CHROMATIN STRUCTURE OF C-*MYC* IN HORMONE DEPENDENT AND INDEPENDENT HUMAN BREAST CANCER CELLS

By

Teresa Miller

A Thesis submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Department of Biochemistry and Molecular Biology

Faculty of Medicine

University of Manitoba

June 28, 1994

ISBN 0-315-92291-5

Name Teresa Lynn Miller

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Molecular Biology
SUBJECT TERM

0307
SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology ... 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology .. 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics ... 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

CHROMATIN STRUCTURE OF C-MYC IN HORMONE DEPENDENT AND INDEPENDENT HUMAN BREAST CANCER CELLS

BY

TERESA MILLER

A Thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Table of Contents

Abstract

The goal of this thesis was to characterize regulatory proteins interacting with the promoter regions of the c-*myc* gene in ER^+ and ER^- human breast cancer cell lines. This research has demonstrated differential sensitivity to the enzyme deoxyribonuclease I at the hypersensitive site associated with the P_0 promoter, (DH site II_2). This differential nuclease hypersensitivity is indicative of differences in the chromatin structure of the 5'-flanking region of the c-*myc* gene between ER^+ and ER^- human breast cancer cells. It was determined that this difference was not due to differential promoter usage, as the relative promoter usage was similar for estrogen responsive and nonresponsive cell lines. As well, no detectable difference in DNA methylation patterns in the c-*myc* 5'-regulatory region was observed between the cell lines MCF 7 and MDA MB 231, a hormone responsive and nonresponsive human breast cancer cell line, respectively. It was determined that the differential hypersensitivity was most likely due to differences in the nonhistone DNA binding proteins associated with DH site II_2. Characterization of these differences led to the identification *in vitro* of novel sites of DNA-protein interactions in the regions spanning the nuclease hypersensitive sites II_2 (-687 to -607) and III_2 (-1 to +66) of c-*myc*.

This study also provides evidence that Sp1-like proteins interact with ME1a1 and ME1a2 binding sites of c-*myc* DH site III_2 and confirms the presence of Sp1 binding sites in the c-*myc* DH II_2 region in human breast cancer cells. A model for the juxtapositioning of proteins associated with the DH II_2, III_1 and III_2 regions is proposed through the interaction of Sp1 multimers associated with these and other consensus sites throughout the regulatory regions of c-*myc*. The ER is suggested to be involved in this

interaction in hormonally dependent breast cancer cells while other factors are predicted to be involved in hormonally independent breast cancer cells. In conclusion, this thesis demonstrates that nuclear extracts of ER^+ MCF 7 and ER^- MDA MB 231 cells contain a similar, but not identical, spectrum of sequence specific DNA-binding proteins that interact with the DH II_2 and the DH III_2 regions of the human c-*myc* gene.

Research Acknowledgments

I would like to thank my supervisors Dr. Leigh C. Murphy and Dr. James R. Davie for the opportunity to pursue a PhD under their guidance, the wisdom that prevented me from quitting and the friendship we shared. I would also like to thank the other members of my PhD committee, Dr. Klaus Wrogemann and Dr. Janice G. Dodd for their advice and friendship throughout this degree. I would especially like to thank Dr. Dodd for her encouragement to continue graduate studies, her perspective on science and for the initial training I received in her lab as a MSc student. I would finally like to thank the members of my PhD committee and my external thesis examiner, Dr. Alain Nepveu, for reading and examining this thesis.

I would like to thank and acknowledge Dr. Yan Jin, Norman Huzel, Jian-Min Sun, Tim Salo and Helmut Dotzlaw for their technical assistance. The DNase I hypersensitivity studies were initiated by Tim Salo and completed by Norman Huzel and myself. Helmut Dotzlaw demonstrated and assisted with the RNase protection assays. Dr. Yan Jin performed the restriction enzyme methylation studies and the work on the c-*myc* DH III_2 region under my guidance and with my collaboration. Jian-Min Sun demonstrated and assisted with the DNase I footprinting analyses. I enjoyed the friendship we shared and wish you all continued success.

This research was supported by the National Cancer Institute of Canada, the Medical Research Council of Canada, the Cancer Research Society Inc. and the H. E. Sellers Trust Fund. I would like to thank the Manitoba Health Research Council and the University of Manitoba for studentship stipend support throughout my PhD.

Personal Acknowledgments

I would like to thank the following members of my family for all their encouragement, support and tolerance since I entered graduate studies; my husband Gary, my parents Don and Pat, my two brothers Darren and Wesley, my sister Lorraine and her husband Paul, and my grandmother Isabel. You were all important in my ability to continue on and without all of you I would have lost all perspective. I would also like to thank my friends Kathy and Ian Power for their friendship, ability to occasionally get me out of the lab and tolerance of my tardiness when they did. In addition, I would like to thank my husband's family for accepting and supporting my need to continue with my degree in spite of his transfer out of Winnipeg.

There are a number of friendships I have gained over the years as a PhD student and for these I am truly grateful. I would like to thank the following people for their continued friendship and support; Dr. Mike Hendzel, Amanda Coutts, Dr. Denise Belsham, Dr. Fred Pereira, and Dr. Barb Triggs-Raine. In particular, I would like to thank Mike Hendzel for the many hours of enlightening and stimulating conversation, even if it was at times distracting.

This thesis is dedicated to my husband Gary Dotzlaw.

Thank you for your support and encouragement through the four years of my PhD program. Although our dreams moved us apart for three of those years our love always held us together. Always remember that I love you not only for who you are, but for who I am because of you.

The force of the waves is in their perseverance.

--- Gila Guri

List of Figures

List of Abbreviations

Chemicals:

NaCl	Sodium chloride
NaAc	Sodium acetate
KCl	Potassium chloride
NH_4Ac	Ammonium acetate
$MgCl_2$	Magnesium chloride
DTT	Dithiothreitol
DMS	Dimethyl sulfate
DMSO	Dimethyl sulfoxide
GITC	Guanidinium isothiocyanate
$NaHCO_3$	Sodium Bicarbonate
NaOH	Sodium hydroxide
$MgSO_4$	Magnesium sulfate
NaH_2PO_4	Sodium mono-phosphate
$ZnCl_2$	Zinc chloride
$CaCl_2$	Calcium chloride
CsCl	Cesium chloride
EDTA	Ethylenediaminetetraacetic acid
SDS	Sodium dodecyl sulfate
BSA	Bovine serum albumin
Tris	Tris (hydroxymethyl) aminomethane
Pipes	1,4-Piperazinediethanesulfonic acid
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
Mops	Morpholinopropanesulfonic acid
Temed	N, N, N', N',-tetramethylenediamine
NP40	Nonidet P-40
PMSF	Phenylmethylsulfonyl fluoride
Mg^{2+}	Magnesium ions
Zn^{2+}	Zinc ions
Adr	Adriamycin (doxorubicin)

Units:

IU	International units
rpm	Revolutions per minute
cpm	Counts per minute
^oC	Degrees celsius
aa	Amino acids
bp, kbp	Base pairs, kilo base pairs
g, mg, µg, ng	Gram, milligram, microgram, nanogram
l, ml, µl	Liter, milliliter, microliter
M, mM, µM, nM	Molar, millimolar, micromolar, nanomolar
%	Percent
#	Number

Nucleotide Bases:

G	Guanine
A	Adenine
T	Thymine
C	Cytosine
U	Uridine
G+A	Guanine plus adenine
G+C	Guanine plus cytosine
A+T	Adenine plus thymine
dI-dC	Deoxy Inosine-deoxy cytosine polymers

Amino Acids:

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic Acid
C	Cysteine
Q	Glutamine
E	Glutamic Acid
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

Acronyms:

DNA, cDNA	Deoxyribonucleic acid, complementary DNA
RNA, mRNA, tRNA	Ribonucleic acid, messenger RNA, transfer RNA
NTP,dNTP	Nucleotide triphosphate, deoxy NTP
NDP	Nucleoside diphosphate
AMP	Adenosine monophosphate
DNase I	Deoxyribonuclease I
DH	DNase I hypersensitive
RNase	Ribonuclease
HRE	Hormone response element
ER, ER^+, ER^-, ERE	Estrogen receptor, ER positive, ER negative, estrogen response element
AR, ARE	Androgen receptor, androgen response element

D_3R, DRE	Vitamin D receptor, vitamin D response element
GR, GRE	Glucocorticoid receptor, glucocorticoid response element
MR, MRE	Mineralcorticoid receptor, mineralcorticoid response element
PR, PRE	Progesterone receptor, progesterone response element
RAR, RARE	Retinoic acid receptor, retinoic acid response element
T_3R, TRE	Thyroid hormone receptor, thyroid hormone response element
GnRH	Gonadotropin releasing hormone
CAT	Chloramphenicol acetyl transferase
MMTV	Mouse mammary tumor virus
LTR	Long terminal repeat
Wap	Whey acidic protein
PRA	Prolactin receptor associated protein
AdrR	Adriamycin resistant
MAR	Matrix associated region
SAR	Scaffold associated region
S/MAR	Scaffold or matrix associated region
LCR	Locus control region
TAF, TAF-1, TAF-2	*Trans*-activating function
TFIIA, TFIIB, TFIID	Transcription factor II A, II B, II D
TFIIE, TFIIF, TFII-I	Transcription factor II E, II F, II-I
TFIIIA	Transcription factor III A
TATA	TATA box
TBP	TATA box binding protein
INR	Initiator element
ARS	Autonomous replicating sequence
HBD	Hormone binding domain
NLS	Nuclear localization signal
DBD	DNA binding domain
BR	Basic region
HLH	Helix-loop-helix
LZ	Leucine zipper
USE	Upstream element
FUSE	Far upstream element
CBP	CAAT binding protein
Rb	Retinoblastoma protein
RCE	Rb control element
H1, H2A, H2B, H3, H4	Histones 1, 2A, 2B, 3 and 4
MBP-1	Myc binding protein 1
Oct-1	Octamer binding protein 1
POU	Pit-1, Octamer, Unc 86
POU_S	POU specific
POU_{HD}	POU homeodomain
Ig, IgL, IgH	Immunoglobulin, Ig light chain, Ig heavy chain
CK-II	Casein kinase II

BL	Burkitt's lymphoma
REF	Rat embryo fibroblast
TNM	Tumor size, node involvement, metastisis
EMSA	Electrophoretic mobility shift assay
NPB	Nuclei preparation buffer
CLB	Cell lysis buffer
DMEM	Dulbecco's minimal essential medium
PRF	Phenol red free
FBS	Fetal bovine serum
CS-FBS	Charcoal stripped FBS
SSC	Standard sodium citrate
SSPE	Standard saline phosphate EDTA
MW	Molecular weight/mass
x g	Times gravity
v/v	Volume per volume
w/v	Weight per volume
α	Alpha
ß	Beta
γ	Gamma
κ	Kappa
λ	Lambda
μ	Mu
^{32}P	Phosphorous 32

Introduction

Development and Pathology of the Breast

The physiological function of the mammary gland is to produce milk proteins which are secreted during lactation. The development of mammary tissue is regulated in part by the presence of estrogen and progesterone. Estrogen controls the growth and branching of the ductal system in cooperation with at least growth hormone, prolactin, glucocorticoid and insulin. Similarly, progesterone works synergistically with these hormones to regulate lobule growth, alveoli budding and development of the secretory characteristics of the alveoli. The mammary gland is therefore an endocrine target tissue which responds to a variety of hormonal stimuli. Pathological conditions of the breast, principally breast cancer, are therefore potentially responsive to the same hormonal signals. Significantly, assays for measuring estrogen and progesterone receptors are used to classify breast tumors. Additionally, the anti-estrogen, tamoxifen, is used therapeutically as a post-operative treatment for breast cancer and as an experimental chemopreventive measure for women at high risk for this disease. The development of breast tumors which are resistant to tamoxifen is one disadvantage of its chemotherapeutic use. A better understanding of the molecular actions of estrogen may eventually lead to an understanding of how to prevent the development of estrogen independent cells and therefore limit treatment failure.

Normal Breast Development

The embryonic mammary gland arises from two dense bands of ectodermal cells which thicken by 6 weeks post-conception into the primary mammary bud (reviewed in Topper and Freeman, 1980). This structure

further differentiates under the influence of the surrounding mesenchyme into secondary buds which later form a characteristic branching pattern of ducts. At this point in development male and female mammary gland morphology is histologically and functionally identical. Destruction of the ovaries does not block this development, suggesting mammogenesis during embryonic life may be independent of ovarian hormones (Raynaud, 1950). During the third trimester, canalization of the lactiferous ducts towards the apex of the nipple occurs in the female. Inhibition of this process in the male coincides with the emergence of androgens from the fetal testes (Drews and Drews, 1977). Destruction of the fetal testes allows the female mammary gland phenotype to prevail (Raynaud, 1950).

The cells of the fetal mammary gland consist of an inner secretory epithelium and an outer layer of myoepithelial cells. The secretory cells are already determined as mammary cells and can respond *in vitro* to a combination of insulin, glucocorticoid and prolactin to synthesize milk proteins (Ceriani, 1970). Similarly, fetal prolactin and placental steroids stimulate partial differentiation and limited secretion from female ductal secretory cells for a period just prior to parturition and continuing for approximately four weeks after birth (Friesen, 1973). After this time period, the newborn mammary gland reverts from this partially differentiated state to an undifferentiated state until the onset of puberty.

With the onset of sexual maturation mammary gland development primarily involves duct formation. Additionally, the deposition of adipose tissue and the development of stromal connective tissue also ensues. The production of a determined ductal architecture is dependent on both estrogen and the presence of growth hormone or prolactin. Conversely, the pituitary hormones are not necessary for ductal maintenance providing ovarian or

adrenal secretion is present. The exogenous addition of estrogen to ovariectomized mice stimulates the growth of ductal end buds (Bresciani, 1968). However, estrogen is unable to stimulate ductal growth when ovariectomy is combined with hypophysectomy and adrenalectomy (Lieberman *et al.*, 1978; Lyons, 1958). In the dwarf mouse, thyroid hormones are able to substitute for the pituitary hormones (Pissott and Nandi, 1961). Ductal growth appears to be independent of both progesterone and insulin (Topper and Freeman, 1980). Additionally, extensive growth occurs in the absence of glucocorticoids even though this hormone is required for maximal growth (Richardson, 1955).

Although the nonlactating breast undergoes some cyclical development during each follicular and luteal phase of the menstrual cycle (Bresciani, 1965), the final stages of mammary gland maturation occur during pregnancy. The ductal end buds differentiate into lobuloalveolar structures, consisting of clusters of spherically organized epithelial cells, which fill the interductal spaces (Kaplan and Schenken, 1990). At this stage, estrogen and progesterone are both required for mammary gland development (Topper and Freeman, 1980). Estrogen is involved in stimulating the general growth response of the gland, while progesterone primarily stimulates lobuloalveolar development. The presence of growth hormone, prolactin or placental lactogen appears essential to the maturation and maintenance of these structures as well (Nandi, 1958; Talwalker and Meites, 1961; Topper, 1970; Franks *et al.*, 1977; Neilson *et al.*, 1979). Ovariectomy or hypophysectomy causes a complete loss of alveoli, while adrenalectomy has only a slight effect (Nandi, 1958). Thyroid hormones have been shown to be stimulatory for alveolar development (Vonderhaar and Greco, 1979). Similarly, glucocorticoids, although not essential, enhance

lobuloalveolar development (Nandi, 1958). Conversely, insulin is not required for ductal or alveolar growth, but may be necessary for the manifestation of the ultimate mammary phenotype (Topper and Freeman, 1980). A cyclic acquisition and loss of insulin responsiveness during development has been reported (Oka *et al.*, 1974). During pregnancy the mammary cells become sensitive to insulin and retain this phenotype to the end of lactation. Upon cessation of lactation the mammary gland undergoes involution and these cells revert to an insulin insensitive state. It has been suggested that progesterone is responsible for sensitizing cells to insulin (Topper and Freeman, 1980). In addition to the mentioned hormones, the insulin-like growth factors I and II, epidermal growth factor and transforming growth factor alpha are all stimulatory for lobuloalveolar growth (Tonelli and Soroff, 1980; Schreiber *et al.*, 1986).

In the later stages of pregnancy the mammary cells mature to their full secretory potential. While progesterone plays a role in this maturation process, this steroid simultaneously blocks secretory activity (Topper and Freeman, 1980). The initial induction of milk protein secretion is primarily regulated by placental lactogen and occurs just prior to parturition (Egli *et al.*, 1961). Estrogen is believed to promote the initiation of lactation but is nonessential for its maintenance (Topper and Freeman, 1980). Declining levels of progesterone coupled with increased levels of prolactin and glucocorticoid are responsible for maintaining postpartum lactation (Lyons, 1958; Topper and Freeman, 1980). Although estrogen promotes synthesis of prolactin and its receptor, levels of this steroid decline after parturition. This is likely due to an inhibitory effect of estrogen on postpartum lactation. Conversely, increases in oxytocin, which also stimulate prolactin release (Lumpkin *et al.*, 1983), are required for myoepithelial contraction and

expulsion of milk from the breast. Continuous emptying of the alveoli maintains lactation by stimulating both prolactin and oxytocin release from the pituitary (Peaker, 1980).

A large body of research therefore suggests that the development and maturation of the mammary gland is coordinately regulated by a number of different hormones. Although the development of the embryonic mammary gland appears to be independent of ovarian function (Raynaud, 1950), the growth and development of the ductal architecture during sexual maturation and pregnancy requires ovarian hormones (Nandi, 1958; Bresciani, 1968; Topper and Freeman, 1980). In particular, estrogens have an essential role in the initiation and maintenance of this process which continues up to the initiation of lactation (Topper and Freeman, 1980). This steroid hormone is therefore necessary to the normal development of mammary tissue and is potentially involved in the development of mammary neoplasia.

Breast Cancer

Breast cancer is the most common cancer in women of western society. Accurate records of breast cancer cases in Connecticut since 1940 indicate that the incidence of this disease has increased steadily by about 1% a year (Marshall, 1993a). In 1988 alone, 4,480 women and 33 men died from breast cancer in Canada (NCIC, 1990). Even more striking, in the United States 180,000 women are diagnosed and 46,000 succumb to breast cancer each year (Marshall, 1993a). In total, the World Health Organization has estimated that breast cancer causes the death of a quarter of a million women each year throughout the world (Logan, 1975). For North American women, the lifetime risk for developing this malignancy used to be one in twelve but more recently has been adjusted to one in nine (Volkers, 1992). This statistic is

based on the combined probabilities of a woman born in 1992 living to the age of 85 and developing breast cancer between birth and age 85.

Increased use of mammography initiated in the 1980s has led to earlier and better detection of breast cancer, which was reflected in larger yearly increases (up to 4% by 1987) in breast cancer incidence (Marshall, 1993a). The detection of stage I breast tumors (< 2 cm) and those that have not yet invaded the surrounding breast tissue (*in situ*) has increased, while the incidence of more advanced cancers has declined or remained constant since 1982 (Miller *et al.*, 1991). This may be expected to promote a decline in the breast cancer mortality rate for those cancers detected early. However, deaths from breast cancer have so far remained steady at 32-40 per 100,000 women since 1931 (NCIC, 1990).

Pathophysiology of Breast Cancer

The phenotypic characteristics shared by all cancer cells include unrestricted growth, the ability to invade across basement membranes and metastasize to distant sites, and adaptability at the genome level to environmental changes (Lippman, 1993). Some of these characteristics are used to predict disease outcome. In particular, tumor size, lymph node involvement plus presence and site of metastasis (TNM grading) are the traditional prognostic variables for breast cancer. The additional measurement of estrogen and progesterone receptors (ER and PR, respectively) for prognostic and treatment response information was initiated in the 1970's. It has been clearly demonstrated that patients having breast tumors which express these hormone receptors enjoy a better overall survival rate which is independent of their exposure to anti-hormone therapy (Blanco *et al.*, 1984).

The primary cause of breast lesions are hyperplastic changes in the lobular and ductal epithelial cells. Most of these lesions are not neoplastic and represent common physiological abnormalities resulting from inflammation or fibrocystic alterations within the breast. Conversely, atypical hyperplasia is a marker associated with an increased risk for developing neoplasia (Vorherr, 1980). This preneoplastic condition is characterized by pleiomorphism, nuclear enlargement and chromosomal abnormalities, as well as increased and homogeneous expression of the ER. The presence of this receptor in preneoplastic mammary disease has suggested that it is involved in the promotion of breast neoplasia (Nenci *et al.*, 1988). Only about 10% of all palpable breast lesions are diagnosed as breast cancers. Of these, 75-80% are carcinomas originating from glandular epithelial cells of ductal origin and 10-20% are carcinomas arising from epithelial cells of lobular origin (Ramzy, 1990). Neoplastic cells, which have not invaded the basement membrane of the duct or lobule ,are considered to be ductal or lobular carcinomas *in situ.* Conversely, those breast tumors, which have penetrated the surrounding tissue, are known as invasive or infiltrating ductal or lobular carcinomas.

Risk Factors Associated with Breast Cancer

The factors that have been clearly linked with an increased risk in the development of breast cancer are family history, reproductive events and certain subtypes of benign breast disease (reviewed in Marshall, 1993b). Although a number of other associations that have been characterized suggest some external influence in promoting breast cancer, the factors involved have not been definitively identified. In particular, women born in North America or Northern Europe and those in a higher socioeconomic class experience an increased relative risk of developing breast cancer by a factor

of 2.0-4.0 (Figure 1). The geographical coincidence of a high fat diet and high cancer rates seemed to provide circumstantial evidence for a dietary role in breast cancer risk. This was supported by the increased incidence of breast cancer in subsequent generations of women who moved from low risk to high risk countries (McMichael and Giles, 1988). However, the epidemiological data did not support a relationship between fatty diets and breast cancer, although the evidence did support this relationship with colon cancer (Willett, 1989). Conversely, daily alcohol consumption was found to be correlated with an increased risk for breast cancer.

For a minority of breast cancer cases a family history of the disease suggests a genetic predisposition (Anderson, 1974). A family history of premenopausal bilateral breast cancer increases a woman's relative risk by a factor of 4.0 (Figure 1), while a woman having a first-degree relative with breast cancer has a relative risk increase of 2.0-4.0 fold (Marshall, 1993b). Additionally, the probability of breast cancer recurrence after an extended disease-free interval is very high (Brinkley and Haybittle, 1975). A family history of fibrocystic breast disease or history of primary cancer in the ovary or endometrium similarly increases the relative risk by a factor of 2.0-4.0.

Weak linkage to the estrogen receptor gene on chromosome 6 has been observed in a late-onset breast cancer family (Zuppan *et al.*, 1991). Similarly, evidence now exists that one woman in 200 inherits a defective gene found on chromosome 17q that results in an 80-90% chance of developing breast cancer (Hall *et al.*, 1990; Narod *et al.*, 1991; Roberts, 1993). This defect is believed to be responsible for about 5% of all breast cancers, specifically the rare inherited form which strikes women in their thirties or forties. Linkage of this chromosomal region has also been demonstrated for an inherited form of ovarian cancer (Narod *et al.*, 1991; Black and Solomon, 1993; Roberts, 1993),

Risk Factors for Breast Cancer in Females

Factor	High Risk	Low Risk	Magnitude of Differential
Age	Old	Young	●●●
Country of birth	North America Northern Europe	Asia, Africa	●●
Socioeconomic class	Upper	Lower	●●
Marital status	Never Married	Ever Married	●
Place of residence	Urban	Rural	●
Place of residence	Northern U.S.	Southern U.S.	●
Race	White	Black	●
Age at first full term pregnancy	Older than 30	Younger than 20	●●
Ovariectomy	No	Yes	●●
Body build, postmenopausal	Obese	Thin	●●
Age at menarche	Early	Late	●
Age at menopause	Late	Early	●
Family history of premenopausal bilateral breast cancer	Yes	No	●●●
History of cancer in one breast	Yes	No	●●●
History of fibrocystic disease	Yes	No	●●
Any first-degree relative with breast cancer	Yes	No	●●
History of primary cancer in ovary or endometrium	Yes	No	●●
Radiation to chest	Large doses	Minimal exposure	●●

Key to magnitude of risk differential ●●● = relative risk of greater than 4.0; ●● = relative risk of between 2.0-4.0; ● = relative risk of between 1.1 and 1.9.

Figure 1. Risk Factors Associated with Breast Cancer. A variety of factors associated with the development of breast cancer in females and their associated relative risk are compared. The key to the magnitude of risk differential between the low and high risk groups for each factor is given at the bottom of the figure. This figure was reproduced from Marshall, 1993.

while a similar chromosomal region appears to be involved in the development of sporadic forms of breast and ovarian cancers (Futreal *et al.*, 1992; Black and Solomon, 1993; Jacobs *et al.*, 1993).

