevant literature.

Editor: Simon Rallison, Heidelberg
Desk editor: Anne Clauss, Heidelberg
Production editor: Nadja Kroke, Leipzig
Cover design: design & production GmbH, Heidelberg
Typesetting: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Printed on acid-free paper SPIN 11001973 27/3150/YL – 5 4 3 2 1 0

Preface

Scientists often look askance at their colleagues whose research appears too strongly focused on a single gene or gene product. We are supposed to be interested in the "big picture" and excessive zeal in pursuit of a single pixel might seem to border on an obsession that is likely to yield only details. However as this volume of *Current Topics in Microbiology and Immunology* demonstrates, this is certainly not the case for *myc*. Intense study of this enigmatic proto-oncogene over the last twenty years has only broadened our view of its functions and led to insights into mechanisms relating to transcriptional regulation as well as to cell growth, proliferation, differentiation, apoptosis and organismal development.

The *myc* gene originally came to light as a retroviral oncogene (v-*myc*) associated with a wide range of acute neoplasms. It was later shown to be a virally transduced cellular gene (c-*myc*) which is a member of family of oncogenes (c-*myc*, N-*myc*, L-*myc*). These family members are themselves subject to a bewildering assortment of genetic rearrangements associated with many different types of tumors derived from many different types of cells. These rearrangements (including chromosomal translocation, viral integration, and gene amplification) act to uncouple expression of the *myc* family genes from their normal physiological regulators. The chapter by LIU and LEVENS describes the key pathways leading to regulation of *myc* expression, showing that such regulation occurs at several different levels and through multiple mechanisms.

The early findings on *myc* regulation and its involvement in tumorigenesis suggested that *myc* plays a fundamental role in cell behavior and also served to attract a great deal of interest in understanding *myc*'s biological and molecular functions. One outcome of the strong research interest in *myc* was the realization that its encoded protein (Myc) does not function alone, but rather acts as part of a network, or module, of interacting proteins. Myc is a member of the of basic-helix-loop-helix-zipper (bHLHZ) class of proteins and forms a heterodimer with the bHLHZ protein Max. Myc-Max heterodimers recognize the sequence CACGTG and, with lower affinity, other related E-box

sequences. Binding of Myc-Max at E-boxes activates transcription from promoters in the vicinity of the binding sites. Underlying this transcriptional activity is the ability of Myc-Max to recruit several higher order chromatin modifying complexes to its binding sites, the major topic of the chapter by COLE and NIKIFOROV. Importantly Myc-Max have also been demonstrated to repress transcription of a number of genes, many of which are involved in cell cycle arrest and adhesion. The mechanism underlying Myc mediated repression is in part related to Myc's ability to interact with and inhibit the activity of other transcription factors, such as the BTB-POZ domain protein Miz-1. Repression by Myc and its biological consequences is described in the chapter by KLEINE-KOHLBRECHER, ADHIKARY, and EILERS.

The discovery of Max as an obligate dimerization partner for Myc led to the identification of other Max binding proteins. These include a novel group of bHLHZ proteins known as Mad proteins (now renamed Mxd), the closely related Mnt protein, and Mga. All these proteins have been associated with transcriptional repression at E-box binding sites. Indeed, Mad/Mxd and Mnt act as partial antagonists of Myc function. ROTTMANN and LÜSCHER review in detail the complex molecular and cellular biology of these proteins.

The Myc/Mad/Mad network then is defined by, and functions through, the interactions between individual Myc and Mad family proteins with Max as well as by interactions between Myc and Mad family proteins with higher order co-repressor and co-activator complexes. The structural biology of a number of these key interactions is reviewed in the chapter by NAIR and BURLEY who also discuss the basis for the high degree of specificity in complex formation. One concept that has emerged from these studies is the notion that a balance between Myc and Mad proteins may act to control key cellular events. A critical question that has haunted the Myc field for some time concerns the number and nature of the genes regulated by the network. This is the subject of the chapter by LEE and DANG who describe the approaches used the delineate target genes for Myc and how the thinking about Myc target genes has evolved.

Another major area of research interest relates to the biological consequences of both normal and abnormal Myc function. The chapter by WADE and WAHL describe evidence for the relationship between deregulated Myc protein expression and altered DNA repair and genomic instability. BLANCK, PIRITY and SCHREIBER-AGUS review the role of Myc/Max/Mad in embryonic development based on studies carried out in mice bearing targeted deletions of these genes. The chapter by GALLANT summarizes what research in invertebrate orthologs has taught us about the evolution of the Myc/Max/Mad network. Just as the Myc and Mad proteins do not function alone, the Myc/Max/Mad network is also unlikely to function in isolation. Mlx is a Max-like protein, which interacts with a subset of Max network

family proteins as well as with several bHLHZ proteins. BILLIN and AYER review the evidence for the Mlx network and describe its functions in energy metabolism.

The chapters included in this volume illustrate the complexities of the Myc/Max/Mad network and how its functions impinge on fundamental biological processes through regulation of transcription. I am grateful to all the authors for their efforts in putting together comprehensive and provocative chapters as well as for their patience during the long time it took for this volume to come to fruition. I thank Peter Vogt for suggesting the volume on the network and Ms. Anne Clauss for her help in assembling the volume. I hope that the exciting research described here will stimulate others to explore the functions of transcription factor networks.

Seattle, Washington, July 2005 *Robert N. Eisenman*

List of Contents

Making Myc 1
J. Liu and D. Levens

Transcriptional Activation by the Myc Oncoprotein 33
M. D. Cole and M. A. Nikiforov

Mechanisms of Transcriptional Repression by Myc 51
D. Kleine-Kohlbrecher, S. Adhikary, and M. Eilers

The Mad Side of the Max Network: Antagonizing the Function of Myc and More . 63
S. Rottmann and B. Lüscher

Structural Aspects of Interactions Within the Myc/Max/Mad Network 123
S. K. Nair and S. K. Burley

Myc Target Transcriptomes 145
L. A. Lee and C. V. Dang

c-Myc, Genome Instability, and Tumorigenesis: The Devil Is in the Details 169
M. Wade and G. M. Wahl

Lessons Learned from Myc/Max/Mad Knockout Mice 205
M. Pirity, J. K. Blanck, and N. Schreiber-Agus

Myc/Max/Mad in Invertebrates: The Evolution of the Max Network 235
P. Gallant

The Mlx Network: Evidence for a Parallel Max-Like Transcriptional Network
That Regulates Energy Metabolism 255
A. N. Billin and D. E. Ayer

Subject Index 279

List of Contributors

(Addresses stated at the beginning of respective chapters)

Adhikary, S. 51
Ayer, D. E. 255

Billin, A. N. 255
Blanck, J. K. 205
Burley, S. K. 123

Cole, M. D. 33

Dang, C. V. 145

Eilers, M. 51

Gallant, P. 235

Kleine-Kohlbrecher, D. 51

Lüscher, B. 63
Lee, L. A. 145
Levens, D. 1
Liu, J. 1

Nair, S. K. 123
Nikiforov, M. A. 33

Pirity, M. 205

Rottmann, S. 63

Schreiber-Agus, N. 205

Wade, M. 169
Wahl, G. M. 169

CTMI (2006) 302:1–32
© Springer-Verlag Berlin Heidelberg 2006

Making Myc

J. Liu · D. Levens (✉)

Gene Regulation Section, Laboratory of Pathology, NCI, DCS, Bldg. 10, Rm 2N106, Bethesda, MD 20892-1500, USA
levens@helix.nih.gov