Almost a century has passed since ovariectomy was first demonstrated to cause a temporary regression of breast cancer (Beatson, 1896). Significantly, women without functional ovaries and males rarely develop breast cancer (Vorherr, 1980; Thomas, 1986). A role for ovarian function in the development of breast cancer is supported by numerous other observations. In particular, reproductive factors, which determine a woman's exposure to the steroid hormones estrogen and progesterone, contribute significantly to the risk of developing breast cancer (Marshall, 1993b). Early age at menarche and late onset of menopause independently increase the relative risk of developing breast cancer by a factor of 1.1-1.9 (Figure 1). Additionally, the protective effect of early age at first full-term pregnancy results in a reduced relative breast cancer risk of 2.0-4.0 fold. This is believed to be partly due to the early cellular differentiation of breast tissue and to the temporary cessation of the ovulatory cycle during pregnancy, both of which suspend the proliferative effects of estrogen on breast epithelium (Cairns, 1975; Preston-Marin *et al.*, 1990).

The use of steroid hormones for contraception and hormonal replacement therapy may also increase the risk of breast cancer. In general, no increase in risk has been associated with oral contraceptive use, however an increased risk of premenopausal breast cancer from long term (10 years) use prior to first full-term pregnancy has been reported (Miller *et al.*, 1989a; Romieu *et al.*, 1990). This increased risk may reflect a cohort effect since women with a long duration of contraceptive use are only now entering their postmenopausal years (Romieu *et al.*, 1990). Significantly, the simultaneous

presence of estrogen and progestin in the combined oral contraceptive markedly decreased the risk of endometrial and ovarian cancer (Henderson *et al.*, 1993). Similarly, although estrogen-replacement therapy has been associated with an increased risk in endometrial cancer, hormone-replacement formulations which include progestins negate this effect. Most studies suggest that postmenopausal estrogen-replacement therapy does not impose an increased risk for breast cancer (Armstrong, 1988). Use of this hormone over several decades has been reported to cause only a modest increase in risk (Brinton *et al.*, 1986). However, very few studies have focused on the effects of combined hormone-replacement therapy. The risk associated with this therapy appears to be higher than estrogen alone (Persson *et al.*, 1992) supporting the "estrogen augmented by progesterone" hypothesis regarding breast cancer etiology (Key and Pike, 1988; Henderson *et al.*, 1993).

Treatment of Breast Cancer

Historically, treatment of breast cancer occurred hierarchically, with surgical modalities ranking above clinical intervention (reviewed in Bonadonna and Valagussa, 1988). Surgical removal of the tumor and surrounding tissue was the first stage of defense but was used only when the disease was considered limited to local-regional areas. Those patients with technically unresectable breast cancer were referred to radiation therapy, while those neoplasms thought to be untreatable by radiation were referred to the clinical oncologist for treatment with growth inhibiting chemicals. In 1966 it was determined that the regional lymph nodes did not constitute an effective barrier to tumor cell dissemination (Fisher and Fisher, 1966). This led to the concept of a biological approach to cancer treatment which evolved into the multidisciplinary approach we see in use today.

The use of radical mastectomy has been replaced with lumpectomy for easily resectable tumors. As well, these surgical treatments have been combined with various post-operative therapies which have been analyzed by clinical trials (reviewed in Bonadonna and Valagussa, 1988). The benefits and optimal duration of combination or single agent chemotherapy has been studied for almost two decades. Particularly the treatment performance of the concurrent use of cyclophosphamide, methotrexate and fluorouracil (CMF) has been compared to the individual action of each of these and other chemicals and to CMF treatment in combination with prednisone, vincristine and/or adriamycin. Additionally, the use of postoperative irradiation, immunotherapy or endocrine therapy to supplement adjuvant chemotherapy has been considered.

Temporary regression of breast cancer through ovariectomy was demonstrated almost a century ago (Beatson, 1896), although the use of ovariectomy as an adjuvant therapy to mastectomy was suggested by Schnizinger seven years prior to this (Bonadonna and Valagussa, 1988). Even though ovariectomy was used as an effective treatment for breast cancer, it wasn't until the 1970's that hormones began to receive renewed attention as potential adjuvant treatments (Bonadonna and Valagussa, 1988). Since its introduction, the antiestrogen, tamoxifen, has been used extensively to treat advanced breast cancer and as an adjuvant treatment for early stages of the disease. This estrogen antagonist is a nonsteroidal aminoether derivative of polycyclic phenols which exerts its cytostatic effect principally by interacting with the ER (Furr and Jordan, 1984; Epstein, 1988). It is the only antiestrogen which has been shown through clinical trials to be at least as effective as surgical methods of hormone ablation (Furr and Jordan, 1984). The benefits of such therapy include increased over-all

and disease-free survival in addition to a decreased incidence of contralateral breast cancer (Nayfield *et al.*, 1991). These beneficial effects were most apparent in postmenopausal women with ER positive tumors, but have also been seen in premenopausal patients and those with ER negative breast cancer. Other hormonal manipulations used to treat breast cancer include gonadotropin-releasing hormone (GnRH) agonists, aromatase inhibitors and progestins (Epstein, 1988). The use of GnRH agonists has been shown to simulate the effects of hypophysectomy and induce remissions as efficiently as ovariectomy (Harvey *et al.*, 1985). In addition, the controversial antiprogestin, RU 486, has recently entered limited clinical trials in Canada to determine its efficacy in the treatment of metastatic and inoperable local breast cancer (Jenks, 1992).

Although the risk of breast cancer recurrence decreases over time, the possibility of treatment failure continues beyond even 15 years. This unsettling fact therefore defines breast cancer "cure" as the disease-free survival time (Bonadonna and Valagussa, 1988). Substantial evidence implicates the emergence of drug-resistant tumor cells as one of the major factors responsible for the clinical failure of breast cancer (Goldie and Coldman, 1984). This phenomenon limits the use of current therapy protocols, such as treatment with the nonsteroidal antiestrogen, tamoxifen. Although approximately 70% of breast tumors are classified as ER positive only half of these will actually respond to tamoxifen or other forms of endocrine therapies (McGuire, 1980; Arafah and Pearson, 1986) and are therefore hormonally dependent. The regression of primary breast cancer through hormonal manipulation remains a significant advance in the clinical management of the disease, in spite of the possible progression of the tumor to a hormonally non-responsive state (Epstein, 1988). However, a major

problem in breast cancer biology concerns the changes in breast cancer cells that result in their progression from hormonal dependence to hormonal independence.

Prevention of Breast Cancer

A number of preventative interventions have recently been initiated in the United States and Canada (reviewed in Henderson, 1993). The Women's Health Initiative has a number of investigative goals. These include an analysis of low dietary fat and high fruits and vegetables on the incidence of breast cancer and the effect of combined (estrogen plus progestin) hormone-replacement therapy on the development of breast and endometrial cancers, prevention of coronary heart disease and prevention of osteoporosis. The Canadian National Breast Cancer Screening Study is aimed at comparing the efficacy of annual mammography with clinical breast examination. Similarly, a joint study between China and the United States is aimed at assessing the impact of monthly breast self-examination on breast cancer mortality, since mammography is essentially unavailable in China.

Until recently, the only preventative actions for women at high risk of developing breast cancer were prophylactic bilateral mastectomy or early detection programs (Nayfield *et al.*, 1991). The development of various chemopreventive treatments has focused on the interruption of estrogen action. The use of a peptide to block ovarian production of estrogen in women with a high risk of developing the disease is presently in clinical trials (Spicer *et al.*, 1991). This GnRH agonist is predicted to afford a similar protective effect for ovarian and endometrial cancer as combined oral contraceptives, while reducing the lifelong breast cancer risk by 31%, 47% and 70% if used for 5, 10 and 15 years, respectively (Henderson *et al.*, 1993). Similarly, treatment with the nonsteroidal antiestrogen, tamoxifen, which binds to and

blocks ER action, is also in clinical trials (Nayfield *et al.*, 1991; Henderson *et al.*, 1993; Henderson, 1993). The major reason for determining the efficacy of this antiestrogen as a chemopreventive agent was the observation that women with primary breast cancer that have been treated with adjuvant tamoxifen therapy have a significantly decreased incidence of contralateral breast cancer (Henderson *et al.*, 1993). The concern of using tamoxifen as a chemopreventive agent has been the reported increased risk of endometrial cancer similar to that seen with postmenopausal estrogen-replacement therapy (Nayfield *et al.*, 1991).

The identification of additional factors, which contribute to breast pathogenesis, will likely promote increased research and development into other chemopreventive therapies. The significant breast cancer protection attributed to an early first full-term pregnancy has lead to the suggestion that young women should be encouraged to bear children in their late teens and early twenties (Henderson, 1993). It is more plausible that our knowledge of the hormonal control of breast differentiation during pregnancy may provide the rationale for designing other chemopreventive strategies. Additionally, mechanisms to postpone early menarche is another avenue for breast cancer prevention which could be explored.

The Role of Estrogen in Breast Cancer

Estradiol is a potent mitogen for human breast cancer cells (Darbre *et al.*, 1983). Even though only 7% of normal breast epithelial cells express the ER, 70% of diagnosed breast cancers are classified as ER positive (Isotalo *et al.*, 1983). Significantly, the hormonally dependent phenotype correlates strongly with the presence of ER in the tumor (McGuire *et al.*, 1975). In addition, the presence of the PR is a marker for the expression of a normal, functional ER (Horwitz and McGuire, 1978; McGuire, 1980). The use of ER

and PR as markers for tumor differentiation have therefore become important for predicting patient prognosis and response to hormonal treatment (McGuire *et al.*, 1975; Blanco *et al.*, 1984; Barbi *et al.*, 1987; Bonadonna and Valagussa, 1988).

Radioreceptor and immunoassays are primarily used to detect the presence of ER and PR in breast tumors (Holmes *et al.*, 1990). Both of these methods rely on the integrity of the carboxy-terminal region, particularly the steroid binding domain. Breast tumors containing receptors, which are defective for hormone interaction, would be detected as ER negative (ER^-) even though the receptor may actually be present and constitutively active in the absence of hormone. Similarly, receptors, which are structurally or functionally abnormal in other areas, such as the DNA-binding or *trans*-activation regions, would result in an ER positive (ER^+) classification. The significance of accurate receptor classification is exemplified by the variability of tumor response to endocrine therapy. Sixty percent of tumors characterized as ER^+ by hormone binding assays respond to endocrine treatment, while an additional 10% classified as ER^- by the same assay will show a similar response (McGuire *et al.*, 1975). Furthermore, 80% of those tumors, which are both ER^+ and PR^+, will respond to anti-hormone therapy, while breast cancers, which contain ER but not PR, respond in only 40% of the cases (McGuire, 1980).

The discovery that at least 20% of ER^+/PR^+ tumors and 60% of ER^+/PR^- tumors fail to respond to anti-estrogen treatment suggests that these breast tumors may contain defective ER. One hypothesis explaining the failure of endocrine therapies on ER positive breast cancers is therefore the presence of abnormal forms of the receptor. The detection of variant forms of the ER mRNA and protein have been described for a number of

breast tumors and cell lines (Murphy and Dotzlaw, 1989a; Fuqua *et al.*, 1991; Scott *et al.*, 1991; Foster *et al.*, 1991). In particular, the loss or truncation of the DNA-binding and steroid-binding domains has been detected. Although particular alterations in the ER *trans*-activating domains have not been detected, the classification of some tumors as positive for hormone and DNA binding but negative for the expression of the PR suggest defects in the transcriptional activation function of the ER or the ability of the PR gene to respond to ER regulation (Foster *et al.*, 1991). Additionally, it has been proposed that the use of DNA binding assays could aid in the prognostic value of steroid receptor classifications of breast tumors (Scott *et al.*, 1991; Foster *et al.*, 1991).

Research therefore supports a role for estrogen in regulating the proliferation of both normal breast tissue (Topper and Freeman, 1980) and human breast cancer cells (Darbre *et al.*, 1983). Significantly, ovarian function is a major factor in predicting breast cancer risk (Marshall, 1993b). The molecular mechanisms responsible for determining this hormonal dependence are complex and in most cases unresolved (McGuire, 1975). However, some of the most effective treatments for breast cancer exploit the cellular dependence of the mammary gland on estrogen. In particular, the use of the anti-estrogen tamoxifen primarily interferes with the action of the ER and is as effective as hypophysectomy and ovariectomy in blocking estrogen action (Jordan and Furr, 1984). The development of resistance to endocrine therapy limits the use of these treatments (McGuire, 1980) and promotes the development of estrogen independent breast tumors (Goldie and Coldman, 1984). An understanding of the pathways by which estrogens and antiestrogens regulate the proliferation of hormonally dependent breast cancer is therefore an essential prerequisite to an understanding of how

breast cancers develop resistance to endocrine therapies and eventually become hormonally independent.

Models for Breast Cancer

Breast cancer cell lines have been established from both solid tumors and pleural effusions of patients with the disease, with the latter yielding the majority of the cell lines isolated (reviewed in Engel and Young, 1978). The first breast cancer cell line to be established, BT 20, was isolated from a solid tumor in 1958 (Lasfargues and Ozzello, 1958). Subsequently, many other cell lines have been established, including the cell lines ZR 75-1 (Engel *et al.*, 1978), MCF 7 (Soule *et al.*, 1973), MDA MB 231 (Cailleau *et al.*, 1974) and T47 D (Keydar *et al.*, 1979), all of which were isolated from pleural effusions. Additionally, an immortalized "normal" breast cell line, HBL 100, has been isolated from breast milk (Gaffney, 1982). All of the above breast cancer cell lines were isolated from infiltrating ductal carcinomas except MDA MB 231, which was classified only as an adenocarcinoma.

The breast cancer cell lines ZR 75-1, T47 D and MCF 7 have been characterized as ER, PR and glucocorticoid receptor (GR) positive, while the BT 20 and MDA MB 231 breast cancer cell lines were only GR positive. The ER positive breast cancer cell lines ZR 75-1, T47 D and MCF 7 proliferate in response to estrogen (Lippman *et al.*, 1976; Chalbos *et al.*, 1982), while progestins exert an antiproliferative effect (Murphy and Dotzlaw, 1989b; Vignon *et al.*, 1983). These cell lines and the ER negative MDA MB 231 cells are also growth inhibited by glucocorticoids. The expression of α-lactalbumin, considered by some to be a definitive marker for mammary cell origin, has been conclusively demonstrated for the MCF 7, BT 20 and MDA MB 231 cell lines (Engel and Young, 1978).

Some additional cell lines have been subcultured from the aforementioned breast cancer cell lines. These include the T47 D5 cell line, which was initially described as a faster growing subpopulation of T47 D cells (Reddel *et al.*, 1988), and a progestin resistant cell line T47 D5 RP established from the T47 D5 cell line by long term culture in the presence of the synthetic progestin, medroxyprogesterone acetate (Murphy *et al.*, 1991). These three cell lines display differences in their response to progestins and in their levels of ER and PR. Steroid receptor positive sublines of MDA MB 231 cells have been isolated and their parental origin has been confirmed by analysis of hypervariable minisatellite sequences (Crépin *et al.*, 1990). Additionally, the transfection of MDA MB 231 cells with an ER expression vector has resulted in the S30 cell line, which although regaining its hormonal responsiveness, is growth inhibited by estrogen (Jiang and Jordan, 1992).

The long term treatment of MCF 7 cells with adriamycin (doxorubicin), which is an antibiotic with antineoplastic properties, has led to the establishment of a multidrug resistant cell line (MCF 7/AdrR) which no longer expresses the ER (Vickers *et al.*, 1988). Additionally, the long term growth of MCF 7 cells in estrogen depleted conditions either *in vitro* (Katzenellenbogen *et al.*, 1987; Clarke *et al.*, 1989) or *in vivo* by passage through ovariectomized athymic nude mice (Clarke *et al.*, 1989; Brünner *et al.*, 1993) has produced the MCF 7 sublines MIII, BSK-2 and BSK-3. These cell lines were able to proliferate in the absence of estrogen both *in vitro* and *in vivo*, but remained responsive to estrogen as measured by a decreased tumor doubling time.

Chronic infection of murine mammary tissue by the mouse mammary tumor virus (MMTV) has been associated with the spontaneous induction of

mammary tumors (Callahan, 1989; Callahan and Campbell, 1989). Mammary tumors are induced between 4 and 9 months after viral infection by the insertional mutation and subsequent inappropriate activation of silent cellular genes. Four unrelated genes (*int-1*, *int-2*, *int-3*, *int-4*), which may play a role in mammary carcinogenesis, have been detected near the MMTV integration site. The enforced expression of the *int-1* gene linked to the MMTV long terminal repeat (LTR) has been shown to induce mammary hyperplasia and subsequent adenocarcinoma (Tsukamoto *et al.*, 1988). Similarly, expression of *int-1* and *int-2* has been observed in some human breast tumors (Ali *et al.*, 1989).

The MMTV promoter and LTR enhancer have also been used to enforce tissue-specific expression of the c-*myc* proto-oncogene (Stewart *et al.*, 1984). Mice carrying this transgene show no obvious alterations in mammary gland development, however the deregulated expression of the c-*myc* gene results in pregnancy induced mammary-specific adenocarcinoma. Similarly, the 5'-regulatory region of the whey acidic protein (*Wap*) gene has been used to direct mammary-specific c-*myc* expression (Schöenenberger *et al.*, 1988). An 80% incidence of mammary adenocarcinomas is observed in transgenic animals expressing the *Wap-myc* fusion gene. Significantly, the expression of this transgene also results in the hormonally independent expression of the endogenous *Wap* and *ß casein* milk protein genes, which persists even when the tumors are transplanted into nude mice.

Other models for studying breast cancer include the use of the athymic nude mouse and culturing of cell lines on extracellular matrices. The passage of breast cancer cell lines in nude mice has allowed researchers to study the metastatic potential of each cell line, the factors involved and the related biology of the host animal. Studies such as these have led to the

identification of *nm23*, expression of which blocks cellular migration and prevents metastasis (reviewed in Marx, 1993). This protein, which is a nucleoside diphosphate kinase, has recently been shown to be related to the c-*myc* DNA-binding protein PuF (Postel *et al.*, 1993). The culturing of cells on extracellular matrices has similarly aided studies on the three-dimensional architecture of breast tissue and the interaction of these cells with the surrounding basement membrane. Interestingly, these studies have demonstrated that normal breast cells as well as breast cancer cell lines grown on this substrate will form sac-like structures similar to alveoli. Some of the differences between the normal and abnormal cells include the larger, disorganized nature of the breast cancer structures and the ability of normal but not abnormal "alveoli" to vectorally secrete milk proteins (Peterson *et al.*, 1992).

Genetic Alterations Associated with Breast Cancer

Homozygous deletion or loss of heterozygosity (LOH) has been observed in a number of breast cancer cells. This type of genetic alteration has been detected for chromosomes 1q, 3p, 7q, 11p, 13q, 13p, 16q, 17p, 17q and 18q (Callahan, 1989; Callahan and Campbell, 1989; Miyagi *et al.*, 1992 and references therein). Most of these loci are hypothesized to contain tumor suppressor genes. Significantly, structural changes to the retinoblastoma gene (Rb) located on chromosome 13q14 have been observed in 25% of breast tumor cell lines and 7% of primary tumors (T'Ang *et al.*, 1988). In almost all cases these changes were due to total or partial deletion of the gene, which resulted in loss or truncation of the transcript.

Recently linkage to a locus called *BRCA1*, on the long arm of chromosome 17 (17q12-21), has been demonstrated for familial breast and ovarian cancer (Hall *et al.*, 1990; Narod *et al.*, 1991; Black and Solomon, 1993;

Roberts, 1993). It has been determined that for nearly all the families studied with both breast and ovarian cancer and half of those with only breast cancer, the disease is linked to *BRCA1* (Black and Solomon, 1993; Easton *et al.*, 1993). Additionally, approximately 80% of the sporadic breast and ovarian cancers studied demonstrate LOH near this region (Futreal *et al.*, 1992; Foulkes *et al.*, 1991; Black and Solomon, 1993). These deletions provide evidence that the responsible gene may be acting as a tumor suppressor (Black and Solomon, 1993; Roberts, 1993). Some of these cancers may be due to the loss of p53 or the neurofibromatosis gene, both of which are tumor suppressor genes located on the short (17p13.1) and long (17q11.2) arms of chromosome 17, respectively. Significantly, families with germ-line mutations of p53 often have multiple breast cancers which represent about 1% of familial breast cancer cases (Black and Solomon, 1993). Additionally, a deletion on chromosome 17, which is distal to the *BRCA1* region (17q22-23), has been described for sporadic ovarian cancer (Jacobs *et al.*, 1993).

Current research in this area is focused on the identification of the exact gene responsible for the *BRCA1* familial cancers. Three candidate genes within this region, *NME1* (A chain of NDP kinase *nm23*), *THRA1* (thyroid hormone receptor gene) and *PHB* (prohibitin gene) have already been eliminated (Black and Solomon, 1993). Two remaining candidates, the *RARA* (retinoic acid receptor α gene) and the *EDH17β* (17ß-estradiol dehydrogenase gene), which are known to map within the *BRCA1* region, are currently under intense investigation. The significance of characterizing the *BRCA1* gene is further supported by the observation that male relatives of women with breast cancer have an increased risk of developing prostate cancer (Tulinius *et al.*, 1992) and carriers of the *BRCA1* deletion have an increased frequency of other cancers (Black and Solomon, 1993).

Amplification of proto-oncogenes has been reported by a number of research groups (reviewed in Callahan, 1989; Callahan and Campbell, 1989). The amplification of the c-*erbB2* gene (also called c-*neu* and *HER-2*) and the *int-2* gene has been observed in 10-40% and 9-23% of primary breast tumors, respectively. The c-*erbB2* gene, which encodes an epidermal growth factor receptor-like protein, is frequently co-amplified with the closely linked c-*erbA1* proto-oncogene, which is a member of the steroid/thyroid hormone receptor family. Similarly, the *int-2* gene, which is related to fibroblast growth factor, is often co-amplified with the *bcl-1* and *hst* loci of chromosome 11 (Callahan, 1989; Callahan and Campbell, 1989).

Biopsies of human breast cancer also frequently display amplification of the c-*myc* gene (Escot *et al.*, 1986; Mariani-Costantini *et al.*, 1988; Münzel *et al.*, 1991; Berns *et al.*, 1992; Pavelic *et al.*, 1992). Amplification of this gene has been reported to vary from 6-56% of breast neoplasias (Callahan, 1989; Callahan and Campbell, 1989), although one group has summarized the overexpression of c-*myc* to be between 70 and 100% of breast tumors (Berns *et al.*, 1992). Varying levels of c-*myc* have been found in tumors with similar ductal grade and histological appearance (Mariani-Costantini *et al.*, 1988; Pavelic *et al.*, 1992). Some of these variations have been attributed to the infiltration of lymphocytes in the tumor biopsy (Mariani-Costantini *et al.*, 1988). Alterations in c-*myc* other than amplification have sporadically been reported for breast cancer tumors (Escot *et al.*, 1986; Münzel *et al.*, 1991). Although often found associated with c-*myc* amplification, rearrangements and hypomethylation have been observed in 4-16% and 36% of breast biopsies, respectively (Escot *et al.*, 1986; Münzel *et al.*, 1991). In particular, the breakpoint of one c-*myc* rearrangement has been mapped near the 3'-end of exon 3 (Escot *et al.*, 1986). Although co-amplification of c-*myc* and c-*erbB2*

has been reported for 1-23% of breast tumors, no correlation between the two proto-oncogenes appears to exist (Münzel *et al.*, 1991; Berns *et al.*, 1992).

Some association of c-*erbB2* amplification and patient relapse has been observed, however amplification of c-*myc* is strongly associated with an early recurrence and early death (Berns *et al.*, 1992). In particular, 90% of patients with tumors expressing amplified c-*myc* relapsed as compared to 57% with normal c-*myc* levels. Similarly, only 31% of the patients with tumors expressing amplified c-*myc* survived the 10 year follow-up, compared to 60% with normal c-*myc* levels. One study found that 72% of breast biopsies with c-*myc* amplifications occurred in patients that were over 50 years of age, however correlations with receptor status or TNM grading were negative (Escot *et al.*, 1986). Conversely, in another study a positive association between c-*myc* amplification, tumor size and lymph node involvement was reported, but no correlation was detected for age, receptor status, tumor grade or menopausal status (Berns *et al.*, 1992). The measurement of c-*myc* amplification has therefore been suggested to significantly increase the power of other prognostic indicators.

Research on the general biology of breast cancer has been furthered by the development of cell lines and other models for this disease. Estrogens have been shown to regulate the proliferation of estrogen responsive breast cancer cell lines (Lippman *et al.*, 1976; Chalbos *et al.*, 1982). This suggests that many breast tumors originate from cells which express the ER and respond to ligand binding. The characterization of breast cancer cell lines that do not respond to estrogen suggests that these cells may be the result of cellular evolution to hormonal independence. The development of sublines from these two extreme breast cancer phenotypes (Vickers *et al.*, 1988; Clarke *et al.*, 1989; Jiang and Jordan, 1992) may provide interesting models for

characterizing the molecular mechanisms of breast cancer progression from hormonal dependence to independence. The use of mammary specific promoters has identified genes whose inappropriate expression leads to mammary tumors. Some of these genes have been shown to be important in the development of human breast cancer. In particular, the c-*myc* gene has been identified by experimental overexpression in the mammary gland of transgenic mice (Stewart *et al.*, 1984; Schöenenberger *et al.*, 1988) to be involved in breast cancer development. Significantly, the overexpression or amplification of c-*myc* in human breast cancer biopsies has also been correlated with a poor patient prognosis (Berns *et al.*, 1992). An analysis of the regulation of c-*myc* is therefore necessary to an understanding of its role in breast cancer development.