1 *c-myc* Regulation 2
1.1 The Problem 2
1.2 Do Myc Levels Matter? 2
1.3 Promoters 4
1.4 Chromatin Changes 7
1.5 Where Are *c-myc* Regulatory Elements? 9

2 *cis*-Elements and Transacting-Factors Regulating *c-myc* Expression 9
2.1 *cis*-Elements 9
2.2 Traditional *trans*-Acting Factors 10
2.3 Funny DNA: The Role of Topology and Conformation 14
2.4 Strange Factors 15
2.5 FUSE-Binding Protein 18
2.6 FBP-Interacting Repressor 19
2.7 Special AT-Rich Binding Protein 1 19
2.8 A Scheme to Regulate *c-myc* Transcription 20

References 22

Abstract Myc regulates to some degree every major process in the cell. Proliferation, growth, differentiation, apoptosis, and metabolism are all under Myc control. In turn, these processes feed back to adjust the level of *c-myc* expression. Although Myc is regulated at every level from RNA synthesis to protein degradation, *c-myc* transcription is particularly responsive to multiple diverse physiological and pathological signals. These signals are delivered to the *c-myc* promoter by a wide variety of transcription factors and chromatin remodeling complexes. How these diverse and sometimes disparate signals are processed to manage the output of the *c-myc* promoter involves chromatin, recruitment of the transcription machinery, post-initiation transcriptional regulation, and mechanisms to provide dynamic feedback. Understanding these mechanisms promises to add new dimensions to models of transcriptional control and to reveal new strategies to manipulate Myc levels.

1
c-myc Regulation

1.1
The Problem

As discussed elsewhere in this volume, Myc-Max heterodimers operating directly as a transcription factor recruit effector complexes to activate and repress transcription. Alternatively, by binding with E-boxes, Myc-Max competes with the other HLH-bZIP complexes to modify target gene action. Counted among Myc targets are genes essential for proliferation, growth, the cell cycle, apoptosis, metabolism, and both intra- and intercellular signaling. Thus the c-Myc network ensnares prey from virtually every important cellular activity (Grandori et al. 2000; Levens 2002, 2003). In turn, it would seem that *c-myc* expression should be coupled with direct or indirect feedback from many intra- and extracellular systems and subsystems. These systems regulate Myc at every level—from transcription, RNA processing, messenger (m)RNA half-life, and translation to protein turnover (Cole and Mango 1990; Wisdom and Lee 1991; Laird-Offringa 1992; Lavenu et al. 1995; Yeilding and Lee 1997; Brewer 1999; Creancier et al. 2001; Lemm and Ross 2002; Kim et al. 2003b). Although a number of factors bind to *c-myc* mRNA to influence its turnover and translation, it appears that most *c-myc* regulatory pathways are channeled through transcriptional control (though not necessarily exclusively so). Indeed, as will be discussed below, many important pathways reach the *c-myc* promoter through a variety of canonical, non-canonical, and atypical *cis*-elements. The central and elusive problem in *c-myc* regulation is discerning how multiple, and often disparate, signals are integrated to determine the final level of Myc.

1.2
Do Myc Levels Matter?

A number of observations indicate that cellular and organismal physiology and pathology are sensitive to slight alterations of Myc levels. The body sizes of mice bred to generate every diploid combination of normal, hypomorphic, and null *c-myc* alleles scaled with the amount of Myc; only the null-homozygotes were inviable, succumbing during development as reported (Davis et al. 1993; Trumpp et al. 2001). Somatic knockout of one or both *c-myc* alleles showed that the cell cycle length varies inversely with the dose of *c-myc* (Shichiri et al. 1993; Mateyak et al. 1997; Schorl and Sedivy 2003). These same studies indicate that there is no upregulation of the normal *c-myc*

allele to compensate for the impaired or absent expression of its partner; there is no *c-myc* signal upregulating basal expression acting in *trans*. In contrast to underexpression, it is likely that Myc overexpression depresses expression from normal *c-myc* alleles (first appreciated by the silence of the unrearranged allele in Burkitt lymphoma, this phenomenon is not universal in all cases of the disease) (Siebenlist et al. 1984; Facchini et al. 1997). A variety of studies indicate that Myc-targeted gene expression varies quantitatively, if not qualitatively, as Myc levels are altered (Levens 2002, 2003). Chromosomal translocations, rearrangements, and viral insertions that deregulate *c-myc* expression without activating mutations within the Myc protein indicate that failing to confine Myc levels within physiological bounds is an important step in the carcinogenesis of many, if not most, tumors. Even in Burkitt lymphoma, the malignancy most closely associated with abnormal *c-myc* expression, the range of *c-myc* expression in some cases barely exceeds (1.47-fold) the levels found in normal tissues (Saez et al. 2003). It seems that Myc levels do matter.

The kinetic features of *c-myc* mRNA and protein indicate that the protein must be tightly regulated. Both the mRNA and the protein possess short half-lives (20–30 min), but may be stabilized in some pathological or physiological circumstances (Hann and Eisenman 1984; Dani et al. 1985; Rabbitts et al. 1985; Sears et al. 1999). In normal resting cells, *c-myc* mRNA levels are low: as low as one molecule per cell. It has been estimated that about 40% of the mRNAs in cells are single copy (Lockhart and Winzeler 2000). If these are Poisson distributed, as expected, at any given moment every cell has a unique expression profile. Rapidly dividing tissues support higher levels of *c-myc*, but even in embryonic cells, *myc* RNA levels have been estimated to be approximately five mRNAs per cell (Evingerhodges et al. 1988; Warrington et al. 2000). During mitogenic stimulation of normal cells or in tumors, *c-myc* transcripts may transiently rise to higher levels. During the G0–G1 transition, *c-myc* mRNA levels spike before dropping to steady-state levels (Dean et al. 1986). Apparently less Myc is required to sustain than to initiate proliferation; if so, fluctuating Myc expression might prove deleterious. However, because of rapid turnover and low abundance, it would seem that cells lack a sufficient reservoir to buffer the stochastic noise expected to buffet *c-myc* mRNA levels (Elowitz et al. 2002; Swain et al. 2002). Somehow all of the signals converging on *c-myc* must be integrated in a manner ensuring sufficient stability to prevent abnormal proliferation or apoptosis, yet responsive enough to allow Myc to fulfill its role as an immediately-early gene.