Steroid Hormone Receptors

Steroid hormone receptors are nuclear proteins which interact with *cis*-acting hormone response elements to enhance the expression of steroid regulated genes. Individual members of this family of *trans*-acting factors bind various ligands, including estrogen, progesterone, dihydrotestosterone, glucocorticoids, mineralcorticoids, retinoic acid, vitamin D and thyroid hormone. There appear to be two subgroups within this receptor family based on amino acid homologies of the DNA binding domain and DNA sequence similarities of the steroid response elements. This classification places the receptors for progesterone, androgen, glucocorticoid and mineralcorticoid in one subfamily and the receptors for estrogen, vitamin D, retinoic acid and thyroid hormone in a second subfamily. There are additional members of this family of proteins, whose ligands are yet to be identified, which appear to also belong to the second subgroup. Recent

studies have cast doubt over the hypothesis that ligand binding induces receptors to form homodimers which are then able to bind their response elements. Ligand interactions which appear to be required only for the *trans*-activating function of the receptors and the observation of monomeric and heteromeric receptor interactions with DNA suggest that the classical model for transcriptional regulation by the members of this receptor family needs to be reassessed.

Structural Organization

Our understanding of steroid hormone action advanced considerably with the cloning of the genes for the ER and other members of the steroid/thyroid hormone receptor family (Walter *et al.*, 1985; Green *et al.*, 1986). The mRNA for the ER is between 6.2-7.5 kilobase pairs (kb), depending on the species, and encodes a 589-600 amino acid protein with a molecular mass of approximately 66.5 kDa. The structural and functional domains of the ER are encoded by eight exons located on human chromosome 6q24-27 (Ponglikitmongkol *et al.*, 1988). The 5' untranslated sequences and the A/B region are encoded by the first exon, while exons 2 and 3 encode region C, with each zinc finger of the DNA binding domain represented by a separate exon (Figure 2). Exon 4 encodes all of region D and part of region E, with the balance of region E encoded by exons 5 to 8. The remaining domains, region F and the large 3' untranslated region, are encoded by exon 8. With the exception of the thyroid hormone (T3R) and retinoic acid (RAR) receptors, which consist of multiple types each encoded by different genes present at more than one chromosomal location (Parker, 1988), the other members of the steroid/thyroid hormone receptor family are transcribed from single genes.

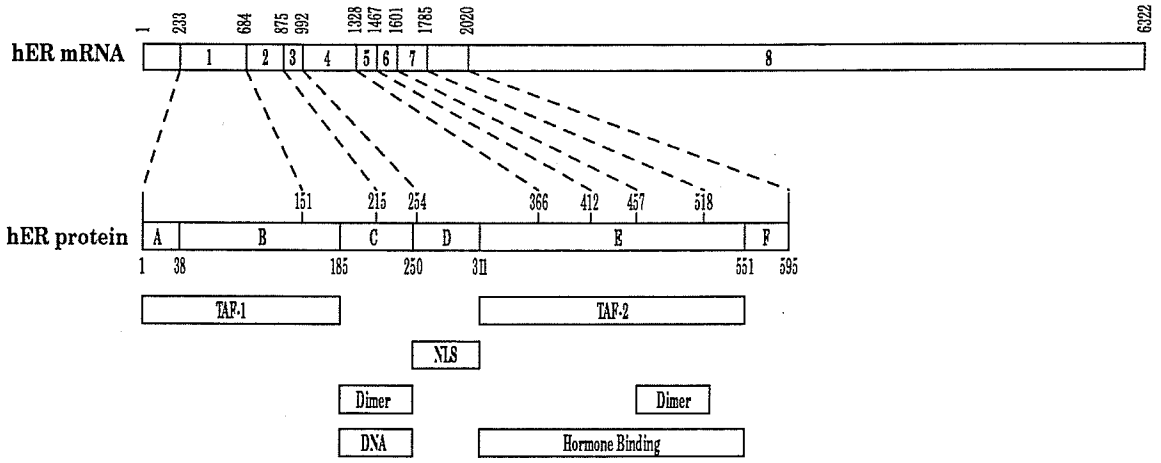


Figure 2. Estrogen Receptor Domain Structure and Associated Functions. The human estrogen receptor gene consists of eight exons which are shown spliced into the hER mRNA. This mRNA contains a 5' untranslated region encompassing nucleotides 1-233 and a large 3' untranslated region spanning nucleotides 2020-6322. The coding sequences of the mRNA (bp 233-2020) are aligned with the amino acids (aa 1-595) of the ER protein domain structure. The functional regions of the ER protein are diagrammed under the ER protein and represent the *trans*-activation regions (TAF-1, TAF-2), the nuclear localization signal (NLS), the dimerization region (dimer), the DNA binding domain (DNA) and the hormone binding domain (hormone binding). This figure was modified from Beato, 1989 according to information derived from references listed in the text. .

Various structure and function analyses have demonstrated that the steroid hormone receptors consist of a modular array of discrete functional domains (reviewed in Green *et al.*, 1986; Evans, 1988; Beato, 1989). Comparisons between the human and chicken ER sequences have revealed three regions of significant homology (A, C and E) separated by regions of lesser identity (B, D and F) (Krust *et al.*, 1986; Green *et al.*, 1986). The hormone binding domain has been assigned to region E, while dimerization motifs have been localized to both regions E and C (Fawell *et al.*, 1990). Region E also contains sequences (TAF-2) involved in the transcriptional activation function of the receptor (Kumar *et al.*, 1987; Webster *et al.*, 1989). Additional regions having *trans*-activation function (TAF-1) are found within the A/B domain of the ER (Tora *et al.*, 1989a). Region C is responsible for the sequence-specific DNA binding function of the receptor, while a centrally located nuclear localization signal is found in region D (Didier *et al.*, 1990).

The functional regions of the other steroid hormone receptors share similar domain organization but are differentially conserved between the members of this receptor family (Green *et al.*, 1986; Evans, 1988; Beato, 1989). In particular, region C, which contains DNA binding function, is the most highly conserved, while limited homology is retained in the hormone binding domain contained in region E. Conversely, the A/B domain exhibits considerable difference between the various steroid/thyroid hormone receptors. This region along with domains D, E and F share vary little sequence similarity between the members of the steroid/thyroid hormone receptor family.

Functional Models of Steroid Hormone Action

The classical model for steroid receptor action proposes that ligand interaction with the receptor is required before nuclear transportation, receptor dimerization and sequence-specific interaction with the hormone response element (HRE) can occur. However, controversy now exists over the hypothesis that ligand interaction with ER is necessary for interaction with the estrogen response element (ERE). Additionally, there is speculation that this receptor is able to interact with DNA in its monomeric form. A more recent model suggests that the ER binds to its response element as a monomer and is stabilized by interaction with an additional protein(s) to form heteromeric complexes (Gorski *et al.*, 1993). Additionally, this model proposes that interaction of the monomeric ER with its response element occurs in the absence of hormone. Estrogen interaction with the ligand binding domain of the receptor induces conformational changes in the receptor, which stimulate the *trans*-activation functions and result in the transcriptional activation of estrogen responsive promoters.

It was previously thought that the binding of estrogen to its receptor induced dimer formation. This interaction was proposed to unmask the DNA binding domain of the receptor by the displacement of receptor associated proteins and/or structural changes in receptor conformation (reviewed in Green and Chambon, 1986). The ability of an ER lacking the hormone binding domain to form dimers was suggested to provide evidence that the unoccupied steroid binding domain destabilized the dimerization mediated by the DNA binding domain (Kumar and Chambon, 1988). Additionally, differences in protein isolation methodology have been suggested as reasons for ER binding to DNA in the absence of estrogen (Klein-Hitpass *et al.*, 1989). Estradiol or tamoxifen interaction has been reported to induce dimer

formation in the intact receptor, which resulted in a stronger association with the ERE (Kumar and Chambon, 1988). However, the ER construct (HE0) used to undertake the above studies was later confirmed to have a single mutated amino acid which altered ligand affinity and resulted in a more labile receptor (Tora *et al.*, 1989b; Gronemeyer, 1991). Significantly, it has been known for at least twenty years that the unliganded receptor is more labile than the hormone bound receptor (Gorski *et al.*, 1993). Subsequent observations with a wild type receptor have not demonstrated any effect of estrogen on ER affinity for the ERE by *in vitro* mobility shift or quantitative avidin-biotin-DNA complex assays (Gorski *et al.*, 1993). Significantly, one group has reported that in the absence of magnesium ions (Mg^{2+}) the ER is able to interact with its binding site in the absence or presence of estrogen, but hormone is required when Mg^{2+} is present (Brown and Sharp, 1990).

Although *in vivo* estrogen binding studies support monomeric binding of the ER (Gorski *et al.*, 1993), at least two independent groups have reported that the ER interacts with its DNA recognition site as a dimer *in vitro* (Kumar and Chambon, 1988; Klein-Hitpass *et al.*, 1989). The consensus ERE consists of the sequence 5'-AGGTCA-n3-TGACCT-3', where n is any nucleotide (Figure 3A). Direct interaction of the ER with the guanine residues of both half-sites on both faces of the DNA has been demonstrated by *in vitro* methylation interference (Klein-Hitpass *et al.*, 1989). This has led to the suggestion that ER binding occurs in the dimeric form. Both intact ER and receptors lacking the hormone binding domain have been demonstrated to interact with the ERE as homodimers (Kumar and Chambon, 1988). However, a more recent report has demonstrated monomeric ER interaction with an intact ERE and an ERE with a one nucleotide substitution (Figure 4) which destroyed the dyad symmetry of the element (Medici *et al.*, 1991).

	Response Element		Consensus Sequence	
A	GRE/PRE/ARE/MRE		GGTACA-n_3-TGTTCT	(1)
	TRE/RARE/DRE (half site)		TCAGGTCA-	(1)
	ERE		AGGTCA-n_3-TGACCT	(1)
B	Vitellogenin A1	*Xenopus*	GGTCA-n_3-TGACC	(2,3)
	Vitellogenin A2	*Xenopus*	GGTCA-n_3-TGACC	(2,3)
	Vitellogenin B1a	*Xenopus*	AGTTA-n_3-TGACC	(2,4)
	Vitellogenin B1b	*Xenopus*	AGTCA-n_3-TGACC	(2,4)
	Vitellogenin II	Chicken	GGTCA-n_3-TGACC	(2,3)
	ApoVLDL II	Chicken	GGTCA-n_3-TGACT	(2)
	Ovalbumin (half sites)	Chicken	$(TGACC)_4$ GGTCA	(5)
	Prolactin	Rat	TGTCA-n_3-TGTCC	(6)
	β-Leutenizing Hormone	Rat	GGACA-n_5-TGTCC	(7)
	Uteroglobin	Rabbit	GGTCA-n_3-TGCCC	(8)
	pS2	Human	GGTCA-n_3-TGGCC	(9)
	Oxytocin	Human	GGTGA-n_3-TGACC	(10)
	Lactoferrin	Human	GGTCA-n_3-CGATC	(11)
	c-fos	Human	CGGCA-n_3-TGACC	(12)

C Consensus ERE

(top strand)

```
  GG       C  TT
NGTCA-n3-5-TGNCC
  AT
```

(bottom strand)

```
  TC       G  AA
NCAGT-n3-5-ACNGG
  CA
```

Figure 3. Hormone Response Elements of the Steroid Receptor Family. A. The consensus response elements of the glucocorticoid, progesterone, androgen and mineralcorticoid receptors (GRE/PRE/ARE/MRE) are compared to the response element of the estrogen receptor (ERE) and the half site element of the thyroid hormone, retinoic acid and vitamin D receptors (TRE/RARE/DRE). **B.** ERE sequences identified in a variety of estrogen responsive genes and their associated references are shown. **C.** A more complete consensus ERE based on these published sequences is also shown. Numbers in brackets correspond to the following references: (1) Beato, 1989; (2) Walker *et al.*, 1984; (3) Klein-Hitpass *et al.*, 1988a; (4) Martinez and Wahli, 1989; (5) Kato *et al.*, 1992b; (6) Waterman *et al.*, 1988; (7) Shupnik *et al.*, 1989; (8) Slater *et al.*, 1990; (9) Berry *et al.*, 1989; (10) Richard and Zingg, 1990; (11) Teng *et al.*, 1992; (12) Weisz and Rosales, 1990. This figure was partly reproduced from Dubik, 1991.

Multiple mutations destroyed both monomeric and dimeric ER binding. Additionally, the DNA binding domains of a number of steroid receptors, including the ER, have been shown to interact first with the downstream half-site of the HRE (Tsai *et al.*, 1988; Ponglikitmongkol *et al.*, 1990). This binding subsequently facilitated the association of the receptor with the upstream half-site of the response element. Significantly, the interaction of the ER with half-site EREs separated by more than 100 nucleotides has recently been described (Kato *et al.*, 1992b).

In support of monomeric ER association with DNA is the observation that the orientation and spacing of various response element half-sites can dictate the oligomeric form of the interacting receptor and the transcriptional response at the promoter (Näär *et al.*, 1991; Umesono *et al.*, 1991; Forman *et al.*, 1992). The half-site repeats of the consensus ERE are identical to the T3R, RAR and vitamin D3 receptor (D3R) elements except for differences in orientation and spacing (Figure 4). Binding and transcriptional activation at half-sites in the palindromic (head to head), direct repeat and inverted palindromic (tail to tail) orientations was observed for the ER, RAR and T3R, respectively (Näär *et al.*, 1991). The D3R, T3R and RAR were found to specifically *trans*-activate reporter genes containing half-site elements arranged as direct repeats with a spacing interval of three, four and five nucleotides, respectively (Umesono *et al.*, 1991). Additionally, palindromic half-sites and direct repeats, which lacked spacing nucleotides, were bound and activated by both RAR and T3R, while direct repeats separated by 2 nucleotides no longer bound RAR and were repressed by T3R binding (Näär *et al.*, 1991).

Similarly, the ability of the RAR and T3R to interact with HREs as monomers, homodimers or heterodimers is controlled by the characteristics of

Orientation	Spacing	Sequence	Receptor Binding	Receptor Form			Fold Response
				Monomer	Homodimer	Heterodimer	
Palindrome	0	AGGTCATGACCT	T_3R	+	+	+	87
			RAR		+	+	48
Palindrome	+3	AGGTCA-n_3-TGACCT	ER	+	+		5
			T_3R	+			59
			RAR		+		35
m-Palindrome	+3	AGGTCA-n_3-TCACCT	ER	+			ND
Palindrome	+5	AGGTCA-n_5-TGACCT	T_3R	+		+	82
			RAR	+		+	32
Palindrome	+10	AGGTCA-n_{10}-TGACCT	T_3R	+			8.3
			RAR	+			2.6
Inverted Pal	+4	TGACCT-n_4-AGGTCA	T_3R				ND
Inverted Pal	+7	TGACCT-n_7-AGGTCA	T_3R				3.0
Direct Repeat	0	AGGTCAAGGTCA	T_3R	+			4.2
			RAR	+			2.3
Direct Repeat	+2	AGGTCA-n_2-AGGTCA	T_3R				-3.5
Direct Repeat	+3	AGGTCA-n_3-AGGTCA	D_3R				11
Direct Repeat	+4	AGGTCA-n_4-AGGTCA	T_3R				6.0
Direct Repeat	+5	AGGTCA-n_5-AGGTCA	T_3R	+	+		7.6
			RAR		+		13
Half Site	-	AGGTCA	T_3R	+			4.4
			RAR	+			

Figure 4. Orientation and Spacing of Half Site Response Elements. The interactions of the estrogen, thyroid hormone, retinoic acid and vitamin D receptors with two half site elements in the palindrome, inverted palindrome and direct repeat orientations are shown (m-palindrome refers to a mutation in one half site). For consistency between the different studies, the spacing between the two half site AGGTCA elements was used to generate the separation differential of 0-10 nucleotides. The receptor form involved in DNA binding and fold response was not determined (ND) by some of the studies. This figure was derived from information found in Näär *et al.*, 1991; Medici *et al.*, 1991; Umesono *et al.*, 1991; Forman *et al.*, 1992.

the core DNA binding motif (Forman *et al.*, 1992). Comparisons of the relative abundance between the monomeric and dimeric forms of the receptors indicated that predominantly monomeric binding was observed for RAR at direct repeats without spacing and palindromic repeats separated by five or ten nucleotides, while dimers were more efficiently formed on palindromes separated by zero or three nucleotides and dimeric repeats separated by five nucleotides (Figure 4). Conversely, equivalent binding of monomeric and dimeric forms of the T3R was seen on palindromic repeats without spacing and direct repeats separated by five nucleotides, while monomeric forms predominated in unspaced direct repeats and palindromic repeats separated by three, five or ten nucleotides. Both receptors were able to bind as monomers on isolated half-sites, while heterodimers between T3R and RAR were observed on palindromic repeats with and without a five base pair spacing. Interestingly, both of these receptors were able to interact with the various HREs in the absence of hormone. These observations suggest that subtle differences in the response element can direct the interactions of steroid hormone receptors with responsive promoters and modify the ultimate response (Forman *et al.*, 1992).

The *trans*-activation function (TAF-2) of the hormone binding domain has been reported to require induction by steroid interaction with the receptor (Kumar *et al.*, 1987; Webster *et al.*, 1988). Contrary to receptors lacking this domain, receptors, which are mutated in this region, cannot form dimers, bind ERE weakly if at all and are void of *trans*-activation function. The *trans*-activation function of receptors, which completely lack the hormone binding domain, has been considered to be constitutive (Didier *et al.*, 1990). The interaction of estrogen with the steroid binding domain has been shown to induce conformational changes in the ER which result in a tighter

association with the nucleus. Moreover, these structural changes appear to differ from those induced by antiestrogen binding (Gorski *et al.*, 1993). Interestingly, the interaction of the ER DNA binding domain with the ERE has been shown to decrease the rate of estrogen dissociation from the steroid-binding domain (Fritsch *et al.*, 1992). It has subsequently been proposed that these structural alterations contribute to receptor interactions with additional *trans*-activators, the general transcription factors or other chromatin-associated proteins (Gorski *et al.*, 1993). In support of this hypothesis is the observation that the PR interacts with intermediary factors found in the TFIID chromatographic fraction (Shemshedini *et al.*, 1992). Additionally, both forms of the PR have been shown to interact with its response element in an identical manner, but only form A was able to enhance expression of the ovalbumin promoter (Tsai *et al.*, 1988). Significantly, a role for intermediary factors in the *trans*-activation function of the steroid hormone receptors has been suggested (Bocquel *et al.*, 1989).

The data presented therefore support a functional role for the ER as a transcriptional activator. The observation that the ER interacts with its response element as a monomer both *in vitro* (Medici *et al.*, 1991) and *in vivo* (Gorski *et al.*, 1993) supports the suggestion that homodimer formation may not be necessary for DNA interaction. Additionally, this interaction does not appear to be dependent on ligand binding (Gorski *et al.*, 1993; Brown and Sharp, 1990), however the *trans*-activation functions of the receptor do require hormone interaction (Kumar *et al.*, 1987; Webster *et al.*, 1988). Possibly the binding of ligand allows the DNA bound receptor to interact with other transcriptional regulators to bring about the activation of estrogen regulated genes.

The Components of Transcriptional Regulation

Characterizing the factors involved in the pathway towards transcription initiation is a prerequisite to unraveling the mechanisms of gene activation. The components of the general transcriptional machinery aid RNA polymerase II in the recognition and initiation of transcription. The synthesis of messenger RNA (mRNA) is regulated in part by sequence-specific DNA-binding proteins that interact proximal and distal to the promoter at defined *cis*-acting elements. These *trans*-activating factors are involved in relieving the repression exerted by the nucleosomal packaging of DNA. Alterations in chromatin structure precede or accompany gene activation, converting a transcriptionally repressed gene into a competent state. These observable differences are believed to reflect the disruption of histone interactions with the DNA and the binding of other transcriptional regulators. The entire process is spatially confined within the nucleus by the association of chromatin and the components of transcriptional regulation with the nuclear matrix.

Chromatin Organization

Chromosomes of higher eukaryotes are thought to be organized into a series of discrete and topologically independent higher-order domains (for reviews see Getzenberg *et al.*, 1991; Bonifer *et al.*, 1991; Jackson, 1991). DNA is first compacted by its association with the nucleosomal core histones to form the 10-nm fiber (Figure 5). The association of the linker histone with the nucleosomal DNA and proteins further compacts the chromatin into the 30-nm fiber. The attachment of DNA sequences to components of the nuclear matrix subsequently allows the formation of the chromatin loop domain. At least three types of nuclear matrix association, reflecting different functional roles, are involved in maintaining these domains. These interactions are

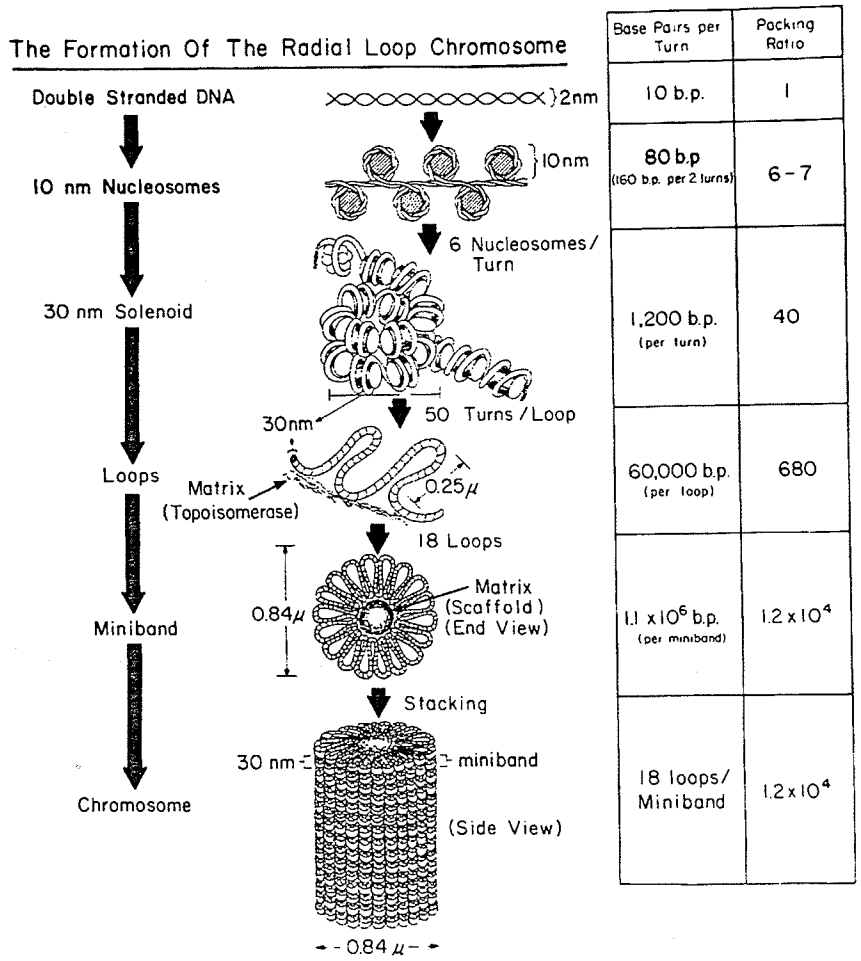


Figure 5. Chromatin Organization. A diagram of the various levels of DNA organization and their associated packing ratio. The 2 nm double stranded DNA helix is wound twice around the histone octamer to form the 10 nm fiber known classically as the "beads on a string" form of DNA. The continued winding of 6 nucleosomes per turn further compacts the chromatin into the 30 nm solenoid fiber which associates with the nuclear matrix to form the loop domains. These loops are wound along a central axis to form minibands which are stacked into the chromosome. This figure was reproduced from Getzenberg *et al.*, 1991.

involved in controlling the transcriptional activity of the chromatin loop domain.

The Nucleosome

Packaging into nucleosomal structures compresses the length of DNA 30 to 40 fold, whereas, the higher-order folding observed in the chromosome has a packaging ratio of 1:10,000 (Gasser and Laemmli, 1987). The nucleosomal proteins, which consist of the core histones H2A, H2B, H3 and H4, and the linker histone H1 are involved in DNA packaging (reviewed in Hayes and Wolffe, 1992). Recently, X-ray crystallography has confirmed that the structure of the nucleosome consists of an octamer of two histone H2A-H2B dimers bound to a centrally located histone H3-H4 tetramer (Arents *et al.*, 1991). Histone H1 binds across the entry and exit points of the DNA wrapped around the nucleosome (Boulikas *et al.*, 1980). The highly basic tails of this histone neutralize the negatively charged phosphodiester backbone of linker DNA allowing arrays of nucleosomes to be folded into the 30-nm fiber (Hayes and Wolffe, 1992). The length of supercoiled DNA associated with the core histones varies widely in eukaryotes, but is approximately two full turns which is equivalent to about 165 bases (Grunstein, 1990). In one study, 196 base pairs of DNA was protected by the nucleosome from micrococcal nuclease digestion in all of the mature tissues analyzed (Compton *et al.*, 1976). Fewer bases were protected from nuclease attack with cells in culture, for instance the DNA repeat length in HeLa cells appeared to be 188 bp. This variation in length is likely due to differences in the length of linker DNA between nucleosomes.

The Chromatin Loop Domain

The next level of chromatin organization is the folding of nucleosome packaged DNA into chromatin loops (Figure 5). These domains are visible in electron micrographs of interphase nuclei and metaphase chromosomes (Goldman, 1988; Van Driel, *et al.*, 1991). The size of a chromatin loop domain is variable and may consist of a single gene, such as the chicken lysozyme domain (Bonifer *et al.*, 1991) or several genes, such as the Drosophila histone and human globin gene families (Van Driel *et al.*, 1991). They have been described to range from 5 to 100 kb in length with highly expressed genes located in a subset of small loops up to 13 kb (Gasser and Laemmli, 1987). The major factor influencing the nuclease susceptibility of chromatin is the conformation in which it is held by the nucleosomal proteins (Whitlock *et al.*, 1977). A characteristic feature of transcriptionally active chromatin is the increased susceptibility to digestion by nucleases, especially to digestion by deoxyribonuclease I (DNase I), compared to the relative insensitivity of inactive or repressed chromatin. The preferential DNase I sensitivity of active genes is not restricted to the coding portion of the gene but extends far upstream and downstream into adjacent non-transcribed DNA sequences (Reeves, 1984; Phi-Van and Strätling, 1988). This sensitive chromatin domain constitutes a chromosomal loop and thus represents a unit of gene regulation (Yaniv and Cereghini, 1986).

Nuclear Matrix Association

The nucleus is organized around a structural network or framework (reviewed in Berezney, 1991; Van Driel *et al.*, 1991) called the nuclear matrix in interphase nuclei (Berezney and Coffey, 1974) and the nuclear scaffold in metaphase chromosomes (Paulson and Laemmli, 1977). Two of the major

functions of this nuclear architecture are to organize discrete loops of chromatin (Figure 5) and to compartmentalize nuclear operations (Van Driel *et al.*, 1991). The processes of replication, transcription, RNA processing and RNA transport are all activities associated with the nuclear matrix and appear to be discretely organized. Additionally, the nuclear matrix may be structurally linked to the cytoskeleton, the cell periphery and the extracellular matrix (Getzenberg *et al.*, 1991). Although the list of proteins associated with the nuclear matrix is far from complete, a number of significant proteins have already been described, in particular, the enzymes DNA polymerase α, RNA polymerase II and topoisomerase II (Cook, 1989; Jackson, 1991; Berezney, 1991). Additionally, a variety of non-histone DNA-binding proteins (Dworetzky *et al.*, 1992; Bidwell *et al.*, 1993; van Wijnen *et al.*, 1993), including steroid hormone receptors (Barrack and Coffey, 1980; Alexander *et al.*, 1987) and c-Myc (Eisenman *et al.*, 1985), have been associated with the nuclear matrix. Significantly, nuclear matrix protein patterns have been shown to be influenced by hormonal status (Getzenberg and Coffey, 1990) and differ between normal and neoplastic tissue (Partin *et al.*, 1993). The association of these enzyme activities and regulators of transcription with the nuclear matrix and its fluctuating composition are important indications of the association of nuclear function with this structure.

A significant body of research supports the hypothesis that chromatin domains are maintained by their association with the nuclear matrix (reviewed in Berezney, 1991; Bonifer *et al.*, 1991; Van Driel *et al.*, 1991; Boulikas, 1992; Felsenfeld, 1992). Three classes of attachment are generally described; those that are stable and define structural domains, those that control transcriptional competence, and those involved in transient

associations conferred by RNA polymerase dependent interactions. The structural attachments of chromatin domains are maintained through DNA elements variously called matrix attachment regions, MARs, (Cockerill and Garrard, 1986), scaffold associated regions, SARs, (Mirkovitch, *et al.*, 1984; Gasser and Laemmli, 1987) and A-elements (Phi Van and Strätling, 1988) Conversely, locus control regions, LCRs, (Grosveld *et al.*, 1987) are dominant acting elements which appear to regulate transcriptional competence. Finally, the transient interaction of *trans*-acting factors with their DNA elements allows for the formation of functional regulatory complexes which subsequently interact with a matrix bound RNA polymerase (Boulikas, 1992).

Stable Matrix Attachment

MARs, SARs and A-elements appear to perform similar functions by forming permanent attachments of chromatin with the nuclear matrix. MARs and SARs were initially detected as the residual DNA associated with nuclear matrix structures isolated by a variety of techniques, including hypertonic salt and ionic detergent extractions and encapsulation in agarose beads, all in combination with nuclease treatment (Jackson, 1991). In all cases DNA fragments ranging in size from 0.5 to 2 kb in length were found to be firmly associated with nuclear structural components. These sequences do not share extensive sequence similarity but are generally long stretches of adenine and thymine (A+T) rich DNA containing topoisomerase II consensus sequences (Getzenberg *et al.*, 1991; Boulikas, 1992; Bode *et al.*, 1992). The A+T rich sequences of S/MARs are also reminiscent of origins of replication and the core recognition sequences of homeodomain proteins (Boulikas, 1992). Significantly, origins of replication (Razin *et al.*, 1986) and the homeodomain protein Oct-1 (van Wijnen *et al.*, 1993) have been found associated with the nuclear matrix. Conversely, A-elements were identified by their ability to

protect randomly integrated reporter genes containing weak promoters from position effects (Bonifer *et al.*, 1991). They contain S/MAR sites but are generally longer in length (Grunstein, 1990).

All of these attachment sites are located in the 5' and 3' distal regions of DNase I sensitive genes and are believed to mark the boundaries of a functional chromatin domain (Bonifer *et al.*, 1991). Significantly, S/MARs/A-elements confer elevated, position-independent but copy number dependent expression of permanently integrated transgenes (Grunstein, 1990; Phi-Van *et al.*, 1990; Bonifer *et al.*, 1991). The transcriptional activation of promoters by these attachment sequences was not detected in transient transfection assays. Additionally, their full activity in stable systems required a complete transcriptional unit. Together these studies suggest that these elements differ from traditional enhancers and only act in a normal chromatin environment (Phi-Van *et al.*, 1990; Bonifer *et al.*, 1991).

In general, these structural elements anchor chromatin domains to the nuclear matrix thereby protecting them from the influences of neighbouring chromatin (Felsenfeld, 1992). Some described functional features of S/MARs are to constrain the chromatin domain, relax positive supercoiling of the domain, recruit topoisomerases and prevent superhelical strain from being transmitted to neighbouring domains (Bode *et al.*, 1992). They are often found in close proximity to *cis*-acting regulatory DNA elements identified either genetically or functionally (Cockerill and Garrard, 1986; Gasser and Laemmli, 1987; Phi Van *et al.*, 1990). A number of proteins, which bind these elements, have been described (Phi-Van *et al.*, 1990; von Kries *et al.*, 1991; Berezney, 1991). Two such factors (NMP-1 and NMP-2) interact with similar sequences to the binding sites for the ATF and C/EBP transcription factors, respectively (Dworetzky, *et al.*, 1992; Bidwell *et al.*, 1993), while another

(ARBP) bears little resemblance to typical *trans*-activators but is suggested to induce DNA loop formation (von Kries *et al.*, 1991).

Transcriptional Competence Attachments

Exogenous genes which lack any matrix associating sequences but contain promoter sequences are unable to drive efficient expression in transfection and often fail to give any expression in transgenic mice. Similarly, transgenes, which contain only promoter and enhancer elements, succumb to position effects characterized by low expression which is not temporal or tissue specific (Dillon and Grosveld, 1993). LCRs are dominant control regions, which have presently been described only for the *globin* gene domain. They are located at the 5'-end of the domain and induce sensitivity to DNase I over an extensive region of chromatin (Bonifer *et al.*, 1991; Dillon and Grosveld, 1993). They are *cis*-acting elements which yield position-independent but copy number dependent gene expression (Grosveld *et al.*, 1987; Dillon and Grosveld, 1993). Additionally, their presence is required for correct temporal activation and cell-type specific expression (Jackson, 1991; Bonifer *et al.*, 1991). The activity of LCRs therefore appears similar to S/MARs/A-elements since they allow gene expression to be independent of the surrounding chromatin environment. Similarly, they confer high level expression but only have this enhancer-like function when stably integrated into chromatin (Grosveld *et al.*, 1987; Dillon and Grosveld, 1993). Additionally, LCRs contain binding sites for *trans*-acting factors, suggesting they act as powerful *cis*-acting elements which exert a dominant effect. Some of these functions are partially separable suggesting the LCR encompasses a number of functionally distinct elements (Felsenfeld, 1992).

The difference between the activity of an LCR and the structural attachments facilitated by S/MARs/A-elements, appears to be the ability of an

LCR to exert its effect when only placed at the 5'-end of a gene. Conversely, the structural attachments of S/MARs/A-elements are normally found at both the 5' and 3' extremes of a chromatin domain (Felsenfeld, 1992). The sequences described for S/MAR/A-elements range in size from 0.5 to 2 kb, suggesting it is possible that some of the functions ascribed to these elements may be due to the inclusion of an LCR in the 5' matrix associated DNA. The primary role of an LCR may be to act as a dominant signal to override nucleosome inhibition and keep the promoter free of histones by interacting with the classical promoter/enhancer elements of individual genes (Jackson, 1991; Bonifer *et al.*, 1991; Felsenfeld, 1992). Significantly, LCRs induce the formation of tissue specific, developmentally stable DNase I hypersensitive sites, which reflect nucleosome free DNA (Dillon and Grosveld, 1993).

Transient Matrix Attachments

It has been suggested that temporal associations of a gene with the nuclear matrix occur through proteins interacting at enhancer and promoter sequences (Boulikas, 1992). These attachments would be defined by the transcriptional activity of a particular gene and the developmental stage of the cell. Recently a number of *trans*-acting factors have been isolated as components of the nuclear matrix (Dworetzkey *et al.*, 1992; Bidwell *et al.*, 1993; van Wijnen *et al.*, 1993). Significantly, proteins related to the *trans*-acting factors Sp1 and ATF were found in both nuclear matrix and nonmatrix protein fractions, while factors interacting with the CCAAT and C/EBP motifs were nuclear matrix specific and cell type dependent (van Wijnen *et al.*, 1993). Additionally, these same authors found that factors from both nuclear matrix and nonmatrix extracts bound AP1 consensus sequences but differences were detected in the association of Fos and Jun proteins. The Jun-B and Jun-D proteins were associated with both nuclear

matrix and nonmatrix fractions, however Fos proteins were detected only in the nonmatrix component. These discoveries lend support to the previously described association of steroid hormone receptors (Barrack and Coffey, 1980; Alexander *et al.*, 1987) and c-Myc (Eisenman *et al.*, 1985) with the nuclear matrix.

Although no association of the general transcription factors with the nuclear matrix has been reported, the association of RNA polymerase II with this structure is well documented (Abulafia *et al.*, 1984; Cook, 1989). The interaction of matrix associated *trans*-acting factors with their *cis*-acting elements and the transient association of these regulatory components with a matrix bound polymerase supports a functional role for nuclear matrices in supporting the transcription process (Cook, 1989). Significantly, the ability of the transcription factor ATF to facilitate the formation of a preinitiation complex (Horikoshi *et al.*, 1988) combined with the association of this factor with nuclear matrix protein preparations (van Wijnen *et al.*, 1993) and MAR sequences (Dwortzky *et al.*, 1992) is strongly suggestive of a role for *trans*-acting proteins in catalyzing these interactions.

A significant body of research therefore suggests that the DNA molecule as it exists in the nucleus is structurally organized for spatial and functional reasons. The association of DNA with histones and the nuclear matrix restricts the DNA to the limited space of the nucleus and assists in the formation of transcriptionally active regions of the genome. In addition to providing a mechanism for chromatin compaction, these higher-order domains are thought to define independent units of gene activity. Models for the remodeling of chromatin from repressed into active domains must address the need for activator proteins and the transcription machinery to overcome the obstructive packaging of DNA into nucleosomes and higher

order structures. The maintenance of these domains by their interaction with the nuclear matrix and the concomitant association of various enzyme activities with this nuclear framework challenges the accepted principles of nuclear function.

Classical Regulation of Transcription

The conventional model of transcription presumes that a mobile polymerase progresses along a fixed DNA template to synthesize an RNA molecule (reviewed in Cook, 1989). This model presupposes that the polymerase, the associated transcription factor complex and the *trans*-acting DNA binding proteins are all soluble and able to freely diffuse towards the DNA template. More recent studies suggest that this supposition is incorrect and that the polymerase, topoisomerase II, transcription factor complex, *trans*-acting factors and the RNA processing apparatus are all associated with the nuclear matrix (reviewed in Cook, 1989; Berezney, 1991; Carter and Lawrence, 1991; Van Driel, *et al.*, 1991). Additionally, evidence already described suggests that the DNA template is also tethered to the nuclear matrix (reviewed in Berezney, 1991; Bonifer *et al.*, 1991; Van Driel *et al.*, 1991; Felsenfeld, 1992). From these studies it is possible to suggest that nuclear matrix associations play an important role in the remodeling of chromatin.

The predominant view of transcriptional activation revolves around the ability of *trans*-acting factors to regulate transcription initiation through their *cis*-acting elements (reviewed in Mitchell and Tjian, 1989). Usually, *cis*-acting elements are located near the promoter, but some elements can exert their control over great distances in a position independent manner (Ptashne, 1986; Mitchell and Tjian, 1989). These sequence elements are

described as promoter proximal when located in the near vicinity of the promoter or distal when located at a substantial distance from the promoter. Interactions between factors associated with unconnected elements are facilitated by the looping out of the intervening DNA (Ptashne, 1986). Unique combinations of *cis*-acting elements provide a foundation for the assembly of complicated arrays of transcriptional regulators which interact directly or indirectly with the transcription machinery to confer on each gene an individual spatial and temporal transcription program (Mitchell and Tjian, 1989). The recent description of several effector proteins which facilitate the *trans*-activation functions of DNA-binding proteins provide evidence that the complex arrangements of proteins that form at DNA sites function together to activate transcription (reviewed in Martin, 1991).

Transcriptional Machinery

The assembly of the RNA polymerase II transcription initiation complex occurs at a region of the promoter referred to as the core promoter. This region contains the TATA box sequence and possibly other accessory sequences near the site of transcription initiation which aid in the assembly of the preinitiation complex. The general transcription factors are essential to the formation of this complex and assist RNA polymerase in recognizing the promoter (reviewed in Saltzman and Weinmann, 1989; Kadonaga, 1990; Ham *et al.*, 1992). Recognition of the TATA box sequence by the binding of the TATA box-binding protein (TBP) is the first event in the sequential assembly of the preinitiation complex (Figure 6). This interaction is essential to the binding of the remaining general transcription factors to the core promoter. Factors associated with TBP form a multiprotein complex known as TFIID, which was originally identified as the chromatographic fraction required to initiate the formation of the preinitiation complex (reviewed in

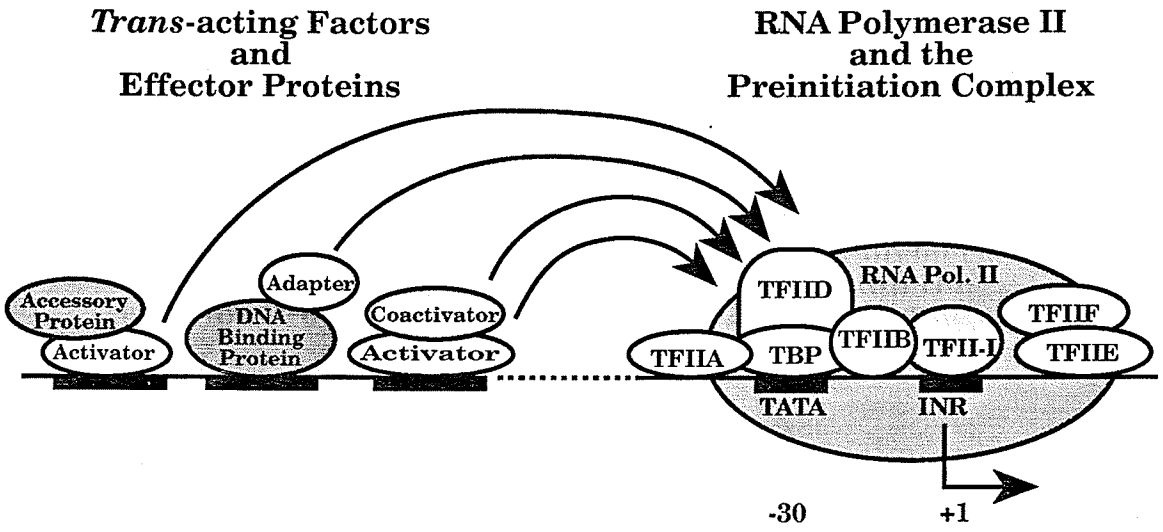


Figure 6. Regulators of Transcription: The Transcription Initiation Complex, *Trans*-acting Factors and Effector Proteins. The components of the transcriptional machinery include the transcription factor complexes TFIIA, TFIID, TFIIB, TFIIF and TFIIE. TFIID is a multiprotein complex containing the TATA binding protein (TBP) which interacts with the promoter DNA at the TATA box (TATA) to trigger the formation of the preinitiation complex. TFII-I is necessary for the association of TFIID at TATA-less promoters and interacts with an initiator element (INR). The order of transcription factor binding is described in the text. The various types of *trans*-acting factors and effector proteins include activators, accessory proteins, DNA binding proteins, adapters and coactivators. The arrows indicate those factors which can directly interact with the transcriptional machinery. This figure was derived from Ham *et al.*, 1992; Martin, 1991.

Saltzman and Weinmann, 1989). TFIIA is unable to bind the promoter on its own and requires the initial binding of TFIID for its assembly upstream ofthe TATA box. The binding of TFIIA to the promoter stabilizes the interaction of TBP with the TATA box. The binding of TFIIB over the initiation site downstream of the TATA box is the next event in the assembly of the preinitiation complex. These factors are targets for *trans*-activators and promote RNA polymerase II recognition of the promoter. Binding of this enzyme occurs in conjunction with the binding of the RNA polymerase associated factors, TFIIE and TFIIF. An additional factor, TFII-I, can functionally replace TFIIA on TATA box containing promoters. This factor complex binds downstream of TFIIB at an initiator element with the consensus sequence YAYTCYYY (Y represents pyrimidine). At TATA-less promoters binding of TFII-I is required for the association of TFIID with the core promoter (Roy *et al.*, 1991).

In theory, any step in the formation of the preinitiation complex could be modified by the interaction of *trans*-acting proteins (Ham *et al.*, 1992). Basal transcription is observed *in vitro*, however it is undetectable *in vivo*. This is presumably due to repression by chromatin packaging and/or the presence of negative factors that also interact with the transcription machinery (Hahn, 1993). Significantly, the depletion of nucleosomes *in vivo* can dramatically increase basal transcription (Durrin *et al.*, 1992). The interaction of activators or their mediators with particular proteins in the transcription complex may compete with these repressors (Felsenfeld, 1992). Additionally, the direct or indirect contact of the *trans*-activation domains of *trans*-acting proteins with RNA polymerase II or any of the general transcription factors could alter the rate of formation or activity of the preinitiation complex (Ham *et al.*, 1992; Herschlag and Johnson, 1993).

Significantly, both TFIID and TFIIB have been implicated as targets for activator proteins (Ptashne and Gann, 1990; Lin and Green, 1991; Hahn, 1993). As TBP does not allow the equivalent activation as TFIID it has also been suggested that the factors associated with TBP could bridge the interaction between *trans*-acting proteins and this general transcription factor (Pugh and Tjian, 1990; Hoey *et al.*, 1993).

Transcriptional Regulators

Impressive progress has been made into the identification of *cis*-acting DNA elements and their associated *trans*-acting proteins. Analysis of transcription from selectively mutated DNA that has been transiently introduced into cells and the use of *in vitro* transcription assays have been essential to characterizing the components involved in transcriptional regulation. Additionally, the development of *in vitro* DNA binding assays, such as the electrophoretic mobility shift and the nuclease footprinting assays, have been implemental in identifying regulatory DNA sequences. Transient transfection assays are used to localize regulatory subregions, while *in vitro* footprinting aids in the characterization of the exact DNA binding site. Concurrently, the development of sequence-specific DNA affinity chromatography has enabled the purification of many *trans*-acting proteins. Structure-function studies of these proteins are made possible by deletion and mutation analysis coupled with the co-transfection of a vector containing the *cis*-acting element. Together these studies have determined that some *trans*-acting factors are active only when their *cis*-acting elements are located close to the promoter, within a few hundred base pairs, while others can exert their effects from many thousands of base pairs away (Serfling *et al.*, 1985; Mitchell and Tjian, 1989; Seipel *et al.*, 1992). Additionally, some of these factors are able to interact directly with the

transcriptional machinery, while others require the presence of additional proteins to assist in the *trans*-activation process (Ptashne and Gann, 1990; Martin, 1991).

Cis -acting Elements

The existence of multiple proteins which recognize related sequences may be reflected in greater precision and flexibility in the regulation of gene expression. This could be achieved by subtle differences in sequence recognition, DNA-binding affinities and dimerization partners (Strühl, 1989). Dimerization is one method for determining differences in gene regulation by the Fos/Jun family of proteins. Homo and heterodimers of various members of this family are able to bind a common sequence usually known as an AP1 site (Strühl, 1989). Differential expression of these factors further dictates their regulatory potential. This has also been observed for the Myc family of proteins which form homo or heterodimers to recognize a methylated or nonmethylated DNA sequence (Prendergast and Ziff, 1991; Prendergast *et al.*, 1991; Suetake *et al.*, 1993).

Conversely, the differential affinity of Sp1 for multiple related sites may be one mechanism by which this *trans*-acting protein modifies its effect on transcription (Kadonaga *et al.*, 1986). The binding of proteins such as GCF-1(Kageyama *et al.*, 1988), which recognizes a guanine and cytosine (G+C) rich site that overlaps with high affinity Sp1 sites, and the interaction of other Sp1-like proteins may compete with or modulate Sp1 activity. Similarly, there are multiple members of the NF1/CTF family which bind CCAAT-related sequences, possibly with differing affinities (Jones *et al.*, 1987). For the POU domain proteins the ubiquitous factor Oct-1 and the lymphoid specific factor Oct-2 bind the same *cis*-acting element (Staudt *et al.*, 1986), while the protein Pit-1 recognizes a similar A+T rich binding site that

differs by only two bases (Elsholtz *et al.*, 1990). The differential affinities of these proteins for their related binding sites is reflected in the regulation of pituitary-specific gene expression by Pit-1 and lymphoid-specific gene expression by Oct-2 (Rosenfeld, 1991).

The primary structure of the DNA-binding site is also a mechanism for dictating binding specificity in the steroid hormone receptor family. Although the DNA binding sites or HREs for these receptors have diverse sequences *in vivo*, there is considerable sequence similarity in their consensus sequences (Schwabe and Rhodes, 1991). In particular, the consensus sequences for the ER and GR have identical half-site spacing but differ in the sequence of each half site (Figure 3A). Conversely, the half-sites recognized by the ER and T3R are identical but the spacing between the half-sites differs (Figures 3 and 4). The spacing between the half-sites of the estrogen and glucocorticoid response elements (ERE and GRE, respectively) results in their location on the same face of the DNA, while the spacing between the two T3R response element (TRE) half-sites places them on opposite faces of the DNA (Schwabe and Rhodes, 1991). Receptors for thyroid hormone can bind to both TREs and EREs with equal affinity, but EREs do not necessarily confer thyroid hormone-dependent gene activation (Glass *et al.*, 1988). Furthermore, the RAR and T3R are able to recognize direct repeats of two half-sites in addition to palindromic repeats (Schwabe and Rhodes, 1991; Näär *et al.*, 1991; Forman *et al.*, 1992).

Trans-acting Factors

Activator proteins are modular in structure (Ptashne, 1988; Frankel and Kim, 1991), having DNA binding and *trans*-activation domains (Klug and Rhodes, 1987; Mitchell and Tjian, 1989; Rosenfeld, 1991; Schwabe and Rhodes, 1991; Treisman *et al.*, 1992), often accompanied with dimerization

motifs (Landschulz *et al.*, 1988; Vinson *et al.*, 1989). DNA binding domains direct the *trans*-acting factor to the vicinity of the promoter, dimerization domains form active proteins and *trans*-activation domains exert the functional role of the activator on the initiation of transcription. The function of these activating regions may be facilitated by the association of various types of effector proteins (Ptashne and Gann, 1990; Martin, 1991).