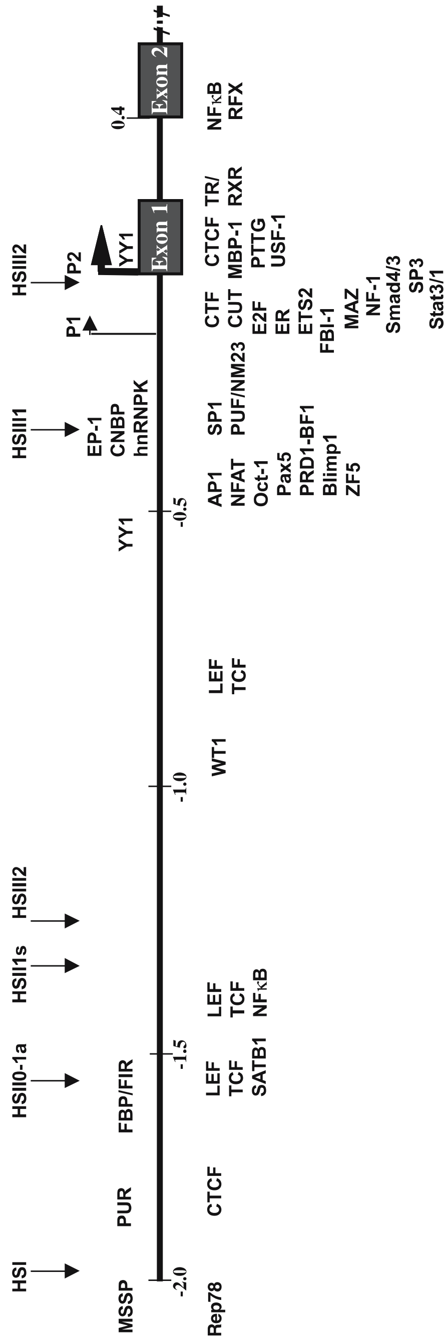
1.3 Promoters

Under most circumstances, *c-myc* transcription is initiated at two promoters, P1 and P2, with the latter supporting about 75% of c-*myc* transcripts (see Fig. 1). Minor amounts of *c-myc* RNA initiate at P0 and P3, and cryptic promoters may become activated following chromosomal translocation, under the influence of pathologically juxtaposed regulatory elements (Spencer and Groudine 1991; Marcu et al. 1992). The discovery of antisense transcription of the murine and human *c-myc* exon 1 presaged the more recent appreciation of opposite-strand transcription as a general phenomenon (Spencer and Groudine 1991, Rinn 2003; Marcu et al. 1992). Usually, only P2-initiated—and to a lesser extent P1-initiated—transcripts contribute significantly to the pool of *c-myc* mRNAs. Under some physiological circumstances (e.g., G0–G1 transition in lymphocytes—Broome et al. 1987) or pathological conditions (e.g., following translocations in Burkitt lymphoma), initiation at P1 nearly equals that at P2. If the minor promoter-initiated—or antisense—transcripts are physiologically relevant, they will most likely serve regulatory roles, contributing to RNA processing or stability [through alternative secondary structures or formation of microRNA (miRNA)] or altering transcription-driven chromatin remodeling and modification. P0-initiated transcription complexes must traverse through *cis*-elements upstream of P1 and P2, and so may potentially contribute to the structural reorganization and exchange of promoter-bound regulatory protein complexes. These possibilities remain largely unexplored. The P2 promoter itself is remarkably resistant to inactivating mutations and deletions. Eliminating both the TATA box and the transcription start-site preserves sufficient information for the transcription machinery to still locate the promoter and initiate transcription (Krumm et al. 1995). But at *c-myc*, most of the time, initiation is not the problem.

Unless the *c-myc* gene has been irreversibly silenced, a transcriptionally engaged RNA polymerase is paused in the promoter proximal region in most cells (Marcu et al. 1992). The presence of this polymerase dictates that *c-myc*

→

Fig. 1 Anatomy of c-myc promoter's DNase I hypersensitive sites and putative *trans*-factor binding sites: The DNA region from ~2.5 kb upstream of P2 to c-*myc* the second exon is shown. Locations of binding sites for over 40 factors directly binding to DNA are indicated relative to the human P2 promoter. Specific binding sites for each factor are listed in Table 1. Conventional double-stranded DNA-binding proteins are listed *below the line* and single-stranded nucleic DNA-binding proteins are listed *above the line*. Exon III and downstream hypersensitive sites are not shown


c-myc anatomy
HSI
HSII0-1a
HSII1s
HSIII2
HSIII1
HSIII2
P1
P2
MSSP
PUR
FBP/FIR
YY1
EP-1
CNBP
hnRNPK
YY1
Exon 1
Exon 2
-2.0
-1.5
-1.0
-0.5
0.4
Rep78
CTCF
LEF
TCF
SATB1
LEF
TCF
NFκB
WT1
LEF
TCF
AP1
NFAT
Oct-1
Pax5
PRD1-BF1
Blimp1
ZF5
SP1
PUF/NM23
CTF
CUT
E2F
ER
ETS2
FBI-1
MAZ
NF-1
Smad4/3
SP3
Stat3/1
CTCF
MBP-1
PTTG
USF-1
TR/
RXR
NFκB
RFX

expression requires a cycle of promoter escape with the transition to elongation before reinitiation occurs. De novo activation from the silent state is not the usual situation for *c-myc* transcription. Paused RNA polymerases, such as first identified at the *hsp70* promoter, customarily have been considered poised for a rapid response (Rougvie and Lis 1988). Although pre-positioning the RNA polymerase might accelerate Myc induction, recent analysis of transcription factor and polymerase dynamics in vivo indicates that DNA–protein interactions are rapid and not usually rate limiting (McNally et al. 2000). In vitro, even in the absence of activators, neither preinitiation complex formation nor elongation is often the rate-limiting step in the transcription cycle. On some promoters the transition from initiation to elongation—promoter escape—is rate limiting (Kugel and Goodrich 1998). The pausing of RNA polymerase at promoters has additional implications for gene expression. Unless activated or removed, a paused polymerase trumps the action of all factors operating to recruit the basal machinery. So factors acting to delay promoter escape provide a check against spurious activation of a vacant promoter.

The site(s) of polymerase pausing on the *c-myc* promoter has not been rigorously defined and the nascent transcripts sprouting from the paused polymerase have never been isolated. Nuclear run-on experiments employing different combinations of nucleotide triphosphates to advance the polymerase incrementally revealed variable pause sites ranging over approximately the first 50 nucleotides (Wolf et al. 1995). Furthermore, the *cis*-elements imposing the pause have also not been rigorously identified. Whether local promoter sequences define intrinsic pause sites (as occurs with RNA polymerase alone), or distant sequences recruit pause-controlling *trans*-factors, is not known (Pal et al. 2001). The control of pausing may be linked with conformational changes demanded by the transition form initiation to elongation. The requirement for DNA melting at all promoters is self-evident. How this melting occurs has not been fully revealed. The structures of all RNA polymerase-template complexes that have been solved (whether phage, bacterial, or polymerase II) reveal a sharp bend with the active site (Zhang et al. 1999; Cramer et al. 2000; Gnatt et al. 2001; Tahirov et al. 2002; Yin and Steitz 2002). Bent DNA melts more easily, and so this feature of transcription complexes may help to open the duplex to permit pairing with the incoming ribonucleotide triphosphates in preparation for phosphodiester bond formation (Kahn et al. 1994). Whether the melted region at the *c-myc* start site is composed only of a transcription bubble sequestered entirely within the active site of RNA polymerase II is not known. When the last general transcription factor, TFIIH, joins the preinitiation complex it carries along the XPB/p89/ERCC3 3′-5′ and XPD/p80/ERCC2 5′-3 helicases; the former is essen-