A number of different mechanisms have been hypothesized for facilitating the activities of *trans*-acting and effector proteins. Direct or indirect interactions of *trans*-acting proteins with the transcription initiation complex has been suggested to aid in the recruitment of additional basal or activating transcription factors. These interactions may modify the preinitiation complex into an active transcription complex. Additionally, the function of some *trans*-acting and co-activating proteins may be to alter the chromatin environment of the promoter by clearing this region of nucleosomal proteins (Felsenfeld, 1992). Each additional component required for creating a functional transcription initiation complex is suggested to provide additional levels of control over gene expression (Martin, 1991). As well, each activator may have multiple targets that differ depending on the promoter context (Ham *et al.*, 1992).

Increased complexity can occur by the multimerization of functionally distinct members of the same *trans*-acting protein family. Additionally, the cooperative or synergistic interactions of multiple *trans*-acting factors would increase the rate of transcription initiation by a greater amount than the sum of the activators individually (Ptashne, 1988). Although the intrinsic affinity of activators and targets may be low, synergistic interactions could stabilize these interactions. Stronger activators would interact more tightly with their target and would therefore be able to function over further distances from the

promoter (Ptashne, 1988). Interactions between factors bound at separated DNA elements would necessitate the looping out of the intervening DNA (Ptashne, 1986).

DNA Binding Domains

The DNA binding domain of a *trans*-acting protein has at least two important functions (Strühl, 1989). First it directs the highly specific interaction of the protein with the DNA thus providing an important mechanism for differential gene regulation. Secondly, this domain positions the protein in a manner which facilitates the interaction of *trans*-activating regions with other regulatory proteins and/or the components of the transcription machinery. A number of these domains have already been characterized (Sturm and Herr, 1988; Mitchell and Tjian, 1989; Strühl, 1989; Rosenfeld, 1991; Schwabe and Rhodes, 1991; Treisman *et al.*, 1992;).

The mammalian helix-turn-helix DNA binding motif is more commonly known as the homeodomain (Treisman *et al.*, 1992). It consists of approximately 60 amino acids which contain three or four α-helical stretches separated by ß-turns. The recognition helix consists of 20 amino acids and lies perpendicular to the first two antiparallel helices. It is the recognition helix and the second helix which form the helix-turn-helix structure. The second helix is believed to hold the recognition helix in place, allowing the carboxy terminal end to fit into the major groove of the DNA binding site. One subclass within the homeodomain containing family is the POU-domain transcription factors (Rosenfeld, 1991). These factors contain two major regions of very high homology, the POU-specific (POU$_S$) region of an additional 69 to 78 amino acids spaced 14 to 25 amino acids amino terminal to the POU-homeodomain (POU$_{HD}$). The POU-specific domain can be further subdivided into two α-helices, POU$_S$-A and POU$_S$-B. Mutagenesis

studies suggest that both the POU_S and POU_{HD} are required to permit high-affinity, site-specific DNA binding (Sturm and Herr, 1988).

Some DNA binding proteins form dimers through the interaction of leucine residues. Originally called the leucine zipper (Landschulz *et al.*, 1988), this dimerization motif forms a two-stranded, parallel α-helical coiled coil (O'Shea *et al.*, 1989). The periodicity of the leucine residues places them on the same side of the α-helix allowing for the interaction of the equivalent region of the second monomer. *Trans*-acting factors containing this dimerization domain fall into two classes; those which have a DNA binding domain consisting of basic residues and those which combine a basic region with a helix-loop-helix region for their DNA binding properties. The proteins Jun and Fos, which dimerize to bind AP1 sites, are two proteins which bind DNA through a basic region (Landschulz *et al.*, 1988; Vinson *et al.*, 1989). The proteins Myc and Max are two proteins which utilize both a basic region and a helix-loop-helix domain for DNA recognition (Dang *et al.*, 1989; Murré *et al.*, 1989a; Blackwood and Eisenman, 1991; Prendergast and Ziff, 1991; Kato *et al.*, 1992a). For both of these classes of leucine zipper proteins the dimerization domain allows the juxtapositioning of the two monomeric DNA-binding regions to form a functional DNA-binding domain.

The zinc finger motif was first identified in the RNA polymerase III transcription factor TFIIIA (Miller *et al.*, 1985). Central to this motif is the tetrahedral coordination of Zn^{2+} ions, which stabilize each of the nine 30 amino acid domains. Additionally, there are a number of conserved hydrophobic residues which may form an inner core in the tertiary structure of the zinc finger (Klug and Rhodes, 1987). The original structural model speculated that the amino acids between the coordination sites would project out as fingers, hence the term zinc finger. A more recent structural model of

TFIIIA consists of an antiparallel ß-sheet, containing one pair of coordinating residues and a short α-helix which encompasses the second pair of coordinating residues (Berg, 1988). The coordination of the zinc ion by two cysteines and two histidines maintains the overall structure and allows the α-helix to interact with the DNA (Frankel and Kim, 1991).

The structure of other zinc containing DNA-binding motifs which have been described is based on the original zinc finger motif. The DNA-binding domain of the *trans*-acting factor, Sp1 (Kadonaga *et al.*, 1988), consists of three zinc finger motifs which utilize two cysteine and two histidine residues (C_2-H_2). The zinc cluster motif has been used to describe the coordination of two zinc atoms with six cysteine residues, as seen in the transcription factor GAL4 (Vallee *et al.*, 1991). Alternatively, these same authors describe the coordination of zinc by four cysteines (C_2-C_2), as seen in the steroid hormone receptors, as a zinc twist.

All of the steroid hormone receptors have only two coordinated zinc structures, but the spacing between the cysteines is similar to that seen in TFIIIA. The sequence differences between the zinc finger motifs of these receptors and TFIIIA include the absence of the conserved hydrophobic residues and a larger spacing between the two motifs. However, despite similarities in the primary structure of the C_2-H_2 and C_2-C_2 zinc finger motifs, the secondary and tertiary structures are quite different (reviewed in Schwabe and Rhodes, 1991). The anti-parallel ß-sheet is absent from the steroid hormone receptor zinc finger structures and the α-helical region extends beyond the coordinating residues rather than between them, as in the C_2-H_2 form. Additionally, the two fingers of the steroid hormone receptors fold together to form a single structural domain. Both fingers consist of an extended loop between the two pairs of metal-binding cysteines,

followed by an α-helix involved in DNA recognition (Berg, 1988). The helices are highly amphipathic and fold together such that they cross at right angles near the midpoint (Schwabe and Rhodes, 1991). The residues 203-212 and 239-251 of the ER were recently suggested through x-ray crystallography studies of the GR (Luisi *et al.*, 1991) and ER (Schwabe *et al.*, 1993) to form helical structures within each zinc finger. The discriminating nature of the residues involved in DNA recognition were also addressed in these studies.

The various amino acid residues, which are responsible for discriminating between different HRE orientations, have been identified by mutagenesis studies of several receptors (reviewed in Schwabe and Rhodes, 1991). In the estrogen receptor, glutamic acid 203, glycine 204 and alanine 207 (Figure 7A) of the first zinc-binding motif are located on the same face of the recognition helix (Schwabe and Rhodes, 1991; Schwabe *et al.*, 1993). These residues are sufficient for differential recognition of the ERE and GRE half-site sequences (Mader *et al.*, 1989) and were determined to interact with the HRE through the major groove (Luisi *et al.*, 1991). Significantly, both the glutamic acid and glycine residues are conserved between the ER, RAR, T_3R and D_3R, while the PR, GR and receptors for androgen (AR) and mineralcorticoid (MR) have conserved glycine and serine residues in these positions (Figure 7B).

Located along the same face of the recognition helix are three basic amino acids (ER; lysines 206 and 210, plus arginine 211) which are conserved between all of the receptors (Figure 7B). These residues were shown to interact with DNA through the acidic phosphate backbone of the minor groove (Luisi *et al.*, 1991). A further five residues, proline 221 to glutamine 226, in the second ER zinc-binding motif are involved in recognizing differences in half-site spacing (Umesono and Evans, 1989). These residues

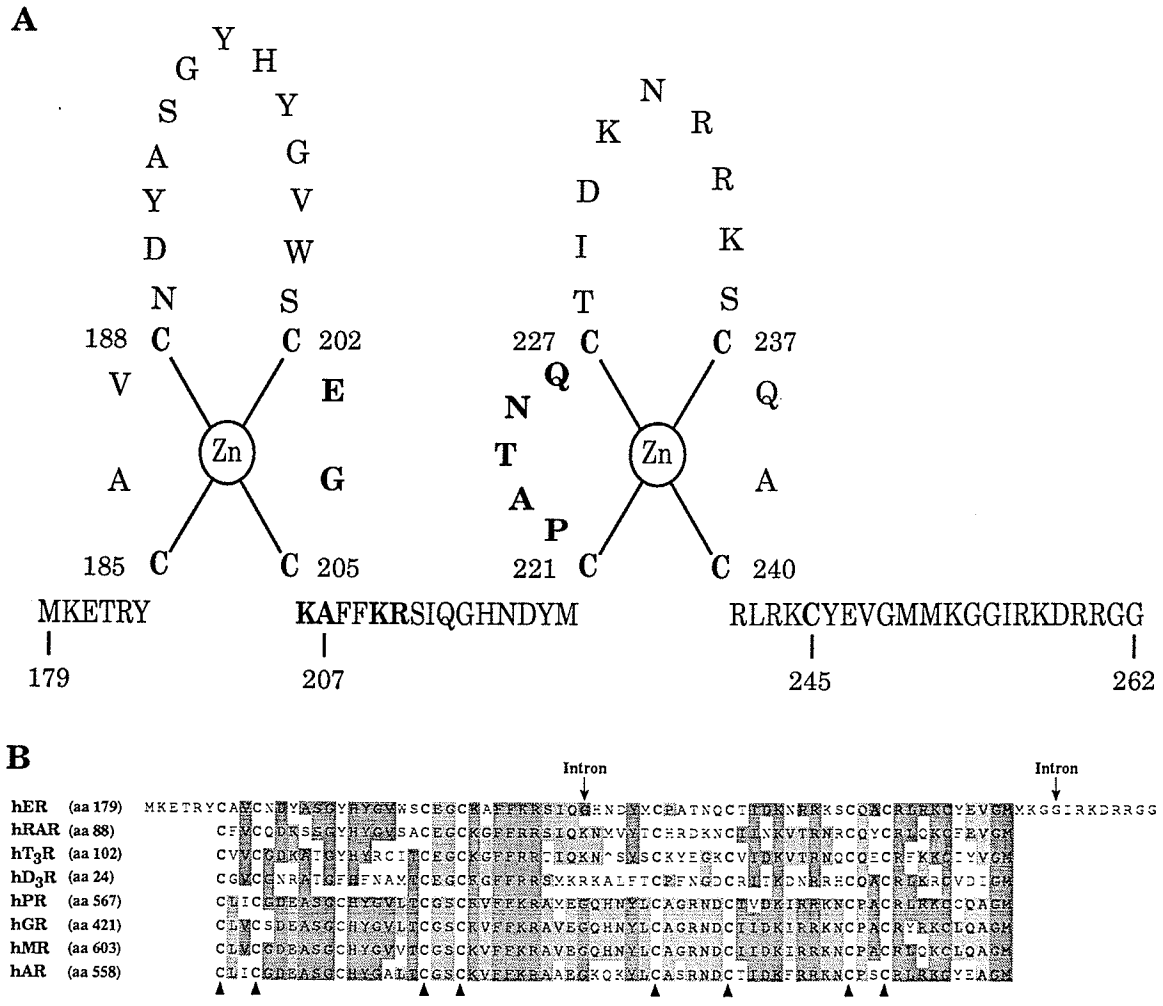


Figure 7. The ER DNA Binding Domain. A. The structure of the ER DNA binding domain is shown in the classical "zinc finger" conformation, although amino acids 203 to 212 and 239 to 251 are now believed to form α helices. Cysteines 185, 188, 202, 205 and 221, 227, 237, 240 are involved in zinc coordination in the first and second "zinc fingers", respectively. Amino acids 203, 204 and 207 are necessary for half site sequence recognition, amino acids 221 to 226 are involved in half site spacing recognition and amino acids 206, 210 and 211 interact with the acidic phosphate backbone of DNA. **B.** The sequence of the ER DNA binding domain is compared to the other members of the steroid and thyroid hormone receptor family. The dark and light shading of the amino acids represent sequence identity between all of the receptors and within the two subgroups of this receptor family, respectively. The conserved cysteines are marked with arrowheads. In the ER exons 2 and 3 are separated by an intron within the codon for glycine 215, while exons 3 and 4 are separated by an intron within the codon for glycine 254. Both introns are indicated with arrows. The one letter code for the amino acids is given in the list of abbreviations. This figure was reproduced from Schwabe and Rhodes, 1991; Mader *et al.*, 1989; Evans, 1988.

are not conserved between the receptors and appear to be located in a loop on the surface of the ER (Schwabe and Rhodes, 1991; Schwabe *et al.*, 1993). Interestingly, the analogous region of the GR, PR, MR and AR is highly conserved (Figure 7B), suggesting a common motif which is not conserved in the ER, RAR, T_3R and D_3R.

Trans-activation Domains

Trans-activating proteins usually have more than one region of amino acids that contribute to transcriptional activation (Mitchell and Tjian, 1989). Two common features, which have been described, are the presence of considerable negative charge and the ability to form an α-helical structure. The *trans*-activation sequences which have been specifically characterized include; acidic domains (Ptashne, 1988; Mitchell and Tjian, 1989; Hahn, 1993), glutamine rich domains (Courey *et al.*, 1989; Mitchell and Tjian, 1989), proline rich domains (Mermod *et al.*, 1989; Mitchell and Tjian, 1989) and serine plus threonine rich domains (Rosenfeld, 1991). Recently, proteins containing these activation domains have been subdivided on the basis of their ability to *trans*-activate from remote or proximal promoter sites (Seipel *et al.*, 1992). These studies classified those factors containing glutamine rich sequences as promoter proximal activators, which could only activate transcription from a position close to the TATA box. Acidic activators and those activators with serine plus threonine rich sequences were found to activate transcription from both promoter proximal and distal enhancer sites, and were subsequently classified as general activators. Factors containing proline rich domains were identified as a third class of activators with considerable promoter activity and low but significant enhancer activity.

Acidic activators include the viral protein VP16 and the *trans*-activator GAL4 (Ptashne and Gann, 1990). Their acidic regions are able to form

amphipathic α-helical structures (Mitchell and Tjian, 1989). Multiple acidic domains have been shown to result in functionally redundant regions in spite of an obvious lack of sequence similarity (Ptashne, 1988). Four regions of Sp1 have been identified by deletion analysis to display *trans*-activating function (Kadonaga *et al.*, 1988; Courey *et al.*, 1989). These include a region with a net positive charge, a region near the carboxy terminus and two regions containing 25% glutamines. Regions rich in glutamines have also been described for the factors OCT-1, OCT-2, Jun, and AP2 (Mitchell and Tjian, 1989). The CTF/NF1 factors use a proline rich domain for *trans*-activating function (Mermod *et al.*, 1989). The carboxy-terminal region contains 20-30% prolines which are required for transcriptional activation but not for stimulating replication. The factors AP2, Jun and Oct2 also have regions which are high in prolines (Mitchell and Tjian, 1989). A number of regions of homology exist within the POU domain proteins, including serine and threonine rich, glutamine rich and glycine and alanine rich (Rosenfeld, 1991). Of these only the region containing high levels of serine and threonine residues has been shown to have *trans*-activating function (Theill *et al.*, 1989). This region is located outside of the POU domain and confers either positive or negative regulatory function (Rosenfeld, 1991).

Two *trans*-activating regions have been described for the steroid hormone receptors (Kumar *et al.*, 1987; Webster *et al.*, 1988; Hollenberg and Evans, 1988; Tora *et al.*, 1989a). One of these (TAF-2) resides within the hormone binding domain (HBD) and does not require the A/B or C regions for *trans*-activation function (Figure 2). Of the five exons encoding the HBD, no single exon encoded this *trans*-activation function for the ER (Webster *et al.*, 1989). Conversely, in the GR this *trans*-activation function was localized to a 30 amino acid region rich in acidic residues, (Hollenberg and Evans, 1988).

An additional *trans*-activating domain (TAF-1) was localized to a second acidic region in the A/B domain of the GR (Webster *et al.*, 1989). A similar acidic region is lacking in the ER, however the A/B domain contains *trans*-activation activity (TAF-1) which cooperates with the HBD *trans*-activating region (TAF-2) to fully activate transcription (Kumar *et al.*, 1987). The identification of acidic residues for the *trans*-activation function of the GR suggests some relationship with other acidic activators. This is partially supported by the ability of acidic activators to inhibit transcription directed by GR but not ER activation regions (Tasset *et al.*, 1990). Conversely, the GR activation regions were unable to inhibit transcription directed by other acidic activators while the opposite was seen with the ER. Together these studies suggest that the target of the steroid hormone receptor *trans*-activation domains may differ from the acidic activators (Martin, 1991). Concurrently, additional *in vitro* competition experiments suggest that the *trans*-activation domains of the ER, GR and T_3R may interact with the same target (Meyer *et al.*, 1989). Whether these domains interact directly or indirectly with the general transcription complex or have some other function remains undefined, but is predicted to be unique (Martin, 1991).

Effector Proteins

A number of effector proteins (Figure 6) have been described to bridge or assist the interactions between *trans*-acting factors and the components of the transcriptional machinery (Ptashne and Gann, 1990; Martin, 1991). The terms co-activator, adapter and accessory protein refer to effector proteins defined on the basis of apparently different functions (Martin, 1991). Co-activators are defined as proteins which do not bind directly to DNA but act with another transcriptional activator to directly interact with the general

transcription factors. Adaptors refer to proteins with *trans*-activating domains which interact with DNA bound factors and facilitate their association with the general transcription factors. Accessory proteins do not directly interact with the general transcription machinery, but are required to indirectly promote transcription activation. Transcriptional activation therefore involves the assembly of numerous different proteins to form a complex protein structure which acts synergistically to stimulate transcription (Lin *et al.*, 1990).

VP16 is an example of a transcriptional adapter which does not bind DNA but has its own *trans*-activation domain. It has been shown to facilitate the activity of the octamer binding protein (Oct-1) and increase the rate of recruitment of TFIIB partly by directly contacting TFIID and TFIIB (Lin and Green, 1991). The transcriptional activator E1a interacts with the sequence specific DNA-binding protein ATF-2. Unlike VP16, E1a is unable to directly facilitate the actions of this *trans*-acting factor. An additional protein factor is required, suggesting that the ATF-2/E1a complex acts through an adaptor or co-activator to stimulate transcription (Martin *et al.*, 1990). Although the transcriptional activators Sp1 and CTF/NF1 contain both *trans*-activation and DNA-binding domains, strong stimulation of transcription is only observed in the presence of TFIID and not with purified TBP (Pugh and Tjian, 1990). This suggests that these *trans*-acting factors require the effector proteins which normally fractionate with TBP for their transcriptional activating functions. Additionally, the interaction of Sp1 with these TBP associated factors has been suggested to be necessary for the stabilization of TBP at a TATA-less promoter (Hoey *et al.*, 1993).

Transcription Factor Interactions

The interactions of *trans*-activators with their DNA elements occurs with varying affinities and specificities, but it is probably specificity that is important for the assembly of functional transcription complexes (reviewed in Frankel and Kim, 1991). Moderate specificity can be overcome by the cooperative interactions of multiple *trans*-acting proteins which assure an overall specificity which is high. Additionally, the use of multi-factorial complexes allows for the flexibility required for differential gene responses to a limited number of transcriptional regulators. Synergism in transcriptional activation is a widely accepted phenomenon and is defined as a more than additive effect of factors bound to multiple transcriptional elements (reviewed in Herschlag and Johnson, 1993). The observation of transcriptional synergy suggests that the individual regulatory factors are influencing the same or converging pathways towards transcript initiation and are therefore functionally dependent activities (Figure 8). This mechanism allows for large responses to small differences in activator levels.

Classical models for transcriptional synergism are based on equilibrium binding interactions (Ptashne, 1988; Ptashne and Gann, 1990; Lin *et al.*, 1990; Herschlag and Johnson, 1993). In general, these are defined by the cooperative binding of two *trans*-activators to their *cis*-acting elements through protein interactions between the two activators (Ptashne, 1988; Ptashne and Gann, 1990), or the cooperative interaction of two activators independently bound to their respective elements through the interactions with a third target protein (Lin *et al.*, 1990). A further classification of equilibrium binding models on the basis of energetic synergy has been described (Herschlag and Johnson, 1993). Positive cooperativity is defined by the increased interaction of one *trans*-activator with the transcriptional

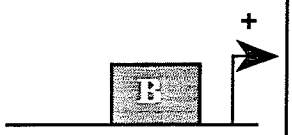
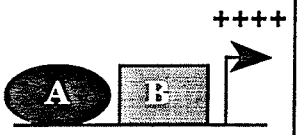
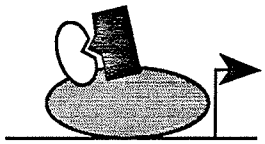
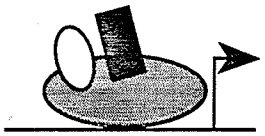
Activation	Physical Description	Simple Model
Independent	Separate Pathways Additive Transcription	 
Synergistic	Same Pathway Greater than Additive Transcription	
Classification	**Energetic Description**	**Physical Model**
Positive Cooperativity	Greater than Additive in Free Energy	
Energetic Independence	Additive in Free Energy	
Negative Cooperativity	Less than Additive in Free Energy	

Figure 8. Models for Transcriptional Synergy. Transcriptional synergy is an indication that two or more transcriptional regulators are acting on the same or merging pathways, rather than independent pathways, to stimulate transcription. This is diagrammed in the top portion of the figure. Traditionally, these interactions were classified as cooperative interactions and were characterized on the basis of equilibrium binding. These synergistic interactions can be further classified on the basis of different types of cooperativity which effect the energetics of factor binding. Positive and negative cooperativity reflect the enhancement and disruption of factor binding, respectively. Conversely, energetic independence suggests that the two factors do not in any way affect the affinity of each others binding, although they could still interact with the same target and along the same transcriptional pathway. This figure was reproduced from Herschlag and Johnson, 1993.

machinery by the direct or indirect interaction with a second transcriptional regulator (Figure 8). Conversely, energetic independence is characterized by the functional dependence of two *trans*-activators which interact with a common target but do not influence each other's binding affinity. Similarly, negative cooperativity suggests functional dependence, but in the presence of activator antagonism. Additional models of transcriptional synergy have been proposed on the basis of kinetics and are characterized by the enhancement of rate limiting activities, the control of converting a nonproductive transcription complex into a productive one and the regulation of post-translational modifications required for transcription initiation (Herschlag and Johnson, 1993).

Cooperative interactions of *trans*-acting proteins at multiple response elements was one of the first descriptions of transcriptional synergy. This phenomenon has been observed for the *trans*-acting factors Sp1 (Courey *et al.*, 1989; Pascal and Tjian, 1991; Anderson and Freytag, 1991), and many of the steroid hormone receptors (Schüle *et al.*, 1988a; Klein-Hitpass *et al.*, 1988b; Tsai *et al.*, 1989; Kato *et al.*, 1992b). Although considered a promoter proximal *trans*-acting factor (Seipel *et al.*, 1992), transcriptional synergy has been observed for Sp1 bound to distal and promoter proximal sites (Courey *et al.*, 1989). The interaction of Sp1 proteins bound at multiple sites has been shown to require the integrity of its glutamine rich *trans*-activation sequences (Pascal and Tjian, 1991). These regions facilitate the formation of Sp1 tetramers at a single element and together with a carboxy-terminal *trans*-activation domain mediate the formation of multimeric complexes between tetramers bound at separate elements. Additionally, the affinities of the various Sp1 binding sites directs the relative contribution of each site to full transcriptional activation (Anderson and Freytag, 1991).

The cooperative binding of steroid hormone receptors to tandemly linked response elements has been characterized for the GR (Jantzen *et al.*, 1987; Schüle *et al.*, 1988a; Strähle *et al.*, 1988), PR (Schüle *et al.*, 1988a; Tsai *et al.*, 1989) and ER (Klein-Hitpass *et al.*, 1988b; Ponglikitmongkol *et al.*, 1990). The DNA binding domains of a number of steroid receptors have been shown to interact first with the downstream half-site of the HRE (Tsai *et al.*, 1988; Ponglikitmongkol *et al.*, 1990). This binding subsequently facilitated the association of the receptor with the upstream half-site of the response element. Additionally, the interaction of two receptors with more than one response element results in an activation which is greatly enhanced over that seen with only one binding site. Mutation of the GRE/PRE disrupts binding of the receptor, however, binding to an adjacent functional site facilitates receptor binding to the altered element (Jantzen *et al.*, 1987; Tsai *et al.*, 1989). The interaction of the ER with a strong ERE is capable of activating expression alone and only additive enhancement is observed with binding to a second site. Conversely, the presence of a weak ERE is incapable of activating transcription, but synergistic activation is observed when ER interacts with two such elements (Klein-Hitpass *et al.*, 1988b; Martinez and Wahli, 1989). Similarly, the interaction of the ER with multiple copies of ERE half-sites results in synergistic activation (Kato *et al.*, 1992b).