tial for initiation and promoter escape, whereas the latter, though inessential for transcription, facilitates promoter escape (Zawel et al. 1995; Hoeijmakers et al. 1996; Ohkuma 1997; Coin and Egly 1998; Frit et al. 1999; Tirode et al. 1999; Akoulitchev et al. 2000). The helicases of the TFIIH core function during nucleotide excision repair to expose damaged bases for removal. TFIIH has been proposed to act as a molecular wrench modifying DNA conformation at start sites from a downstream location, though alternative models for TFIIH action at transcription start sites exist (Robert et al. 1998; Douziech et al. 2000; Kim et al. 2000). The roles of the helicases in modifying promoter structure have been relatively less studied than the role of the TFIIH CAK (cyclin-activating kinase/cdk7) subcomplex. CDK7, the kinase within CAK, plays a major role phosphorylating the hepta-residue repeat comprising the carboxyl terminal domain (CTD) of the large subunit of RNA polymerase II. Depending on the state of CTD phosphorylation, additional factors involved in transcription and RNA processing (including capping) are recruited to early transcription complexes. Several factors regulating *c-myc* transcription [e.g., FUSE-binding protein (FBP), FBP-interacting repressor (FIR), estrogen receptor E2] interact with TFIIH (Pearson and Greenblatt 1997; Chen et al. 2000; Liu et al. 2000, 2001; Keriel et al. 2002). Shortly after initiation, factors such as DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) are recruited and contribute to pausing, at least to some promoters such as the *Drosophila HSP70* promoter (Wu et al. 2003). It should be stressed that the mechanistic role of TFIIH, DSIF, and NELF, as well as other factors involved in promoter escape, have been explored only on a very small number of promoters; whether the details of the disposition of these factors during initiation and promoter escape can be generalized to all promoters from these few cases is not known.

A role of the nascent *c-myc* RNA in regulating transcript growth, either by directly manipulating the transcription apparatus or through the recruitment of sequence or structure-specific RNA-binding proteins [as for human immunodeficiency virus (HIV) via the TAR sequence in the nascent transcript, the TAT protein recruits PTEF, cyclin T-CDK9, which phosphorylates the CTD and stimulates elongation], has not been reported, but should be considered (Garber et al. 1998; Zhou et al. 1998).

1.4
Chromatin Changes

c-myc was one of the first genes analyzed by indirect labeling for DNase I hypersensitive sites (Siebenlist et al. 1984). Constitutive and regulated hypersensitive sites upstream, downstream, and within the *c-myc* gene have been

mapped (Fig. 1). Constitutive DNase hypersensitive site (HS) I is associated with the binding of CCCTC-binding factor (CTCF). This remarkable multi-zinc-finger protein associates with several regions of the *c-myc* gene, including the promoter, under repressed conditions. Since CTCF has enhancer-blocking activity, it is likely to play a role in eliminating or selectively gating the influence of more distant regulatory proteins on the *c-myc* promoter. In cases of Burkitt lymphoma with far upstream translocations, the influence of the immunoglobulin enhancer must penetrate HS I by an unknown mechanism to negate CTCF's barrier function. CTCF also makes protein–protein interactions with other *c-myc* regulators, such as YB1 (Filippova et al. 1996; Chernukhin et al. 2000; Ohlsson et al. 2001; Qi et al. 2003). Although HS II_1 and II_2 map to upstream regions binding various factors, the agents responsible for conferring hypersensitivity have not been unambiguously assigned. In the case of site II_2, cleavage occurs within the CT-element, an entangled mess of overlapping and inter-nested binding sites for conventional and single-strand selective factors; these sites and factors are not easily functionally or biochemically deconvoluted. Sites III_1 and III_2 overlap the P1 and P2 promoters; again, the agent of hypersensitivity has not been ascribed to any single protein or complex, although the ME1a1 site is implicated (Albert et al. 2001). Additional HS sites map 3′ of *c-myc* (Mautner et al. 1995). Whereas some hypersensitive sites are enhanced when *c-myc* is expressed, other sites, such as HS I, persist even in cells with irreversibly silenced *c-myc*.

Actively transcribed *c-myc* genes carry 15 nucleosomes in a 3.6-kb array stretching from upstream of the promoter into intron 1. When inactive, *c-myc* genes harbor an additional four nucleosomes masking segments near HS I, the CT-element, P0, and P1. The major P2 start-site is nucleosome-free irrespective of gene activity (Michelotti et al. 1996b; Pullner et al. 1996; Albert et al. 1997; Schuhmacher et al. 1999). The hypersensitive sites and nucleosome arrangement of *c-myc* promoters embedded in episomal vectors recapitulate those of endogenous *c-myc* genes (Michelotti et al. 1996b; Albert et al. 1997; Madisen et al. 1998; Albert et al. 2001).

The role of chromatin-modifying and remodeling complexes in regulating *c-myc* expression is complicated and confusing. Whereas histone deacetylase (HDAC) inhibitors augment expression from transfected or transgenic *c-myc* promoters consistent with the notion that increased acetylation supports increased transcription, HDAC inhibition paradoxically depresses endogenous *c-myc* in most situations (for example, see Van Lint et al. 1996; Chambers et al. 2003; and many others). The dynamics of HAT, HDAC, and remodeling complex recruitment and dismissal during *c-myc* gene induction and shut-off are incompletely described and likely to prove complicated.

1.5 Where Are *c-myc* Regulatory Elements?

A Hind III site at −2329 relative to P1 has often served as the operational upstream boundary of the *c-myc* promoter; a Pvu II site at the end of exon 1 usually delimits the downstream segment, although sometimes the 5′ portion of intron 1 is also included. Most characterized *cis*-elements and candidate *cis*-elements have been mapped to this interval. Although convenient, these arbitrary choices are poorly justified. *c-myc* promoter-driven reporter genes—whether transiently or stably transfected, integrated, or episomal, as well as transgenes passaged through the germline in mice—have failed to recapitulate proper *c-myc* expression, although certain features of *c-myc* transcription have been coarsely mimicked. Embedding *c-myc* reporters and transgenes in 30 kb or even 50 kb of natural flanking sequence has proved insufficient to confer physiological regulation (Lavenu et al. 1994; Mautner et al. 1996). Why is proper *c-myc* regulation so difficult to achieve? Perhaps important *cis*-elements reside at vast distances from the coding sequence, or perhaps the *c-myc* promoter is particularly sensitive to perturbation of its natural chromosomal context and so requires proper boundary elements to define chromatin and topological domains. Supporting the argument that context is key for *c-myc* governance is the extreme vulnerability of the locus to chromosomal damage, even from vast distances. In the case of Burkitt lymphoma, translocations hundreds of kilobases upstream, or downstream, as well as within the gene, deregulate transcription. In these cases, the cytogenetically juxtaposed, but molecularly remote, immunoglobulin enhancer overrides or usurps all of the locally acting elements with their associated factors to enforce *c-myc* expression. So either context is paramount or vital and remote elements operate on the promoter from either side. Besides translocations, viral insertions, gene amplification, and mutations all deregulate *c-myc* expression.

2 *cis*-Elements and Transacting-Factors Regulating *c-myc* Expression

2.1 *cis*-Elements

There is no evidence for a compact enhancer that confers or explains the physiological patterns of *c-myc* expression. There is no evidence for the assembly of a precisely arranged enhanceosome composed of multiple transcription

factors and architectural DNA-binding proteins, as well as chromatin remodeling and modifying complexes (Thanos and Maniatis 1995). Virtually every major signal transduction pathway impacts directly or indirectly the *c-myc* promoter (Table 1). Some of the *cis*-elements receiving the signals have been well characterized, whereas others have been revealed only in silico. The *c-myc* promoter generally lacks canonical binding elements while relying on atypical binding sites to recruit many of the *c-myc* regulatory proteins. Generally, non-canonical sites are suboptimal for binding transcription factors. Most of the activators and repressors that bind *c-myc cis*-elements recruit coactivators or corepressors, at least in vitro, and in some cases, chromatin immunoprecipitation studies have demonstrated these effector complexes at *c-myc* in vivo. Reliance upon non-canonical *cis*-elements to recruit these effectors has several implications. First, the weak binding may contribute to the observation that *c-myc* levels are adjusted several fold by many, perhaps even most, agents, but very few single signals impel changes in Myc levels of sufficient degree to constitute an on–off switch. Second, higher concentrations of each factor may be required to achieve *cis*-element occupancy, and so a strong or sustained stimulus might be required to activate expression. Third, the stabilization of weak binding factors to their *cis*-elements through cooperatively interacting partners in principle serves to cross-couple signals and promote synergy. Fourth, fractional *cis*-element occupancy also confounds in vivo footprinting and other protection studies that work best at saturation.

A number of *cis*-elements may be densely inter-nested with the promoter. Sequences responsible for negative autoregulation (probably occurring through both direct and indirect mechanisms), as well as sites for binding CTCF, MBP-1 (a protein related to enolase) (Ray and Miller 1991; Subramanian and Miller 2000; Lee et al. 2002), and other factors occur so close to start sites that cohabitation is difficult to imagine; sequential or alternative action at the promoter seems more likely.

2.2 Traditional *trans*-Acting Factors

The literature describing the pathways delivering signals to the *c-myc* promoter via conventional transcription factors constitutes a veritable compendium of gene regulatory phenomena. Although in any one setting or cell line a particular pathway may dominate, in other situations the influence of that same pathway may be minimal or irrelevant. Signaling pathways striking *c-myc* include: MAP kinase, JAK/STAT, Ras, IFN-γ PI3-K, Fas, Wnt, TGF-β, interleukins, cytokines, lymphokines, steroid and peptide hormones, pharmacologic agents, NF-κB-activating pathways, E2F-activating pathways,

Table 1 List of putative transcription factors regulating *c-myc* transcription. Transcription factors directly binding to the *c-myc* promoter are listed alphabetically. Their binding sites are relative to the human P2 transcription start site. Also listed are signal transduction pathways influencing the binding or dissociation of factors to the *c-myc* promoter. A considerable number of factors bind to non-canonical sites

Factor	Binding site (relative to human P2)	Signaling	Canonical	Reference(s)
AP1	(−467/−461)	MAPK, JAK/Stat, TGF-β, FasL/CD95	NC	Takimoto et al. 1989; Hay et al. 1987; Shaulian and Karin 2002; Cippitelli et al. 2003
CEBPα	Through E2F			Johansen et al. 2001
CNBP	CT (−316/−278)		N/A	Michelotti et al. 1995
CTCF	TRE (+2/+48)	TGF-β, MAPK	N/A	Lutz et al. 2002; Ohlsson et al. 2001
CTF	(−61/−47)	TGF-β, JAK/Stat, MAPK	NC	Gronostajski 2000
CUT	ME1a1 (−61/−47)		C	Dufort and Nepveu 1994
E2F	(−77/−69)	Rb, TGF-β, MAPK, focal adhesin signaling	C	Thalmeier et al. 1989; Ogawa et al. 2002; Iakova et al. 2003; Kowalik 2002
ER	(−138/−22)	Ras/Raf	NC	Dubik and Shiu 1992; Shang et al. 2000
ETS2	(−77/−69)	Ras/Raf/MEK/ERK, JAK/Stat	C	Roussel et al. 1994; de Nigris et al. 2001
FBI-1	(−197/−183)		N/A	Hernandez and Pessler 2003
FBP/FIR	FUSE (−1741/−1652)	Wnt	N/A	Michelotti et al. 1996b; Liu et al. 2000, 2001; Braddock et al. 2002
hnRNPK	CT (−316/−278)		N/A	Michelotti et al. 1996a,b

Table 1 (continued)

Factor	Binding site (relative to human P2)	Signaling	Canonical	Reference(s)
LEF	TBE (−1528/−1415, −1330/−1321, −763/−754)	Wnt, PI3 K	NC(?)	Reya et al. 2000; Peifer 2002
MAZ	ME1a1 (−61/−47), ME1a2 (−95/−81)		C	Bossone et al. 1992; Song et al. 2001
MBP1	(−20/−15)		N/A	Ray and Miller 1991; Ray et al. 1995
MSSP1/2	(−2355/−2348)		N/A	Fujimoto et al. 2001
MYB	(strong site at −1154, multiple weak sites)		C and NC	Zobel et al. 1991; Schmidt et al. 2000
NF-1	(−61/−47)	TGF-β, JAK/Stat, MAPK	NC	Gronostajski 2000
NF-κB	(−1392/−1299, +285/+295)	NF-κB signaling, MAPK, INF-γ, BCR Signaling JAK/Stat, SAPK/JNK		Grumont et al. 2002; Shaffer et al. 2001; Arcinas et al. 2001; Jeay et al. 2001
NFAT	(−467/−461)	Ras, JNK/Stat		Rao et al. 1997
Oct-1	(−467/−461)	MAPK, JAK/Stat, Ras, DNA damage	C	Takimoto et al. 1989; Zhao et al. 2000
Pax5	(−453/−428)		N/A	Pasqualucci et al. 2001
PRD1-BF1/Blimp1	(−453/−428)	Jak/Stat Pathway, Fas/CD95	C	Lin et al. 1997; Gupta et al. 2001
PTTG	(−4/+20)	MEK1	N/A	Pei 2001
PuF/NM23	CT (−316/−278)		N/A	Michelotti et al. 1997
PUR	(−1808/−1789)		N/A	Bergemann and Johnson 1992
Rep78	<−2 kb		N/A	Hermonat 1994
RFX	MIE1 (+350/+363)	Jak/Stat,	C	Zhang 1993

Table 1 (continued)