Transcriptional synergy has also been observed between factors which bind different *cis*-acting elements. Several *trans*-acting factors have been shown to contribute to transcriptional synergy with the PR or GR bound at adjacent sites, including NF1, CP1, Sp1, Oct-1, CACCC binding factor, and CCAAT binding factor (Schüle *et al.*, 1988a; 1988b; Strähle *et al.*, 1988). The degree of synergism between the GR and other *trans*-activators was reported to be inversely related to the strength of the GRE (Schüle *et al.*, 1988a).

Additionally, this activity required the GRE to be positioned in close lateral proximity to the *cis*-acting sequence of the cooperating *trans*-activator (Schüle *et al.*, 1988a; 1988b; Strähle *et al.*, 1988). Synergistic interactions between the ER and CTF/NF1 (Martinez *et al.*, 1991) have also been identified. The mechanism for synergy between the ER and CTF/NF1 *trans*-activating domains appears to be mediated through a common limiting factor (Martinez *et al.*, 1991).

Transcriptional antagonism has also been characterized between different *trans*-acting proteins. For steroid receptors these types of interactions have been categorized into three groups of effects based on the observed phenomena (Shemshedini *et al.*, 1991). Hormonal activation has been demonstrated to occur at colocalized binding sites, in the absence (Gaub *et al.*, 1990) and presence (Diamond *et al.*, 1990) of HRE binding, and mutual transcriptional inhibition between these same factors has also been described (Yang-Yen *et al.*, 1990). The interference of the GR with Oct-1 transcriptional activity does not require GRE binding and is mediated by direct protein interactions through the Oct-1 homeodomain (Kutoh *et al.*, 1992). Similarly, functional antagonism between the GR and AP-1 factors has been proposed to occur through protein-protein interactions at AP-1 sites in the absence of DNA-binding by the GR (Yang-Yen *et al.*, 1990). Conversely, composite elements, which bind both the GR and AP-1 factors, are involved in the regulation of positive or negative activity by the GR (Sakai *et al.*, 1988; Diamond *et al.*, 1990). Although originally thought to be a negative GRE (Sakai *et al.*, 1988) subsequent studies have suggested that the relative proportion of c-Fos and c-Jun found in the AP-1 complexes was an important modulator of GR activity (Diamond *et al.*, 1990). In particular, high levels of c-Fos conferred a negative GR response, while high levels of

c-Jun resulted in positive GR activity and the absence of c-Jun made the element inactive. Suppression of the ER transcriptional activity by complexes containing c-Jun, c-Fos and to a lesser extent Jun-B but not Jun-D, has been observed (Doucas *et al.*, 1991). The repression by c-Fos was suggested to be through an interaction with ER bound complexes, while the suppressive action of c-Jun was proposed to be through a glycine rich region adjacent to the c-Jun DNA-binding domain. Significantly, the imperfect ERE found in the regulatory regions of the c-*fos* gene has been shown to bind ER and AP-1 factors (Weisz and Rosales, 1990). In contrast, the mechanism of transcriptional activation by the ER at ERE half-sites does not require the DNA-binding domain and is facilitated by complexes of c-Fos and c-Jun (Gaub *et al.*, 1990). Functional interactions between AP-1 proteins and the RAR, T_3R, PR and AR have also been demonstrated (Schüle *et al.*, 1991; Zhang *et al.*, 1991; Shemshedini *et al.*, 1991). Alternatively, alterations in protein concentration and post-translational modifications of the Myc and Max proteins have also been demonstrated to regulate activation or repression of gene activity (Kretzner *et al.*, 1992; Berberich and Cole, 1992; Amin *et al.*, 1993; Gupta, *et al.*, 1993).

The Role of Transcriptional Regulators

It is widely believed that activators function to facilitate some rate limiting step in the initiation of transcription (Hahn, 1993). One possible mechanism is that the interaction of *trans*-acting factors with the general transcription machinery competes with proteins involved in repressing transcription. By preventing these associations activators would increase the rate of transcription. An alternative but not mutually exclusive model suggests that the general transcription factors assemble into nonproductive complexes at the promoter and require the interaction of *trans*-activators to

form a productive preinitiation complex (Herschlag and Johnson, 1993). Significantly, complete assembly and disassembly of the transcription initiation complex for each initiation event appears to occur in the absence of *trans*-acting factors (Kadonaga, 1990). This is consistent with a role for these factors in the maintenance of the transcriptionally active state. Additionally, the formation of an active preinitiation complex has been shown to require the prebinding of the activator as well as TFIID, TFIIA and TFIIB prior to the addition of polymerase and the remaining general factors (Wang *et al.*, 1992).

Transcriptional regulation is therefore classically viewed as the interaction of soluble *trans*-acting factors with soluble transcription initiation factors which are all bound to a fixed DNA molecule. This misconception prevails because soluble nuclear proteins are active in assays designed to measure the contributions of DNA elements to transcriptional activity. Although this model for transcriptional regulation is incomplete and somewhat flawed, studies using soluble components have contributed greatly to the characterization of *trans*-activators, their *cis*-acting elements and the components of the preinitiation complex. Additionally, this research has been instrumental in characterizing the interactions between these transcriptional regulators and the identification of various effector proteins. This type of research has been supportive of the concept that multiple *trans*-acting factor interactions with each other and regulatory regions of genes contribute to the overall level of transcriptional activation. However, the discovery that sites for well known *trans*-acting factors are found in LCRs and that some of these factors have been isolated from nuclear matrix preparations suggests that the traditional view of transcriptional regulators needs to be modified. Renewed interest in the role of DNA packaging has defined a more complex model for

transcriptional activation which suggests that one role of *trans*-activators is to overcome the repression of chromatin structure and assist in the formation of stable transcription initiation complexes.

Transcriptionally Active Chromatin

The chromatin of inert genes that are not transcribed is extensively packaged into nucleosomal structures. Conversely, the chromatin of active genes, which are engaged in transcription or can be stimulated to initiate transcription, is packaged into nucleosomal structures that are interspersed with non-histone proteins interacting at regulatory and promoter regions (Elgin, 1988; Gross and Garrard, 1988). These interruptions in nucleosomal arrays are detected as nuclease hypersensitive regions. These hypersensitive regions lie within DNase I sensitive chromatin domains and encompass 50 to 400 bp of DNA (Elgin, 1988; Bonifer *et al.*, 1991). DNase I hypersensitive (DH) sites represent a minor (1%) but highly selective fraction of the genome. They are believed to result from the movement of nucleosomal proteins from discrete sites on chromatin which subsequently alters the sensitivity of the DNA to the nuclease by at least two orders of magnitude (Gross and Garrard, 1988). The position of DH regions have been located at or near sequences determined to be important for gene expression. This correlates nuclease hypersensitivity with genetic regulatory loci and binding sites for sequence-specific DNA-binding proteins (Yaniv and Cereghini, 1986; Elgin, 1988; Gross and Garrard).

Some DH sites can be modulated by the presence or absence of an inducer, while others have been correlated with cell-type and cell-stage specific expression (Elgin, 1988; Gross and Garrard, 1988; Bonifer *et al.*, 1991). Constitutive sites are present in genes which are poised for

transcriptional induction. Formation of these sites is independent of gene expression and precedes transcriptional activation (Elgin, 1988; Gross and Garrard, 1988). Protein interactions at these sites do not need to be reestablished with each transcriptional event, but persist as part of the mechanism for maintaining transcriptional competence (Goldman, 1988). Conversely, inducible sites appear as the result of stimulation by an inducing agent and may persist after its removal. These sites are associated with a commitment to transcriptional activation (Gross and Garrard, 1988). Tissue-specific DH sites are a reflection of the potential of a gene to be expressed in certain cell types while stage-specific DH sites appear as a result of developmental and differentiative processes. In all cases, the formation of DH sites appears to precede or accompany gene expression (Elgin, 1988; Gross and Garrard, 1988; Bonifer *et al.*, 1991).

Hypersensitive regions possess multiple internal "cold" and "hot" spots which reflect the presence or absence of bound *trans*-acting proteins, respectively (Gross and Garrard, 1988). Although the underlying DNA sequence is necessary for directing the binding of these factors, the association of *cis*-acting sites with hypersensitive chromatin regions is not sufficient for their functional activity (Gross and Garrard, 1988). The observation that DH sites may persist even when mutations disrupt *cis*-acting elements suggests that particular DNA-binding sites are responsible for promoting the establishment of a DH region (Lee and Garrard, 1992). It has therefore been hypothesized that the role of some *trans*-acting factors is to compete with the packaging of chromatin into nucleosomal structures (Felsenfeld, 1992). This is supported by the finding that the reconstitution of promoter templates with nucleosomes and H1 is

critical for the observation of large responses to activator proteins (Laybourn and Kadonaga, 1991; Workman and Buchman, 1993).

Histones as Repressors of Transcription

The properties of stabilizing the packaging of DNA into nucleosomes and facilitating higher order chromatin structures makes histone H1 an obvious candidate for the generalized repression of chromatin. Additionally, specific nucleosome positioning can alter the accessibility of regulatory sequences to the transcriptional machinery and *trans*-acting factors. An initiated RNA polymerase is able to transcribe through nucleosomal structures, however the formation of nucleosomes over the promoter has been shown to block transcriptional initiation (reviewed in Felsenfeld, 1992; Workman and Buchman, 1993). The assembly of core nucleosomes over the site for transcriptional initiation results in a reduction of basal transcription to 25-50% of that seen with naked DNA (Laybourn and Kadonaga, 1991). This nucleosome mediated repression could not be counteracted by the addition of transcriptional activators. The incorporation of histone H1 further reduced transcription initiation to 1-4% of that seen with reconstituted nucleosomal cores (Figure 9). Transcriptional activators, such as Sp1, were able to prevent H1 mediated repression in the presence (Laybourn and Kadonaga, 1991) and absence of nucleosomes (Croston *et al.*, 1991). Similarly, the selective removal of H1 and nucleosome loss both result in derepression and transcriptional activation (Grunstein, 1990; Garrard, 1991).

One model suggests that the transcription complex and *trans*-acting factors may compete with histones for sites on the promoter and other regulatory elements. This is supported by additional reconstitution studies (Figure 9) which have shown that the order of protein addition is an essential

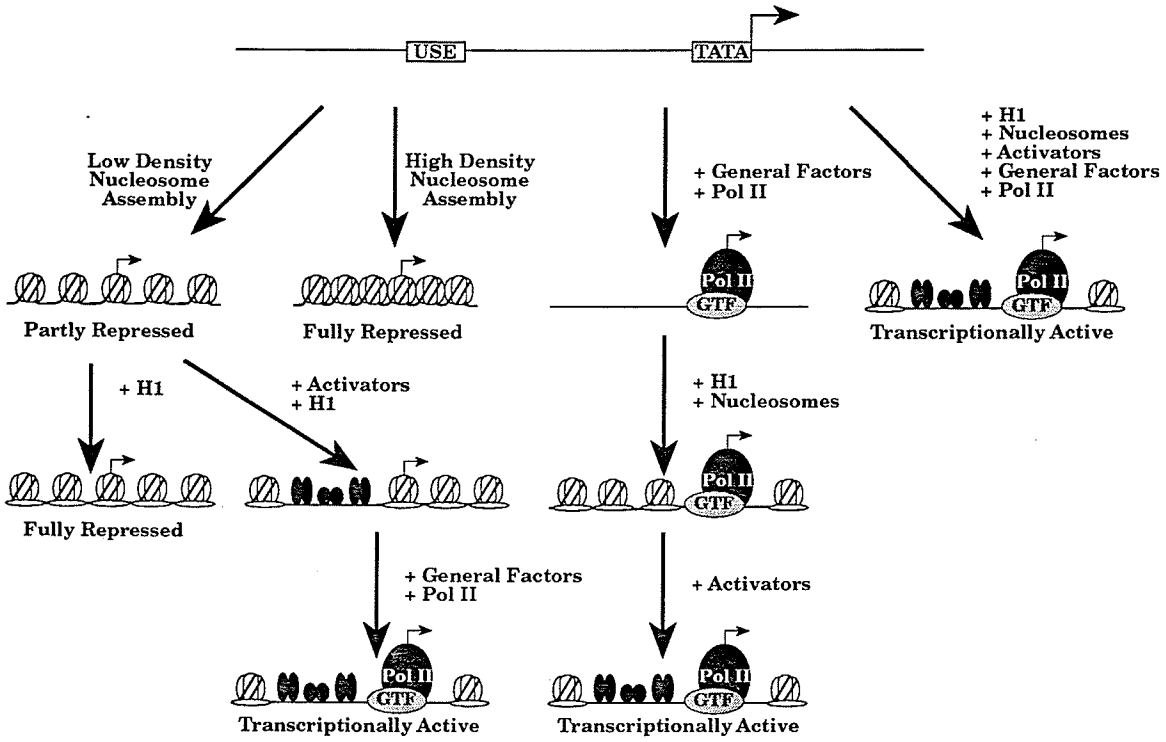


Figure 9. Summary of Chromatin Reconstitution Studies. A summary of experiments which examined the role of the general transcription factors, transcriptional activators and histones in the assembly of active or repressed chromatin. Low density nucleosome assembly partially represses transcription, which can be further repressed by the addition of the linker histone, H1 (ovals below each nucleosome). Full repression by H1 is blocked by the presence of activator proteins. High density nucleosome assembly, even in the presence of the general transcription factors (GTF), fully represses transcription. This can be prevented by the prior assembly of the general transcription factors or by the simultaneous addition of the general transcription factors and activator proteins with the nucleosomal histones. This figure was reproduced from Workman and Buchman, 1993.

component in the determination of an active or repressed DNA template (reviewed in Kornberg and Lorch, 1991; Felsenfeld, 1992; Workman and Buchman, 1993). Significantly, transcription is observed if the preinitiation complex is formed before but not after chromatin assembly, suggesting that TFIID is sufficient to protect the activity of the promoter (Workman and Roeder, 1987). Conversely, the simultaneous addition of TFIID and histones prevents transcription, suggesting that under these conditions TFIID is insufficient for maintaining transcriptional competence (Workman *et al.*, 1988; 1990). Alternatively, the simultaneous addition of TFIID and *trans*-acting factors with histones allows transcription to proceed (Workman *et al.*, 1991). Significantly, the maintenance of *trans*-activating domains is necessary for the exclusion of nucleosome formation over regulatory sequences, suggesting an interaction between *trans*-acting factors and TFIID may be essential to the disruption of adjacent nucleosomes (Felsenfeld, 1992; Workman and Buchman, 1993).

Histone Modifications

Higher-order chromatin folding is a major component in the suppression of gene activity. This compaction prevents accessibility of the polymerase and transcription apparatus to the DNA (Weintraub, 1985; Jackson, 1991). The unfolding of the chromatin domain is therefore an important prerequisite to transcriptional competence. Significantly, the association of H1 with chromatin differs between active and repressed regions (reviewed in Hayes and Wolffe, 1992). Histone H1 remains present in active chromatin, but at a reduced stoichiometry than that seen in repressed regions of the genome (Kamakaka and Thomas, 1990). Additionally, in active chromatin H1 no longer associates with the nucleosomal histones through its central globular domain, only the carboxy and amino-terminal

tails persist in maintaining the chromatin association (Nacheva *et al.*, 1989). Similarly, alterations in the core nucleosomes occur during transcription. The H3/H4 tetramer remains stably associated with DNA during transcription, while the H2A/H2B dimers are freely exchangeable (Hansen and Ausio, 1992). The H3/H4 tetramer is sufficient for nucleosome positioning, suggesting that loss of H2A/H2B dimers would not destroy the nucleosomal organization of chromatin but only increase accessibility of DNA to nucleases and *trans*-acting factors (Hayes *et al.*, 1991).

Nucleosomal DNA is sterically occluded by the histone octamer and the adjacent turn of the DNA (Hayes and Wolffe, 1992). One mechanism for increasing the access of nucleosomal DNA for *trans*-acting proteins and the general transcription factors may be through disruptions in chromatin structure which are regulated by post-translational modifications of core and linker histones. A role in conferring a transcriptionally active chromatin domain has been implied for a variety of different histone modifications including; acetylation, ubiquitination, methylation and phosphorylation (Davie and Murphy, 1990; Ridsdale *et al.*, 1990; Desrosiers and Tanguay, 1985; Grunstein, 1990). It is suspected that acetylation at lysine residues within the core histone amino-terminal tails disrupts nucleosomal structure through conformational changes induced in linker and nucleosomal DNA (reviewed in Hansen and Ausio, 1992). The subsequent dissociation of the linker histone promotes decondensation of the 30-nm fiber by preventing the association of neighbouring nucleosomes. It is possible that H4 amino-terminal mutations, which have been described to impair induction as well as repression, perform a similar function (reviewed in Grunstein, 1990; Kornberg and Lorch, 1991). Alternatively, phosphorylation of histone H1 has been suggested to alter its association with linker DNA (Grunstein, 1990;

Hansen and Ausio, 1992). Amino-terminal modification has been associated with chromatin decondensation, while phosphorylation of the carboxy-terminus is correlated with chromatin condensation (reviewed in Churchill and Travers, 1991). The effect of these and other histone modifications is likely to alter the nucleosomal DNA structure in order to circumvent obstructions to *trans*-acting factor accessibility (Hayes and Wolffe, 1992). Histone modifications are therefore critical controlling mechanisms for regulating chromatin conformation.

Nucleosome Positioning

Precise well defined positioning of nucleosomes within chromatin is known as nucleosome phasing and is important for gene-regulation (Hayes and Wolffe, 1992). This phenomenon is particularly evident in regions flanked by nuclease hypersensitive sites. Constitutive hypersensitive regions are always detected regardless of the transcriptional status of the gene (Gross and Garrard, 1988; Elgin, 1988). These sites may be flanked by strong nucleosome positioning sequences which do not leave enough space to accommodate a nucleosome. Alternatively, such sites may possess DNA sequences that exclude nucleosomes even in the absence of *trans*-acting factors (Hayes and Wolffe, 1992). Three reasons for nucleosome phasing have been suggested (reviewed in Gross and Garrard, 1988). Phasing could allow *cis*-acting sites separated by only one superhelical turn to be placed within the same rather than in neighbouring nucleosomes thus bringing the sites within close lateral proximity. Alternatively, precise positioning could increase accessibility of a binding site for its *trans*-acting factor by placing the site within the linker region between adjacent nucleosomes. Additionally, phasing could define the rotational orientation of the major and minor

grooves relative to the nucleosome, exposing a binding site located within nucleosomal DNA (Gross and Garrard, 1988).

Both active and passive mechanisms for nucleosome positioning have been proposed, however a combination of these methods is likely in use. Passive mechanisms propose that boundaries are created by the presence of other protein-DNA complexes which result in the exclusion of nucleosomes (Kornberg, 1981). This is supported by the observation that extended arrays of nucleosomes often exist between DH sites, although it could also suggest that the formation of a DH site initiates a cooperative unidirectional deposition of nucleosomes. Alternatively, the active process suggests the underlying sequence directs nucleosome positions along the DNA (Travers, 1987). Sequence directed positioning of nucleosomes would be determined by the formation of histone-DNA interactions which expend the least amount of energy for packaging (Hayes and Wolffe, 1992). Support for this model includes the phasing of nucleosomes during *in vitro* reconstitution studies and the observation that deletion of DH sites does not lead to loss of phased nucleosome arrays (Gross and Garrard, 1988). Additionally, A+T rich sequences are preferentially found facing inwards toward the histone octamer, while G+C rich sequences prefer to face outward (Travers, 1987).

Differences in the proportion of methylated bases between active and repressed regions of the genome further support an active role for the underlying DNA sequence in determining nucleosome positioning. Inactive genes are in a nuclease inaccessible condensed conformation which is in part due to the presence of methylated DNA, whereas active chromatin contains a smaller proportion of these bases and is more sensitive to nucleases (reviewed in Adams, 1990). These minor, modified bases are not introduced as DNA is synthesized but are produced shortly after replication by the transfer of a

methyl group from S-adenosyl-methionine by a methylase or methyl-transferase. The most common is 5-methylcytosine, which occurs largely in the dinucleotide sequence mCpG. These modified nucleotides are often found in the 5'-regions of genes (Bird, 1986).

In general, methylation increases the information content of the DNA. The methylation of CpG at promoter sequences is sufficient to repress expression of transgenes and is accompanied by assembly into nuclease-resistant chromatin structures (Cedar, 1988; Gross and Garrard, 1988). Conversely, decreased CpG methylation at many loci *in vivo* correlates with increased gene expression which is perhaps due to increased access for regulatory *trans*-activators (Cedar, 1988). Proteins interacting with DNA form bonds with individual bases of a target sequence, hence methylation has the potential to disrupt sequence-specific proteins from binding. Additionally, hypomethylation is suggested to create a permissive environment for establishing hypersensitive sites, whereas hypermethylation has been shown to suppress DH site formation (Groudine *et al.*, 1981; Saluz *et al.*, 1988). However, methylation of cytosine also enhances the probability that a stretch of DNA will assume a Z configuration (Behe and Felsenfeld, 1981), a conformation which has been associated with active transcription (Wittig *et al.*, 1992).

Nucleosome Exclusion

Changes in the sensitivity and hypersensitivity of chromatin to DNase I occur prior to gene expression and reflect transcriptional potential (Gross and Garrard, 1987). During replication nucleosomes are removed from chromatin. The prevention of nucleosome reassembly during the early stages of this process may be a necessary prerequisite for the assembly of some transcriptionally active domains (Goldman, 1988). Significantly,

transcriptionally competent genes have been associated with early replicating regions, while genetically inert genes are generally found to replicate later (Goldman, 1988). Additionally, chromatin loop domains appear to replicate as a single unit under the control of a single replication origin. This replicative process has been demonstrated to be associated with the nuclear matrix (reviewed in Berezney, 1991; Jackson, 1991). Furthermore, active and repressed chromatin appear to be compartmentalized differently, with the active genes associated with the nuclear periphery and the borders of condensed chromatin masses in close association with the nuclear matrix (Hutchison and Weintraub, 1985; Jackson, 1991).