Factor	Binding site (relative to human P2)	Signaling	Canonical	Reference(s)
SATB1	(−1531/−1507)		N/A	Cai et al. 2003
Smad4/3	TIE (−84/−75)	TGF-β	NC	Kowalik 2002; Chen et al. 2002
Sp1	CT (−316/−278), ME1a1 (−61/−47)	MAPK, INF-γ	NC	Harris et al. 2000
Sp3	ME1a1 (−61/−47)		C	Majello et al. 1997
Stat3/1	(−81/−73)	Jak/Stat Pathway, p38 MAPK, FasL		Bowman et al. 2001; Jenab and Morris 1997
TR/RXR	(Human: −2.2 kb, mouse: +100/+133)	MAPK, PKC	NC	Lutz et al. 2003; Perez-Juste et al. 2000
TCF	TBE (−1528/−1415, −1330/−1321, −763/−754)	Wnt, PI3 K	NC(?)	Peifer 2002; Reya et al. 2000; He et al. 1998
USF-1	(−20/+28)	MAPK	NC	Kiermaier 1999
VDR				Simpson et al. 1987; Katayama et al. 2003
WT1	(−918/−911)	Ras, ERK		Hewitt 1995
YY1	(−553/−423, −20/+28)	TGF-β, MAPK	NC	Riggs et al. 1993; Kurisaki et al. 2003
ZF5	(−453/−403)		N/A	Numoto et al. 1993; Kaplan and Calame 1997

C, canonical;
N/A, not available;
NC, non-canonical

etc. Many of these signals branch and influence more than one *trans*-factor. Each of the *c-myc trans*-factors merits a separate review; some of them are listed in Table 1.

2.3 Funny DNA: The Role of Topology and Conformation

Work from multiple laboratories has contributed to the notion that nonstandard transcription factors binding at atypical binding sites participate in the transcriptional regulation of *c-myc*. Several regions of the *c-myc* gene are associated with non-B-DNA conformation. To understand how these *cis*-elements and their *trans*-factors operate it is important to consider what drives the formation of non-B-DNA.

Conceptually, several processes may directly or indirectly drive conformational changes occurring at *c-myc cis*-elements. First, unwinding torsional stress (negative supercoiling) destabilizes duplex DNA, and certain regions of DNA, especially segments with high A–T content, preferentially melt when the unwinding torque is high enough. Within a topological domain, each hotspot for melting competes with every other hotspot, so the response to torque is inextricably coupled with the creation and destruction of topological boundaries (Benham 1992; Fye and Benham 1999). Protein–protein interactions between DNA-bound factors or attachments to immobile structures restricting rotation of DNA along its helical axis impose topological borders. Topological domains may be nested. Wrapping and fixation of DNA around a nucleosome restrains approximately one supercoil, and unless the grip of the nucleosome is breeched, this DNA constitutes a separate topological domain (Sinden 1994). As long as this DNA is firmly held, the entire protein-DNA assembly may be rotated en bloc, transmitting stress to the unrestrained linker regions. Loosening of histone tails secondary to chromatin modifications such as acetylation would be predicted to expand the amount of linker DNA available to accommodate torsion (Norton et al. 1990; Morales and Richard-Foy 2000). Thus the particular chromatin arrangement within DNA loops may help to focus torsional strain onto DNA segments predisposed to melting or forming alternative structures such as Z-DNA (Rich and Zhang 2003). Second, helicases expend energy from ATP to open bound segments of duplex. As noted previously, the helicase of TFIIH contributes to transcription regulation; whether TFIIH might also contribute to DNA melting at elements other than start sites (directly or as torque generator acting from the promoter) has not been explored except during DNA repair. Several dozen helicase-like open-reading frames reside in the human genome. While many of these are generally presumed to be RNA, rather than DNA, helicases, it is

premature to ascribe a molecule with certainty to one class or the other in the absence of experimental information (Caruthers and McKay 2002). So the possibility that helicases may be recruited to *cis*-elements to facilitate the transition to non-B-DNA remains plausible. Third, the machinery of transcription, replication, recombination, and repair demands transient single-stranded regions at the sites of catalytic activity, apart from supercoil-induced melting driven by complexes translocating along DNA. Auxiliary single-stranded (ss) DNA-binding proteins, chromatin remodeling machines, and topoisomerases accompany each of these genetic transactions, and so all have the potential to reconfigure those *cis*-elements prone to altered states. Fourth, although it is assumed that homeostatic mechanisms maintain a monotonous chemical and physical intranuclear environment, changes in parameters that alter the stability of B-DNA such as ionic strength, pH, divalent cations, polyamines, and temperature may all conspire to alter DNA structure. Utilization of elements responsive to these parameters in principle would directly couple *c-myc* transcription with the maintenance of intranuclear homeostasis (note that the sensitivity of PCR to slight variations of these parameters illustrates that it may not be too far fetched to conceive of physiological or pathological changes in DNA structure due to changes of the intranuclear milieu). Cells embedded in tissues are also subject to considerable mechanical force. If these forces were transmitted to DNA via anchored chromatin, DNA structure could be affected at susceptible sequences. So in principle DNA elements that adopt non-B conformations may serve as *cis*-acting stress sensors acting concertedly with conventional *trans*-acting stress-sensing pathways.

2.4 Strange Factors

Besides the panoply of well-recognized, well-characterized transcription factors binding duplex DNA and operating through conventional mechanisms, *c-myc* promoter recruits a menagerie of strange gene regulators binding to elements assuming unusual DNA structures and conformations. Two regions of *c-myc* sequence are particularly associated with altered DNA structures. First, the CT-element, found 100 to 145 bp upstream of the P1 promoter, has been reported to adopt H-DNA, tetraplex, and single-stranded conformations in addition to the standard B-form duplex (Kinniburgh 1989; Postel 1992; Michelotti et al. 1996b; Simonsson et al. 1998). Each of these states is associated, at least in vitro, with a set of conformation-sensitive binding proteins. The extent of regulatory input in vivo conferred by a particular conformational state and its associated factors is not known. Altered structures or conformation in the region of the CT-element is compatible with the absence of phased

nucleosomes in this region; nucleosomes constrain B-DNA but have not been shown to engage non-B helices or melted DNA.

Candidate *trans*-factors operating through the CT-element are:

Sp1 The predominant duplex-binding protein interacting with the CT-element and a site further downstream. The essential role of Sp1 in regulating housekeeping genes is well-established (DesJardins and Hay 1993; Michelotti et al. 1996b). Acting locally on the *c-myc* promoters, Sp1 seems to be the conduit through which the immunoglobulin enhancer mediates the activation of P1 relative to P2 often occurring in Burkitt lymphoma (Geltinger et al. 1996).

hnRNP K A prototype for the KH-motif, bearing three repeats of this nucleic-acid binding module. Belying its name, this protein binds more tightly and sequence specifically with ssDNA than with RNA (Tomonaga and Levens 1995; Braddock et al. 2002a). Although there is some indication that hnRNP K recognizes duplex CT-elements, the structure of the ssCT-element complexed with hnRNP K provides no insight as to how this interaction might occur (Braddock et al. 2002a). hnRNP K interacts with TFIID as well as numerous other signaling and gene regulatory proteins. Its ability to stabilize single-stranded loops introduces torsional and flexural hinges into promoters, facilitating interactions between flanking sites (Takimoto et al. 1993; Tomonaga and Levens 1995; Geltinger et al. 1996; Michelotti et al. 1996a; Tomonaga et al. 1998). Between sculpting DNA and recruiting diverse partners, hnRNP K seems to be an adapter gating the interactions of other molecules with greater intrinsic transcription effector activity (Bomsztyk et al. 1997). Although hnRNP K is associated with increased *c-myc* expression, on other genes it may play a negative role. Whether hnRNP K plays a positive or negative role is likely to be context dependent in that it would be determined by the intrinsic activities of the more potent effectors it serves.

CNBP Cellular nucleic acid-binding protein. This multi-zinc-finger protein also binds avidly and in a sequence-specific manner with ssDNA and RNA (Rajavashisth et al. 1989; Michelotti et al. 1996b; Pellizzoni et al. 1997; Crosio et al. 2000). Evidence indicates that this protein may function in the translational regulation of some mRNAs. CNBP binds the purine-rich strand of the CT-element. Knockout of CNBP in mice diminishes *c-myc* expression in those zones of the embryo (forebrain most prominently) where CNBP is abundant; *c-myc* levels are unaffected in regions where CNBP is absent indicating either that other factors substitute for CNBP or that alternate mechanisms bypass the CT-element (Chen et al. 2003). Expressing CNBP in the CNBP$^{-/-}$ cells augments the expression of a transfected *c-myc*-reporter. Importantly, these results dramatize the differential utilization of transcription factors to regulate

c-myc expression in different cells. The mechanism of transcription activation by CNBP has not been elucidated.

nm23/NDPK Has been associated with a variety of enzymatic and regulatory activities. Initially identified as a transcript downregulated in metastatic cells, nm23 was shown to have tumor-suppressor activity. It was subsequently discovered in an expression screen designed to identify CT-element (also termed NHE—nuclease hypersensitive element) binding factors. This same protein has been associated with nucleoside diphosphate kinase activity, histidine-kinase activity, and both sequence-specific and generalized DNase activities (Hartsough and Steeg 2000; Postel et al. 2000; Roymans et al. 2002). Recently nm23 has surfaced as a subunit of an S-phase octamer-co-activating complex (OCA-S) (Zheng et al. 2003). The protein lacks intrinsic transcription activating function and may not possess sufficient DNA binding specificity to find its physiological targets in vivo unless complexed with partner proteins (Michelotti et al. 1997).

MAZ A multi-zinc-finger protein first identified binding with the *c-myc* promoter. It interacts at several sites within the vicinity of the promoter including the CT-element and a site further downstream inter-nested with Sp1 and E2F binding sites (Bossone et al. 1992, Sakatsume 1996).

YB-1/NSEP Identified as a component of a *c-myc* promoter binding ribonucleoprotein complex (the RNA component has not been characterized further). YB-1 has also been identified as a ssDNA-binding protein interacting with other promoters in vitro, with supporting evidence for a regulatory role for several genes in vivo (Davis et al. 1989; Kinniburgh 1989). YB-1-related proteins bear cold shock domains and have been reported to bind mRNAs and regulate translation in addition to recognizing DNA (Kloks et al. 2002). Recent studies indicate that signaling pathways potentially leading to specific proteolysis and cytoplasmic-nuclear shuttling might determine whether YB-1 binds to RNA or DNA.

The second region associated with altered structures and DNA conformations is a segment far upstream of P1 and P2 which possesses a peculiar sensitivity to torsional strain: the FUSE that binds FBP resides in an AT-rich segment that is easily melted by application of supercoiling forces (Michelotti et al. 1996b; He et al. 2000). This same region of DNA is hypersensitive to single-strand selective oxidation by potassium permanganate in vivo in cells expressing *c-myc*, but not in cells with silent *c-myc* genes. FUSE in nuclei of *c-myc* expressing cells is also sensitive to S1 nuclease. When *c-myc* is silent, a regular nucleosome array runs through the FUSE region (Michelotti et al. 1996b; Pullner et al. 1996; Albert et al. 1997; Albert et al. 2001). When *c-myc* is expressed, this array

is disturbed in the vicinity of FUSE. A BRG1-containing complex has been proposed to participate in the remodeling of FUSE chromatin. Just as Brahma-related gene 1 (BRG1) action has also been implicated in the generation or stabilization of a Z-DNA segment in the human colony-stimulating factor (CSF)-1 promoter, it is plausible that it may act similarly on *c-myc* (Chi 2003). Immediately upstream of the FUSE region is one of three Z-DNA-forming segments in the *c-myc* gene (Wittig et al. 1992; Wolfl et al. 1997). Antibodies recognizing Z-DNA can be cross-linked to this Z-DNA-forming region in nuclei. Just as negative supercoiling favors melting of FUSE, so conversion of right-handed B-DNA into left-handed Z-DNA is also driven by torsional stress. Melting at FUSE versus Z-DNA formation would compete to absorb torsional stress. Because nucleosomes do not accommodate non-B-DNA structures, both melting of the duplex and Z-DNA formation may contribute to the disturbance of the regular nucleosomal ladder in this region when *c-myc* is expressed. An origin of replication has been mapped to the FUSE/Z-DNA region. As occurs with other origins, nascent strand synthesis maps to a broader zone beyond the FUSE region (Tao et al. 2000; Liu et al. 2003). Functional coordination or direct mechanisms linking DNA synthesis and *c-myc* expression have not been explored.

2.5
FUSE-Binding Protein

FBP engages FUSE through four KH-motifs with each KH domain engaging 4–6 nucleotides (Braddock et al. 2002b). The cognate sequence segments are separated by spacer DNA (due to the intrinsic flexibility of ssDNA, there is no obligatory helical phasing between the segments engaged by each motif). FBP binds tightly with ssDNA and supercoiled DNA, but forms no stable complex with relaxed duplexes. The carboxyl terminus of FBP bears a tyrosine-rich motif that engages TFIIH to activate transcription (Tomonaga and Levens 1995; Duncan et al. 1996; Liu et al. 2001; Braddock et al. 2002a). FBP's activation domain stimulates the 3′-5′ helicase activity of the XPB/ERCC3/p89 subunit of TFIIH and facilitates initiation and advancement to promoter escape (see Sect. 2.8). The amino-terminus of FBP confers repressor activity when transferred to heterologous DNA binding domains (Duncan et al. 1996). FBP has two closely related sibs, FBP2 and FBP3. FBP2 and 3 bind to FUSE through four KH-motifs highly homologous to those in FBP and possess even more potent carboxyl terminal activation domains (Davis-Smyth et al. 1996). An adenovirus vector over-expressing FBP augments *c-myc* mRNA levels, whereas the same vector expressing a dominant-interfering FBP (central DNA binding domain only, devoid of amino and carboxyl terminal effector domains) de-

presses *c-myc* RNA (He et al. 2000). FBP itself is downregulated by the direct Myc target p38/JTV-1 (Kim et al. 2003a). Though a core protein in a multi-transfer (t)RNA aminoacyl-synthetase complex, knockout of p38 surprisingly does not impair protein synthesis (Kim et al. 2002). Rather, mice lacking p38 die in the immediate neonatal period with hyperplastic internal organs and increased *c-myc* levels. p38 targets FBP for ubiquitination and degradation. So, normally FBP augments Myc levels, Myc augments p38/JTV-1, and p38 downregulates FBP, closing a homeostatic feedback loop.