Preemptive or persistent nucleosome exclusion (Figure 10 A) requires replication to remove nucleosomes which block transcription factor access and may have compacted chromatin into a higher order structure (reviewed in Felsenfeld, 1992; Workman and Buchman, 1993). These regulatory regions are resistant to direct nucleosome displacement by *trans*-acting factors. Displacement of nucleosomes during replication results in a chromatin state which requires about 20 minutes to return to the nuclease sensitivity of bulk chromatin (Worcel *et al.*, 1978). The time frame required for nucleosome reassembly is suggested to provide an opportunity for *trans*-acting factors to compete for binding sites with nucleosomal proteins and thus reorganize chromatin into active domains (Svaren and Chalkley, 1990). DNase I footprints of *trans*-acting factors and the initiation complex are 10-20 bp and 75 bp, respectively, compared to at least 146 bp for the nucleosome core proteins (Elgin, 1988; Svaren and Chalkley, 1990; Kornberg and Lorch, 1991). The kinetics of protein-DNA interactions and the differences in binding site size between transcriptional activators and nucleosomal proteins suggest this simplified model may be part of the reprogramming process (reviewed in

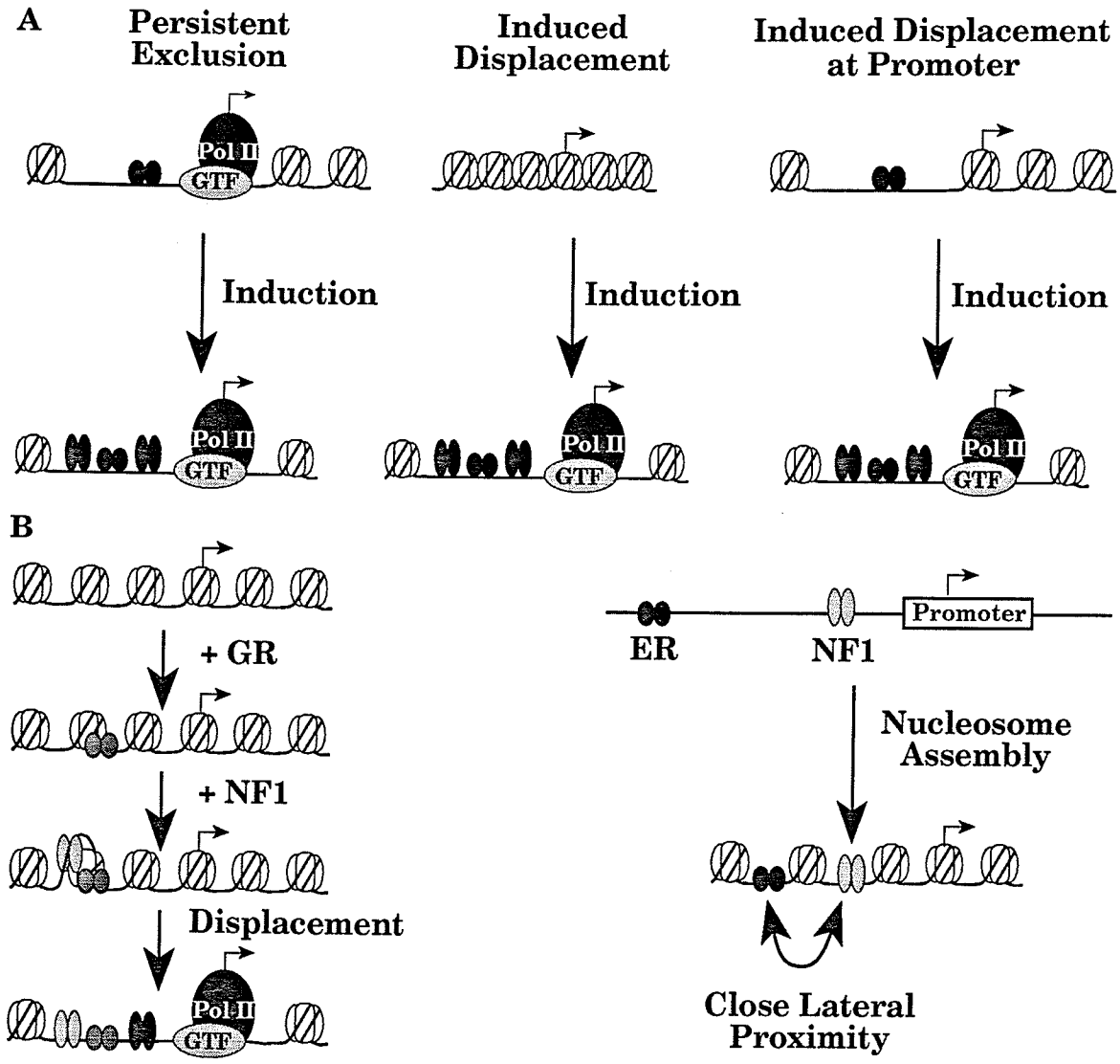


Figure 10. Models for Nucleosome Displacement. A. Persistent exclusion of nucleosomal structures by *trans*-acting factors and the initiation complex occurs shortly after replication when the DNA is free of histones. Displacement of assembled nucleosomes by *trans*-acting factors forms nucleosome free regions allowing formation of the initiation complex. Persistent nucleosome exclusion at *cis*-acting elements may lead to nucleosome displacement at promoter regions. **B.** Two examples of how nucleosomal packaging of DNA assists the formation of transcriptionally active chromatin. The GR is able to bind to the GRE even when it is located within a nucleosome. This interaction assists NF1 binding and the binding of other factors results in nucleosome displacement. The assembly of a nucleosome between the ER and NF1 binding sites places these factors in closer lateral proximity and facilitates their interaction. Part A of this figure was reproduced from Workman and Buchman, 1993. The remainder of the figure was derived from Adams and Workman, 1993; Hayes and Wolfe, 1992.

Svaren and Chalkley, 1990). This model is supported by reconstitution studies (Figure 9) which have demonstrated that the preassembly of TFIID (Workman and Roeder, 1987) or the simultaneous addition of TFIID and *trans*-acting factors (Workman *et al.*, 1991) prevents nucleosome formation over important regulatory regions and allows transcription to proceed.

During the deposition of histones and reassembly of nucleosomes after replication, the proteins assembled at the promoter and other regulatory regions would have to be firmly established to reliably create an altered chromatin structure (Felsenfeld, 1992). The synergistic interactions between individual *trans*-acting factors and their association with the transcriptional machinery may be mechanisms for attaining this stability. Interestingly, two molecules of PR or Sp1 (Klein-Hitpass *et al.*, 1990; Su *et al.*, 1991) or one of Sp1 and an additional *trans*-acting factor (Li *et al.*, 1991) bound at separated sites have been shown to induce loop formation and enhance the formation of a preinitiation complex at the promoter. This supports the notion that the stable interactions of transcriptional regulators with chromatin and each other may be involved in the persistent exclusion of nucleosomes from DNA. The ubiquitous expression of Sp1 and its ability to form multimers (Courey *et al.*, 1989) may be an indication that this *trans*-acting factor performs a specific role in this process.

Mechanisms of Nucleosome Displacement

The precise positioning of a regulatory element in relation to the nucleosome can have a profound effect on the ability of a *trans*-acting factor to recognize its binding site (reviewed in Hayes and Wolffe, 1992; Felsenfeld, 1992; Adams and Workman, 1993; Workman and Buchman, 1993). In some circumstances inducible, tissue specific and developmental DH sites are organized into nucleosomes which are displaced prior to gene activation in

the absence of replication (Grunstein, 1990). These sites suggest additional mechanisms, which do not require replication, are also involved in chromatin remodeling. One prediction is that nucleosome displacement in the absence of replication may require phasing frames that precisely position some recognition sites such that they remain accessible to *trans*-acting proteins (Gross and Garrard, 1988). These binding sites would be located within the linker DNA between phased nucleosomes, positioned within the nucleosome in an orientation facing away from the histone octamer or located at the entry or exit points of nucleosomal DNA (Felsenfeld, 1992). Protein-DNA interactions at these sites are examples of dynamic nucleosome competition which is suggested to displace histones and disrupt the binding of adjacent nucleosomes (Kornberg and Lorch, 1991; Felsenfeld, 1992).

The positioning of regions essential for transcription in nucleosome free DNA would eliminate nucleosome displacement as a requirement for the interaction of a *trans*-acting factor with its binding site (Felsenfeld, 1992; Workman and Buchman, 1993). Conversely, the packaging of some regulatory regions into nucleosomal structures may require displacement of histones prior to binding (Figure 10 A). This could be induced by the association of a *trans*-acting factor with a neighbouring nucleosome free binding site or with a *cis*-acting site which is packaged but still accessible for binding. Similarly, the initial response element may be nucleosome free but access of transcription factors to the TATA box and other critical regulatory sequences may require nucleosome displacement. Specific DNA-binding proteins that are able to directly interact with packaged DNA would be required to facilitate these structural changes (Felsenfeld, 1992).

Evidence suggests that DH sites may not be completely nucleosome free but may represent histone-factor co-occupancy which disrupts but does

not completely displace nucleosomal histones (Adams and Workman, 1993). The formation of ternary complexes between *trans*-activators, DNA and nucleosomes may facilitate the binding of additional factors which further destabilize or alter the conformation of histones for the disruption or displacement of nucleosomes. The detection of 'split' nucleosomes *in vivo* (Lee and Garrard, 1991) may represent the formation of these ternary structures and the subsequent nucleosome disruption. The process of creating a nucleosome free region may be the reverse of nucleosome assembly, with the disruption of H1 interactions followed by the loss of H2A/H2B dimers and possible displacement of the H3/H4 tetramer (Svaren and Chalkley, 1990; Adams and Workman, 1993). The entire process would reorganize chromatin for increased accessibility of the DNA to additional *trans*-acting proteins and the formation of the preinitiation complex. Consequently, the synergistic association of multiple *trans*-acting factors and their interaction with the transcriptional machinery would ensure that transcription proceeds.

The formation of a ternary complex between *trans*-acting factors, histones and DNA is a property of only a subset of transcriptional activators (Adams and Workman, 1993). The GR is one of the few characterized examples of a factor which can interact with an exposed recognition element positioned within a nucleosome (Perlmann and Wrange, 1988). This hormone receptor is able to activate transcription even when its binding site is packaged into a chromatin template (reviewed in Gross and Garrard, 1988; Grunstein, 1990; Hayes and Wolffe, 1992; Workman and Buchman, 1993). There are 6 positioned nucleosomes within the MMTV promoter region, one of which contains a GR binding site (Figure 10 B). Although located within a nucleosome, the GRE is oriented away from the histone core (Hayes and Wolffe, 1992; Workman and Buchman, 1993). Additionally, the GRE

half-sites are separated by one helical turn of the DNA causing both to be positioned on the same face of the DNA helix (Schwabe and Rhodes, 1991). NF1 does not recognize DNA that is in contact with histones. The association of the GR with its recognition sequence facilitates the binding of NF1 (Archer *et al.*, 1992) by displacing H1 and disrupting the repressive nucleosome core. Conversely, ER binding to the vitellogenin B1 promoter region is an example of protein associations that are facilitated by nucleosome positioning (Figure 10 B). A nucleosome is positioned between the ERE sequence and the binding site for NF1. The packaging of DNA around this nucleosome forms a nucleosome dependent loop which decreases the lateral distance between the chromatin bound ER and NF1 (Hayes and Wolffe, 1992; Schild *et al.*, 1993). This nucleosome phasing therefore provides a mechanism for the interaction of these two factors bound to their linker DNA positioned recognition sites and facilitates ER mediated transcription.

Chromatin Remodeling

Decondensation of chromatin for transcription presumably occurs by the sequential loss of higher order packaging which is coincident with increased access to regulatory regions of DNA (Kornberg and Lorch, 1991). This suggests that the mechanisms for nuclear matrix association, nucleosome formation and transcription factor assembly are coordinated activities. Stable (S/MARs, A-elements) and competent (LCRs) associations with the nuclear matrix are necessary to create an appropriate chromatin domain for proper transcriptional regulation by *trans*-acting proteins. The formation of stable nuclear matrix attachments defines the boundaries of chromatin domains while the conversion of repressed chromatin into active domains could be controlled by the interaction of proteins with the dominant acting LCRs (Bonifer *et al.*, 1991). The activities of these elements may be

determined by the methylation status of the underlying DNA sequence (Groudine *et al.*, 1981; Adams, 1990) and the precise positioning of nucleosomes (Gross and Garrard, 1988; Hayes and Wolffe, 1992). Observations of irreversible repression by the core histones and semi-reversible inhibition by H1, which is precluded by the presence of *trans*-activators, supports a role for nucleosomal DNA packaging in the regulation of gene expression (Felsenfeld, 1992). This in turn may be regulated by post-translational modifications of histones induced by the association of *trans*-acting proteins with nucleosomal DNA (Grunstein, 1990; Hansen and Ausio, 1992; Hayes and Wolffe, 1992). Many regulatory sequences and proteins are therefore concerned with catalyzing the remodeling of chromatin structure for transcription (Kornberg and Lorch, 1991).

One hypothesis suggests that reorganization into transcriptionally active domains occurs during DNA replication. Stable nuclear matrix attachments may initially determine the time of replication and therefore the availability of *trans*-acting factors. Chromatin would be packaged into open domains if the deposition of *trans*-acting factors over regulatory and promoter regions occurred prior to nucleosome assembly. The responsibilities of these factors would include the formation and stabilization of large preinitiation complexes which exclude nucleosome formation and direct nucleosome deposition. LCRs, which contain multiple binding sites for different activators, may be the sites responsible for regulating chromatin decondensation. The proteins, which bind to these dominant control elements, may cooperate in the unfolding process of large regions of DNA through effects on histones or nonhistone chromosomal proteins (Kornberg and Lorch, 1991; Felsenfeld, 1992). Nucleosome positioning would facilitate

the synergistic interactions of *trans*-activators bound at laterally separated regulatory sites. Transcription factor synergy may be required to maintain an open chromatin conformation and recruit the transcription machinery.

In later stages of replication or when *trans*-acting factor levels are depleted, nucleosome formation and H1 binding would package some regions of DNA into repressed chromatin as directed by their specific nuclear matrix attachments and methylation status. These compacted domains would require the release of H1 repression and the relaxation of nucleosomal structures prior to transcriptional activation. In the absence of replication, transcriptional competence would therefore require at least two modifications of chromatin; the interaction of *trans*-acting factors and the displacement of histone complexes (Svaren and Chalkley, 1990). The interaction of factors with DNA, which is incorporated within a nucleosome, may mediate the initial steps in destabilizing repressive chromatin structures (Hayes and Wolffe, 1992). These activators would invade nucleosomes and facilitate the breakdown of these structural impediments to transcription by direct or indirect contact with histones (Kornberg and Lorch, 1991; Felsenfeld, 1992). Structural changes in chromatin would require modifications of histones and the subsequent interaction of additional *trans*-acting factors with their DNA-binding sites.

The precise placement of nucleosomes in regulatory regions of the genome is therefore critically important (Felsenfeld, 1992). The ultimate result of *trans*-acting factor deposition at these sites by both preemptive and dynamic mechanisms would be the exclusion, disruption and displacement of histones with the subsequent unfolding of the chromatin domain. The ability to facilitate the disruption of higher order chromatin packaging is one of the important functions of *trans*-activators (Kornberg and Lorch, 1991).

Mutations in *cis*-acting elements within a promoter can inactivate transcription without eliminating DH sites suggesting specific factors are responsible for maintaining nucleosome disruption and displacement (Lee and Garrard, 1992). The purpose of LCR sequences may be to provide the initial sites for *trans*-activator binding which would then direct cooperative interactions with promoter binding factors to keep this region nucleosome free (Felsenfeld, 1992). Additionally, the binding of factors to nucleosomal DNA and their association with the nuclear matrix may induce nuclear matrix bound enzymes to catalyze modifications of histones, which may be necessary for the removal of H1 repression and the induction of nucleosome disassembly.

The ability of the general transcription factors to compete with nucleosomes is dependent on *trans*-activation by upstream regulatory factors (Felsenfeld, 1992; Workman and Buchman, 1993). The binding of additional *trans*-acting factors would provide momentum for the formation of a functional transcription complex with the general transcription proteins associated at the transcriptional initiation site. The conventional activity of *trans*-acting proteins to participate in active complex formation with the general transcription machinery would be facilitated by protein-protein interactions which form smaller functional loops within the larger loop of the chromatin domain (von Kries *et al.*, 1991). This provides a mechanism for the interaction of factors bound at regulatory sites located distal to the promoter.

The c-*myc* Gene and Protein

A vast amount of literature has been published on c-*myc* and its protein since their discovery. Originating in the fields of virology and oncogenesis, research on the c-*myc* gene and its protein continue to intrigue

scientists in these areas, as well as in the areas of gene regulation and development. The c-*myc* gene and its protein are widely expressed and contribute to the functioning of nearly all cell types. Research on the c-Myc protein has focused on its role in proliferation, differentiation and apoptosis, while studies on the c-*myc* gene have been concerned with unraveling its complex regulation and understanding its oncogenic activation.

The first genes discovered to be involved in oncogenesis were originally cloned as the transforming genes of highly oncogenic avian retroviruses. Transformation with these retroviruses resulted in a variety of malignancies, including leukemias, carcinomas, and sarcomas. In 1977, the v-*myc* gene was identified as the transforming gene of the avian sarcoma virus MC29 (Duesberg *et al.*, 1977). The c-*myc* proto-oncogene, the cellular homologue of the retroviral v-*myc* oncogene, was subsequently isolated from the chicken in 1979 (Sheiness and Bishop, 1979). The identification of this highly conserved proto-oncogene led to its isolation from a variety of species, including human (Dalla-Favera *et al.*, 1982; Gazin *et al.*, 1984). As well, a family of *myc*-related genes has been identified due to their considerable homology to the c-*myc* gene (DePinho *et al.*, 1987a; 1987b; 1991).

Studies have shown that the c-Myc protein is localized to the nucleus (Donner *et al.*, 1982; Persson and Leder, 1984) and exists as a phosphoprotein (Lüscher *et al.*, 1989). This protein has a role in a variety of important cellular processes, in particular proliferation, differentiation, apoptosis and transformation (Cole 1986; Cory, 1986; Lüscher and Eisenman, 1990; Penn *et al.*, 1990b; Spencer and Groudine, 1991; Kato and Dang, 1992; Marcu *et al.*, 1992; Evan, *et al.*, 1992). c-Myc's role in these processes probably occurs at the level of replication control (Ariga *et al.*, 1989; Iguchi-Ariga *et al.*, 1988) and/or transcriptional regulation (Blackwell *et al.*, 1990, Prendergast and

Ziff, 1991; 1992). These roles are not mutually exclusive. A large body of data support a transcriptional function of c-Myc as a DNA-binding protein. This activity of c-Myc has gained considerable support with the discovery of a dimerization partner for Myc, which has been named Max (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991).

myc Gene Family

The human *myc* gene family consists of c-*myc*, N-*myc*, and L-*myc* (reviewed in DePinho *et al.*, 1987a; 1991). The N-*myc* and L-*myc* genes were first identified as amplified sequences in human neuroblastomas and small cell lung carcinomas, respectively (Schwab *et al.*, 1983; Kohl *et al.*, 1983; Nau *et al.*, 1985). These three *myc* family members were also identified in the mouse, although the chromosomal locations differ between the two species. The c-*myc*, N-*myc* and L-*myc* genes have been located to human chromosomes 8, 2, and 1, respectively, while the mouse genes have been located to chromosomes 15, 12 and 4, respectively (DePinho *et al.*, 1991). The relationship between these three members of the *myc* gene family extends to the overall architecture of the gene and its transcript, as well as specific regions of considerable sequence homology within the coding region of the gene.

The *myc* genes all share a characteristic three exon structure (Figure 11), with the first exon being untranslated and the polypeptide open reading frame encoded by exons 2 and 3 (Kohl *et al.*, 1986). The c-Myc protein can be initiated at an AUG located at the 5' end of exon 2 as well as at a CUG located at the 3' end of exon 1 (Hann *et al.*, 1988). These translation start sites generate c-Myc polypeptides of 64 and 67 kilodaltons (kDa), respectively (Figure 12). The N-Myc protein is initiated from two inframe AUG codons at the 5' end of exon 2 to yield proteins of 58 and 64 kDa (Ramsay *et al.*, 1986;

Slamon *et al.*, 1986; DePinho *et al.*, 1991; Marcu *et al.*, 1992). Although there are two in frame CUG codons located in the first intron, the L-Myc protein appears to be initiated at an AUG near the 5' end of exon 2 (Marcu *et al.*, 1992; DePinho *et al.*, 1991; DeGreve *et al.*, 1988). The multiple sizes of the L-Myc proteins appear to be the result of differing levels of posttranslational modifications, such as phosphorylation of the 60 kDa protein to the 66 kDa form (Marcu *et al.*, 1992; DePinho *et al.*, 1987; Legouy *et al.*, 1987; DeGreve *et al.*, 1988; Ikegaki *et al.*, 1989). The 5' and 3' untranslated regions of the individual *myc* family genes are unique, but are evolutionarily conserved for a given *myc* gene (DePinho *et al.*, 1991). This lack of similarity between *myc* genes within potential regulatory regions suggests differences in gene expression of the *myc* family members. Indeed, the expression pattern of these *myc* genes is distinct with respect to cell type and developmental stage.

Additional members of the *myc* gene family have been identified. A 34 to 37 kDa L-*myc* protein, which lacks the third exon sequences, has been identified in human small-cell lung carcinoma (Ikegaki *et al.*, 1989), while two L-*myc* pseudogenes, L-*myc-psi* (DePinho *et al.*, 1987b) and p-*myc* (Alt *et al.*, 1986) have been identified on the basis of homology to L-*myc* exon 3. Both pseudogenes have been mapped to the X chromosome (DePinho *et al.*, 1991). The B-*myc* (Ingvarsson *et al.*, 1988) and S-*myc* (Sugiyama *et al.*, 1989) genes have been identified in the rat genome and represent truncated genes. The B-*myc* gene shares considerable sequence identity with the c-*myc* gene throughout the coding and intron sequences, but lacks the third exon sequences (Ingvarsson *et al.*, 1988). This gene encodes a 188 amino acid protein which is widely expressed in most tissues. The S-*myc* gene contains a high degree of sequence similarity to the second and third exon of the N-*myc* gene (Sugiyama *et al.*, 1989). A unique aspect of S-*myc* is that transfection

into a rat tumor cell line suppressed the cells' malignant phenotype, suggesting this *myc* gene may have tumor suppressor activity (Sugiyama *et al.*, 1989). Finally, there are two c-*myc* genes, c-*myc* I and c-*myc* II, in the *Xenopus laevis* genome (Vriz *et al.*, 1989). These two genes differ in their 5' and 3' untranslated sequences, as well as in their patterns of expression. The c-*myc* II gene is expressed from the maternal genome during oogenesis, while c-*myc* I is expressed during oogenesis as well as in the postgastrula embryo (Marcu *et al.*, 1992; Vriz *et al.*, 1989).

c-*myc* Gene Structure

The general structure of the c-*myc* gene has been highly conserved throughout vertebrate evolution from Xenopus to humans (reviewed in Spencer and Groudine, 1991). The normal c-*myc* gene is composed of three exons with the coding region located solely on exons 2 and 3 (Figure 11). All c-*myc* genes contain a long untranslated exon 1, which has termination codons in all three reading frames but no translational initiation codons (Battey *et al.*, 1983; Bernard *et al.*, 1983; Stanton *et al.*, 1983). This exon is 70% conserved between human and mouse (Bernard *et al.*, 1983) but has little similarity with the chicken c-*myc* exon 1 (Shih *et al.*, 1984; Linial and Groudine, 1985; Nottenberg and Varmus, 1986). There are two conserved polyadenylation sites which have been identified in the c-*myc* gene of chicken, rat and human (Spencer and Groudine, 1991). The second site, which is separated from the first by 150 bp, is utilized 6 fold more frequently (Laird-Offringa *et al.*, 1989).

The presence of multiple transcription start sites is another conserved feature of the c-*myc* gene (Figure 11). There are four sense promoters described for the c-*myc* gene, P_0, P_1, P_2 and P_3 (reviewed in Spencer and Groudine, 1991; Marcu *et al.*, 1992). The P_1 and P_2 promoters are located

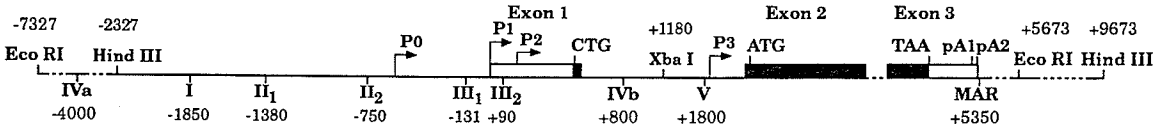


Figure 11. Structure of the c-*myc* Gene. The three exon structure of the c-*myc* gene is shown, with the black area representing the coding region. There are four associated promoters (P0-P3) labelled with arrows. The approximate location of the two translation initiation sites (CTG, ATG), the translation termination site (TAA) and the two polyadenylation signals (pA1, pA2) are shown above the exons. DNase I hypersensitive sites are labelled as roman numerals with their approximate location relative to P1 shown under each site, as described for HL60 cells. The approximate location of a matrix attachment site (MAR) relative to P1 is also shown. A number of restriction sites are labelled for reference. This figure was derived from Siebenlist *et al.*, 1984; Spencer and Groudine, 1991 and information referenced in the text.

within exon 1, lying 161 bp apart in the human gene (Battey *et al.*, 1983; Watt *et al.*, 1983) and 164 bp apart in the mouse gene (Bernard *et al.*, 1983). Both contain classical TATA boxes and initiate transcripts of 2.4 and 2.2 kb, respectively. In normal cells the P_2 promoter is the stronger of the four promoters initiating 75-90% of the c-*myc* transcripts. The P_1 promoter accounts for 10-25% of c-*myc* mRNAs (Stewart *et al.*, 1984; Bentley and Groudine, 1986a; 1986b; Spencer *et al.*, 1990). The minor promoters P_0 and P_3 lack canonical TATA sequences and account for only 5% of c-*myc* transcripts. In the human gene the P_0 promoter lies 550 to 650 bp upstream of P_1 and contains multiple start sites. Transcripts initiating at P_0 have been reported to be 3.1 and 2.5 kb (Bentley and Groudine, 1986a; 1986b). The P_3 promoter is positioned near the 3' end of intron 1 and initiates transcripts of 2.3 kb from multiple start sites (Ray and Robert-Lézénès, 1989). This promoter was initially thought to be utilized only when the other promoters were disrupted by c-*myc* translocation (ar-Rushdi *et al.*, 1983; Bernard *et al.*, 1983; Hayday *et al.*, 1984), however, P_3-RNAs have recently been detected from an unrearranged c-*myc* gene (Eick *et al.*, 1990). Antisense transcription has been observed in the mouse c-*myc* gene (Dean *et al.*, 1983; Nepveu and Marcu, 1986) and a promoter has recently been mapped within intron 2, 175 bp from the start of exon 3 (Spicer and Sonenshein, 1992). No corresponding promoter has been identified in the human c-*myc* gene, however antisense transcripts have been reported for the N-*myc* gene (Krystal *et al.*, 1990).