2.6 FBP-Interacting Repressor

The central DNA binding domain and the amino terminus of FBP bind FIR. FBP, FIR, and FUSE may form a ternary complex possessing both activation and repression moieties. The amino-terminus of FIR engages TFIIH and depresses, but does not abolish, the same XPB/ERCC3/p89 helicase activity augmented by FBP. FIR does not block initiation, but retards the advance of the transcription complex to promoter escape (see Sect. 2.8). *Drosophila* FIR (*puf60, hfp, dFIR*) was first reported to participate in the developmental regulation of alternative splicing (Van Buskirk and Schupbach 2002). More recently, *dFIR* was found to repress *Drosophila c-myc (dmyc)* at the RNA level, and dFIR was implicated in the regulation of cell cycle progression, influencing both G1/S and G2/M progression (Quinn et al. 2004).

2.7 Special AT-Rich Binding Protein 1

Immediately downstream of FUSE is an A–T rich segment. This segment is especially prone to melt at low levels of supercoiling that may nucleate the destabilization of FUSE. This same segment has the properties of a base-unwinding-region (BUR) and binds with special AT-rich binding protein 1 (SATB1), an atypical homeobox protein that nucleates higher order chromatin organization, especially chromatin loops (Cai et al. 2003, Dickinson et al. 1997; Yasui et al. 2002). SATB1 may contribute to cell-type-specific folding of *c-myc* chromatin or the partitioning of *c-myc* upstream sequences into subdomains. SATB1 recruits chromatin remodeling and modifying complexes. SATB1 does not bind to ssDNA, and so FBP and SATB1 actions are likely to prove mutually exclusive, if not antagonistic (Dickinson et al. 1992; Yasui et al. 2002).

Proteins binding “generic” Z-DNA exist, but Z-DNA-binding proteins that are also sequence-specific have not been described, so whether the upstream Z-DNA segment of *c-myc* plays physical roles such as excluding nucleosomes

or as a capacitor storing torsional energy, rather than serving as a platform to recruit special *trans*-factors, is not yet known (Rich and Zhang 2003).

2.8
A Scheme to Regulate *c-myc* Transcription

Where, when, and how much *c-myc* to transcribe must be explained in order to fully understand Myc biology. The cell uses several levels of molecular organization to answer these questions. First, it seems that special mechanisms park *c-myc* loci at particular intranuclear sites. Recent data reveal that *c-myc* genes are non-randomly distributed within nuclei; moreover, frequent-translocation partners with *c-myc*—such as the immunoglobulin heavy chain in Burkitt lymphoma—dwell closer on average to *c-myc* than do cytogenetically indifferent loci (Roix et al. 2003). Whether *c-myc* loci are deployed to stations conducive for proper expression, or whether this localization reflects a more passive partitioning of silent and expressed genes is not known. In terminally differentiated cells *c-myc* expression is irreversibly silenced, but cells retaining proliferative potential preserve the capacity to express *c-myc*. Most of these latter cells have a paused polymerase. The initial events in *c-myc* induction have not been defined. Binding of transcription factors to *c-myc* regulatory sequences and chromatin remodeling are likely to occur concomitantly and are probably interdependent. Depending on the variety and magnitude of signals, transcription factors of all sorts flicker on and off the *c-myc* gene. While some of these factors recruit chromatin remodeling and modifying complexes, the immediate issue for the induction of *c-myc* transcription is to restart the paused polymerase. Operating through TFIIH and perhaps other basal transcription components, the paused polymerase is spurred by activators through a series of otherwise slow transitions. The density of bound factors and the frequency of interactions between their activation domains and the promoter-bound apparatus control progression to the point of promoter escape. Only following escape would the promoter be available for reinitiation. In this scheme, a single intense signal acting repetitively through a responsive transcription factor, or multiple weak signals acting through diverse *cis*-elements, would ratchet the pre-promoter escape transcription complex through its various stages. Signal integration would occur through multiple sequential (but not necessarily ordered) activating events. Utilization of multiple, kinetically equivalent, pre-promoter escape intermediates would reduce the temporal variance and damp stochastic fluctuations when compared with a process regulated at a single rate-limiting step that yields only very small numbers of product. Following the delivery of an activating signal (note that experimentally this has often involved em-

ploying a single agent—few studies have dealt with synergy or antagonism between *c-myc*-regulating signals delivered in combination) *c-myc* transcription stereotypically peaks between 1 and 2 h after stimulation, and declines rapidly thereafter. Upon achieving log-phase, *c-myc* mRNA levels stabilize above resting levels but well below peak levels. Thus, once cells have experienced a pulse of *c-myc* transcription, a lower level sustains proliferation. Overlying the switches that upregulate *c-myc*, a molecular cruise control system may operate to constrain and prevent chaotic fluctuations of Myc levels. A scheme can be jury-rigged from the features of the FBP–FIR–FUSE system to superimpose dynamic, real-time feedback onto the *c-myc* promoter. Upon activation, transcription pumps torsional stress into the DNA upstream of the promoter. Loops between *trans*-factors and the translocating transcription apparatus at least transiently accumulate torsional energy (until either the loop breaks or a topoisomerase landing within the loop relieves all the tension). If the stress is focused mainly into the linker regions, transcribing even a short distance has the capacity, in principle, to drive structural transitions at sensitive sites within the loop. Hence the ability to recruit, hold, and functionally engage topology and/or conformation-sensitive factors would be linked to ongoing gene activity. The effector domains of some of these proteins (such as FBP and FIR) could reach back to the pre-promoter escape transcription complex and influence further transcription, or they might influence the rate of reinitiation on the next round of transcript synthesis. Such a mode of regulation would occur in real-time in response to ongoing RNA synthesis irrespective of the particular pathway activated to drive transcription. In contrast, conventional feedback requiring synthesis of a *c-myc* primary transcript, splicing, processing, mRNA transport to the cytoplasm, translation, dimerization, transport back to the nucleus, protein modification, incorporation into larger chromatin modifying and remodeling complexes, binding at target sites, and finally—if Myc autoregulation is direct via the P2 promoter—transcriptional repression of *c-myc*. For indirect autoregulation, a second cycle of expression would delay feedback repression even further. These delays would limit the ability of end-product feedback to impose tight homeostasis on a rapidly fluctuating system.

Chromatin immunoprecipitation studies have revealed temporal evolution of the spectrum of factors bound to the *c-myc* promoter during induction and shutoff (Shang et al. 2000). These shifting patterns either represent the superimposition of independent pathways activated and repressed with distinct kinetics (as occurs in yeast), or, alternatively, they represent the dynamic progression of a molecular machine through different stages of operation (Bryant and Ptashne 2003). Distinguishing between these alternatives is an experimental challenge with fundamental consequences for understanding