Chromatin Structure of c-*myc*

Analysis of the region 13 kb 5' and 3.5 kb 3' of the human c-*myc* gene locus has revealed a number of DNase I hypersensitive sites (Siebenlist *et al.*, 1984; Dyson *et al.*, 1985; Bentley and Groudine, 1986a; 1986b; reviewed

in Spencer and Groudine, 1991; Marcu *et al.*, 1992). These sites mark the center of regions believed to be free of nucleosomal proteins and bound by regulatory DNA-binding proteins. The location of these DH sites relative to the P_1 promoter (Figure 11) are as follows; IVa (-4000), I (-1850), II_1 (-1380), II_2 (-750), III_1 (-131), III_2 (+90), 4 or IVb (+800) and V (+1800) (Siebenlist *et al.*, 1984; Dyson *et al.*, 1985; Bentley and Groudine, 1986a; 1986b). Using Eco R I endonuclease digestion only sites I, II and III were identified 5' to c-*myc* exon 1 (Siebenlist *et al.*, 1984). High resolution mapping with the endonuclease Sst I determined that sites II and III each corresponded to two sites, II_1, II_2 and III_1, III_2 respectively. Additionally, site IV (IVa) was located upstream of site I just 5' of a Bgl II restriction site. Subsequently, two hypersensitive sites located in intron 1, 4 (IVb) and 5 (V), were detected (Siebenlist *et al.*, 1984; Bentley and Groudine, 1986a). Site IVb mapped near an Sst I site within the first intron, while site V mapped just 5' of the second exon. Use of the appropriate probes and digestion with Bgl II allowed the detection of sites I, II_1, II_2, III_1, III_2, IVb and V, while digestion with Xba I allowed the detection of sites IVa, I, II_1, II_2, III_1 and III_2 (Dyson *et al.*, 1985). An S1 nuclease sensitive site located at the boundary of intron 1 and exon 2 may correspond to site V (Grosso and Pitot, 1985). Some of these DH sites have also been mapped in the c-*myc* chromatin of the mouse, and correspond to -1896 (I), -911 (II_1), -255 (II_2), +15 (III_1) and +183 (III_2), relative to P_1 (Mango *et al.*, 1989).

A function has not been assigned to DH site IVa (Siebenlist *et al.*, 1984; Dyson *et al.*, 1985), and it is not mentioned in any of the recent reviews on c-*myc* (Spencer and Groudine, 1991; Marcu *et al.*, 1992). Its location 4 kilobases 5' of the c-*myc* gene suggests it may correspond to a nuclear matrix attachment region. Similarly, although DH sites have not been identified in

the 3' flanking sequences of the c-*myc* gene, a MAR has been defined in this region which overlaps the two poly-adenylation signals and a topoisomerase II-like sequence (Chou *et al.*, 1990). DH site I is located in a region that shows strong evolutionary conservation, is sometimes retained when c-*myc* is repressed and may correspond to a negative control element (Siebenlist *et al.*, 1984). Sequences near site II_1 correlate with a possible origin of DNA replication (Iguchi-Ariga *et al.*, 1988; McWhinney and Leffak, 1990; Vassilev and Johnson, 1990) and may bind c-Myc (Ariga *et al.*, 1989), nuclear factor 1 (Siebenlist *et al.*, 1984) or some other factors (Saksela *et al.*, 1991). This DH site has also been implicated as a negative regulatory element (Whitelaw *et al.*, 1988; Lang *et al.*, 1988). DH site II_2 is associated with the P_0 promoter (Bentley and Groudine 1986b), while sites III_1 and III_2 are associated with the P_1 and P_2 promoters, respectively (Siebenlist *et al.*, 1984). The relative sensitivity of these sites is reflective of their corresponding transcriptional activity, with the P_2 promoter being utilized more often (Siebenlist *et al.*, 1984). Site IVb is associated with a block to transcriptional elongation, while site V is associated with the P_3 promoter (Siebenlist *et al.*, 1984; Bentley and Groudine 1986a; 1986b; Eick *et al.*, 1990).

Five of these sites (I, II_1, II_2, III_1, and III_2) are present in the c-*myc* chromatin of all cell lines studied to date except those Burkitt's lymphoma (BL) cell lines where the translocation breakpoint disrupts the site (Rabbitts *et al.*, 1983; 1984; Siebenlist *et al.*, 1984; Dyson and Rabbitts, 1985; Cesarman *et al.*, 1987; Spencer and Groudine, 1991). A similar pattern of hypersensitivity was detected over the same region using micrococcal nuclease. This enzyme identified three additional nuclease sensitive sites between DH sites II_2 and III_1 (Kumar and Leffak, 1991). It was therefore suggested that the region between P_0 and P_1 may resist packaging into

nucleosomes. Conversely, an ordered chromatin structure, possibly due to positioned nucleosomes, was evident 5' of the P_0 promoter (Kumar and Leffak, 1989). These hypersensitive regions remained present even when transduced to a new site in the genome, suggesting that the DNA sequence elements are important for organizing the c-*myc* chromatin structure (Kumar and Leffak, 1991). Significantly, during active transcription of c-*myc* the P_0 and P_3 promoter regions have been shown to form Z-DNA which is lost when the cells begin to differentiate (Wittig, *et al.*, 1992).

HL60 cells permanently repress c-*myc* from P_0, P_1 and P_2 after prolonged treatment with differentiating agents. Concurrently, DH sites II_1, II_2 and III_2 are lost (Siebenlist *et al.*, 1988). This is also seen with translocated c-*myc* in BL cell lines (Siebenlist *et al.*, 1984; 1988; Dyson *et al.*, 1985). In contrast all DH sites are maintained if the translocated allele has sequences for the sites. In HL60 cells sites I and V are present in both active and inactive genes, while sites IVb, II_2 and III_1 diminish markedly and III_2 is partially diminished during the induction of differentiation (Dyson *et al.*, 1985). Similarly, three of the c-*myc* DNase I hypersensitive sites (II_1, II_2, III_1) are lost in the nontranslocated allele of Burkitt's lymphoma (BL) cells although these same sites are maintained in the translocated allele (Siebenlist *et al.*, 1984; 1988; Kakkis *et al.*, 1986; Dyson and Rabbitts, 1985). Sites IVa and V were mapped in BL and lymphoblastoid cell lines but their presence seems variable (Siebenlist *et al.*, 1984; Dyson and Rabbitts, 1985). In some BL cell lines (e.g.. Ramos) site II_2 is not detected (Dyson and Rabbitts, 1985) even though the translocation breakpoint is 2 kb 5' of the first exon (Cesarman *et al.*, 1987). Significantly, all DH sites observed with the actively transcribed c-*myc* gene are maintained in the chromatin of quiescent cells.

The c-*myc* gene contains DNase I hypersensitive sites which are associated both with promoter proximal and distal regulatory regions. The entire c-*myc* domain appears to be flanked by DH site IVa at -4000 relative to P_1 and a MAR sequence at +5350 relative to P_1. This suggests that the c-*myc* domain encompasses at least 9.35 kbp of human chromosome 8. Sites II_2, III_1, III_2, and V are all associated with a promoter region, while DH site IVb is associated with premature termination of transcription. The presence of DH site I and V in both active and inactive human c-*myc* chromatin suggests that these two hypersensitive sites may be constitutive. Conversely, DH sites II_1, II_2, III_1, III_2 and IV are lost or diminished in the process of differentiation and repression, suggesting these sites may be stage specific. The observation that an ordered chromatin structure exists between hypersensitive sites I and II_2, suggests that this region of the c-*myc* gene may be involved in positioning nucleosomes. Similarly, the increased accessibility of c-*myc* sequences between DH sites II_2 and III_1 to nuclease digestion indicates this region may be depleted of nucleosomes. Additionally, the conversion of right-handed B-form DNA into high energy left-handed Z-DNA over the P_0 and P_3 promoter regions during transcription could be due to negative supercoiling generated by the removal of nucleosomes or the passage of the RNA polymerase. Together the data support a complex arrangement of transcriptional regulatory regions located within the noncoding regions of the human c-*myc* gene.

c-Myc Protein

The c-Myc protein is a highly phosphorylated nuclear protein which contains a high proline content (Donner *et al.*, 1982; Abrams *et al.*, 1982). This protein has a short half life of 20 to 30 minutes (Hann and Eisenman, 1984). The c-*myc* mRNA is translated into a 439 amino acid protein (Figure

12) when initiated from the conventional AUG start codon located at the 5' end of exon 2 (Hann and Eisenman, 1984; Ramsay *et al.*, 1984). The first 265 amino acids of the AUG initiated c-Myc protein are encoded by exon 2 sequences, with the remaining carboxy terminal amino acids encoded by exon 3. Initiation at an upstream CUG alternative translation start site located at the 3' end of exon 1 adds an additional 14 amino acids to the amino terminus generating a protein of 453 amino acids (Hann *et al.*, 1988). Although the predicted molecular mass of c-Myc is 49 kDa, the proteins run as a doublet of 64 and 67 kDa on SDS polyacrylamide gels (Alitalo *et al.*, 1983; Hann *et al.*, 1988; Persson, *et al.*, 1984). The P_0, P_1 and P_2 promoters initiate transcripts which generate both the 67 and 64 kDa c-Myc proteins, while the P_3 promoter initiates transcripts which restrict translation to the 64 kDa protein (Figure 11). Recent evidence suggests that a third c-Myc protein, with an additional ten amino acids at the amino terminus of the 64 kDa form, which differs from the 67 kDa form, could be translated from P_3 initiated transcripts (Eick *et al.*, 1990). These three forms of the c-Myc protein therefore share the same opening reading frame.

In the c-*myc* gene there are two additional open reading frames, ORF I and ORF II, which are located upstream of and overlapping exon 1, respectively. Both ORF I and ORF II could be encoded by transcripts initiating at the P_0 promoter (Bentley and Groudine, 1986b; Gazin *et al.*, 1984; Gazin *et al.*, 1986; Spencer and Groudine, 1991). The ORF I protein initiates at an AUG located 595 bp 5' of P_1 and terminates at a TGA at -253 relative to P_1. This ORF generates a polypeptide of 114 amino acids which has been identified *in vitro* (Bentley and Groudine, 1986b). The ORF II protein initiates at -23 relative to P_1 and terminates at a codon 12 bp 5' of the exon 1 / intron 1 boundary (Gazin *et al.*, 1984). This ORF generates a 188

amino acid polypeptide which has been identified immunologically in HeLa cells as a dimeric 58 kDa protein (Gazin *et al.*, 1986). The ORF II is not present in mouse or chicken c-*myc* so its biological significance is uncertain (Spencer and Groudine, 1991).

Modular Structure

The p64 and p67 forms of c-Myc are evolutionarily conserved from *Xenopus* to human (King *et al.*, 1986). The amino acid conservation between human c-Myc and mouse, chicken and trout c-Myc proteins is 90%, 70% and 62%, respectively (Bernard *et al.*, 1983; Watson *et al.*, 1983; van Beneden *et al.*, 1986; Kato and Dang, 1992). Similarly, the *myc* family of genes share regions of extensive sequence homology separated by stretches of completely divergent sequence. The N-Myc and L-Myc proteins share an overall 43% and 41% amino acid identity with c-Myc, respectively (Nau *et al.*, 1985; Kohl *et al.*, 1986; Kato and Dang, 1992). The regions of sequence similarity are primarily within the carboxy terminal and amino terminal regions of the gene, although there is an acidic region which traverses the exon 2 and exon 3 splice junction that is also conserved (Figure 12). For all three *myc* genes there are 2 highly conserved regions within exon 2, which are part of the transcriptional activation domain. Additionally three extensively conserved regions are encoded by exon 3, which encompasses the two dimerization domains and the DNA binding domain. Shared between c-Myc and N-Myc but not present in L-Myc are two additional regions of homology (DePinho *et al.*, 1987b). These two regions consist of a casein kinase II phosphorylation site and a nuclear localization signal. The c-Myc protein can therefore be subdivided into a number of functional domains.

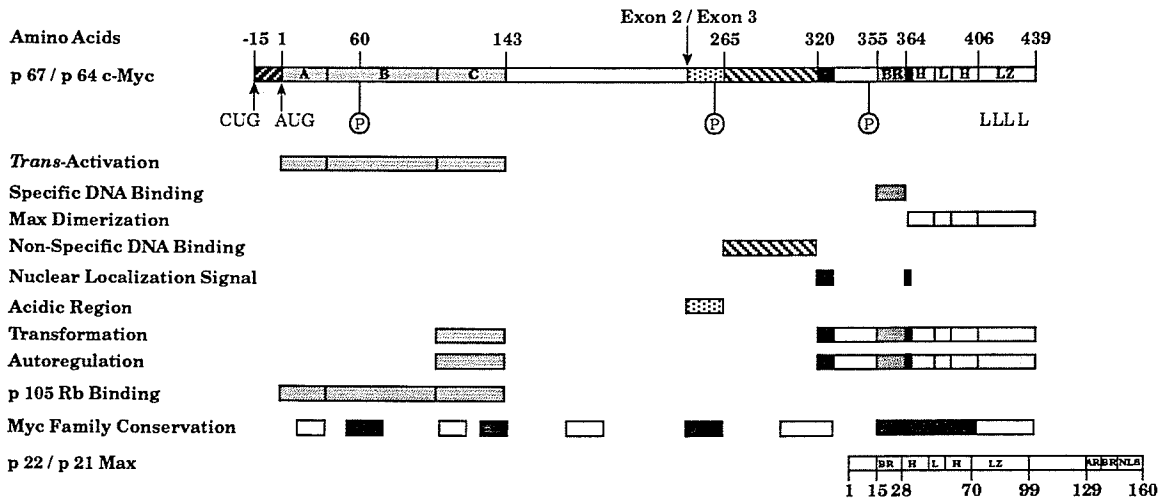


Figure 12. Structure and Features of the c-Myc Protein. The domain structure of the c-Myc protein is shown with the amino acids listed above. The two translation initiation sites (CUG, AUG), the approximate location of the phosphorylation sites (P) and the location of the leucine repeats (LLLL) are shown immediately below the domain structure. The associated functional and conserved regions are aligned with the domain structure of c-Myc and are listed on the left of the diagram. The basic region (BR), helix-loop-helix regions (HLH) and leucine zipper (LZ) are labelled within their respective domains. The *trans*-activation domain is subdivided into regions A, B and C as described in the text. For conserved regions, black and white represent highly conserved amino acid sequence or functional sensitivity to mutation and less pronounced conservation or sensitivity in functional assays, respectively. The structure of the related protein, Max, which dimerizes with the Myc proteins, is shown at the bottom of the figure. Abbreviations are the same as described above, with the addition of the acidic region (AR) and nuclear localization signal (NLS). This figure was derived from Lüscher and Eisenman, 1990; Spencer and Groudine, 1991; DePinho *et al.*, 1991; Cole, 1991; Marcu *et al.*, 1992; Prendergast and Ziff, 1992; Dang, 1992; Kato and Dang, 1992.

Trans-activation Domain

Transcriptional activation domains are defined by their ability to activate transcription when bound to DNA through their associated DNA-binding domain. They are believed to exert their effects by interacting with components of the transcriptional machinery such as the general transcription factors TFIID and TFIIB (Ptashne and Gann, 1990; Lin and Green, 1991; Hahn, 1993). The transcriptional activation domains which have so far been characterized consist of acidic, glutamine rich, proline rich and serine plus threonine rich amino acid sequences (Mitchell and Tjian, 1989; Ptashne and Gann, 1990; Rosenfeld, 1991). In c-Myc, the *trans*-activation domain is located at the amino terminal end, within the first 143 amino acids of the AUG initiated protein (Kato *et al.*, 1989). The amino terminal sequences can be broken into three subregions of amino acids 1-41 (A), 41-103 (B) and 103-143 (C) based on convenient restriction endonuclease sites within the c-*myc* gene (Figure 12). Each of these subregions is capable of activating transcription (Kato *et al.*, 1989). Both region A and B share some similarity to previously described *trans*-activation domains. Region A is glutamine rich and slightly acidic while region B contains a 21 amino acid segment which is highly conserved and contains 33% prolines (Dang, 1991). Within region B there are two potential phosphorylation sites, threonine-58 and serine-62. Replacement of these amino acids with alanine prevented phosphorylation at these sites and resulted in a decrease in Myc-dependent *trans*-activation (Gupta *et al.*, 1993). This suggests that the function of the *trans*-activation domain may be regulated by phosphorylation at these sites. Although perfectly conserved with the other members of the Myc family, region C does not resemble any previously described transcriptional

activation motif (Dang, 1991). This region is one of the domains required for neoplastic transformation and autoregulation (Penn *et al.*, 1990a).

Transforming Domains

There are two assays used for monitoring the transforming activity of an oncogene. The rat embryo fibroblast (REF) co-transformation assay is used to monitor an oncogene's ability to cooperate with *ras* in the transformation of normal cells, while the Rat-1 assay is used to monitor an oncogene's ability to transform immortalized but not transformed cells. The Myc family of proteins can cooperate with Ras to yield transformed foci in the REF assay. Similarly, Rat-1 fibroblasts are neoplastically transformed by transfection with Myc (DePinho *et al.*, 1991). Four functionally distinct regions have been identified using these assays. Amino acids 144-320 are active in the co-transformation assay but nonfunctional in the Rat-1 assay, while amino acids 1-104 are not as essential. The two regions which are indispensable for transformation encompass amino acids 105-143 and 321-439 (Stone *et al.*,1987). Both of these regions have functional significance (Figure 12). The residues 321-439 correspond to the carboxy-terminal portion of Myc and include sequences necessary for nuclear localization, DNA binding and dimerization (Stone *et al.*, 1987; Dang and Lee, 1988; Murré *et al.*, 1989a). Residues 105 to 143 reside within the amino-terminal portion of Myc in a region necessary for transcriptional activation (Kato *et al.*, 1989). These same sequences are required for autoregulation (Penn *et al.*, 1990a; DePinho *et al.*, 1991).

Phosphorylation Sites

Considerable evidence suggests that protein phosphorylation and dephosphorylation plays a key regulatory role in the regulation of the cell

cycle, transcription, and signal transduction (Draetta *et al.*, 1988; Hunter and Karin, 1992; Nishizuka, 1984; Sibley *et al.*, 1988). In particular for transcription factors, phosphorylation has been shown to modulate either the *trans*-activating ability or DNA-binding activity (Berberich and Cole, 1992). The product of the c-*myc* proto-oncogene has been shown to be a nuclear phosphoprotein (Persson and Leder, 1984). It has been suggested that phosphorylation of Myc may be cell cycle-dependent (Alvarez *et al.*, 1991). Significantly, hyperphosphorylation of c-Myc has been observed during the induction of growth arrest (Maheswaran *et al.*, 1991).

Myc is known to be modified by phosphorylation on serine and threonine residues (Hann and Eisenman, 1984; Bister *et al.*, 1987; Lüscher and Eisenman, 1990; Lüscher *et al.*, 1989) but not on tyrosine residues. Phosphorylation has been demonstrated at serine 62 (Figure 12) using $p34^{cdc2}$ (Prendergast and Ziff, 1992;) and MAP kinase (Alvarez *et al.*, 1991), while threonine 58 is phosphorylated by glycogen synthase kinase III (Lüscher and Eisenman, 1990). Threonine 58 may be important for Myc function, since this site is mutated in several retroviral v-*myc* genes (Prendergast and Ziff, 1992). Myc is also a substrate for casein kinase II (CK II) at 2 locations (Figure 12). Amino acid sequences 342-55, which are amino-terminal to the basic region (BR) and conserved in N-Myc and L-Myc, contain potential CK II sites. While amino acids 240-262, which lie adjacent to an acidic region and are only weakly conserved, contain a cluster of CK II sites (Lüscher *et al.*, 1989). Although phosphorylation near the BR has the potential to alter DNA-binding ability, deletion of amino acids 342-357 has no effect on transforming activity *in vitro* (Street *et al.*, 1990).

Nuclear Localization Domains

Nuclear localization sequences were identified by deletion analysis and gene fusion experiments (Stone *et al.*, 1987; Dang and Lee, 1988). A nuclear localization signal (M1) consisting of 9 amino acids is located at amino acids 320 to 328 (Figure 12). This site contains the nonapeptide PAAKRVKLD which is evolutionarily conserved among other c-*myc* genes and related to a similar peptide encoded by the N-*myc* gene (Dang and Lee, 1988). An additional nuclear localization signal (M2), consisting of the amino acids RQRRNELKRSF, is located within the basic region at residues 364 to 374. This peptide, which is also highly conserved, has very limited targeting activity and normally functions as part of the basic region (DePinho *et al.*, 1991). However, in the absence of the first nuclear localization signal this peptide will function to localize c-Myc to the nucleus (Dang and Lee, 1988). Deletion of the M1 peptide yields only weak nuclear targeting, suggesting this peptide contains the dominant nuclear localization signal (DePinho *et al.*, 1991).

During the prophase and metaphase stages of mitosis, c-Myc has been shown to dissociate from condensing chromosomes and become dispersed in the cytoplasm (Winqvist *et al.*, 1984). During telophase and cytokinesis, the later stages of mitosis, c-Myc reassociates with chromatin and very little association with the cytoplasm is seen. Conversely, c-Myc has been shown to associate with the nuclear matrix (Eisenman *et al.*, 1985), however, this association has been disputed (Evan and Hancock, 1985). The first study (Eisenman *et al.*, 1985) demonstrated that 60 to 90% of c-Myc was retained after salt extraction with 0.4M LiCl, 2M NaCl or 0.25M $(NH_4)_2SO_4$ of nuclei depleted of RNA and DNA by micrococcal nuclease and RNase A treatment at 37°C. The second study (Evan and Hancock, 1985) found that the

insolubilization of c-Myc was critically dependent on temperature. Nuclease treatment at 4-35°C allowed c-Myc to be extracted from nuclei with a minimum of 0.2M NaCl while temperatures over 35°C resulted in resistance to high salt extraction. Treatment of these structures with metal ions and chelators or reducing agents was ineffective, however c-Myc was solubilized when the structures were dissolved in SDS or urea (Eisenman *et al.*, 1985; Evan and Hancock, 1985). Recently, an additional study demonstrated that c-Myc association with the nuclear matrix is transient and occurs during the S phase of the cell cycle (Waitz and Loide, 1991). These authors suggested that c-Myc may be a component of the nuclear scaffold.

DNA-Binding Domain

The Myc proteins contain both specific and non-specific DNA-binding domains (Figure 12). The highly conserved non-specific DNA-binding domain is enclosed by amino acids 265 to 318 (Dang *et al.*, 1989; Dang, 1991; Donner *et al.*, 1982; Marcu *et al.*, 1992). This region contains four Ser-Pro-X-X-like sequences, where X is commonly a lysine residue (Dang, 1991). This sequence has been shown to mediate non-specific DNA-binding by histone H1 (Suzuki, 1989; Churchill and Suzuki, 1989). Non-specific DNA-binding has been demonstrated with bacterially expressed and purified c-Myc (Watt *et al.*, 1985). The non-specific DNA-binding region in Myc may serve to facilitate the identification of specific binding sequences (Dang, 1991; Von Hippel and Berg, 1989; Kato *et al.*, 1989).

A stretch of basic amino acids called the basic region (BR) and located amino-terminal to the two Myc dimerization domains is the site of specific DNA-binding (Prendergast and Ziff, 1989). Alignment of the Myc BR with the BR of other helix-loop-helix (HLH) and leucine zipper (LZ) containing proteins uncovers a clear pattern of conserved and non-conserved residues

(Fisher *et al.*, 1991). An α-helical conformation for the BR is suggested by the spacing of three to four conserved residues in the amino-terminal portion of the BR domain. This subdomain consists of an acidic residue, two non-conserved residues, two basic or glutamine residues, two more non-conserved residues and a basic or glutamine residue followed by an asparagine or lysine residue. There is a more highly conserved region of six amino acids near the carboxy-terminal end of the BR which has contiguous spacing. This subdomain consists of one hydrophobic residue and one acidic residue followed by four basic or glutamine residues. Circular dichroism experiments and mutational analysis support the suggestion that the BR of HLH-LZ proteins, including Myc, form an α-helix when bound to DNA. Contacts between the BR and the DNA-binding site were suggested to occur through four of the conserved amino acids on one side of the α-helical face. In addition, it was suggested that the more uniformly conserved subdomain at the carboxy-terminal end of the BR may serve to secure the protein within the major groove of DNA, thus stabilizing and perhaps even nucleating the α-helix formation of the DNA-contact region (Fisher *et al.*, 1993).

Based on binding sites for the BR-HLH family of transcription factors, a putative DNA-binding site for Myc was determined to contain the consensus nucleotides 5'-CA--TG-3' (Murré *et al.*, 1989; Blackwell and Weintraub, 1990). Subsequent experiments utilizing soluble carboxy-terminal Myc fragments, glutathione-S-transferase-Myc chimeric proteins, E12-c-Myc chimeric proteins or purified c-Myc protein revealed that Myc required the sequence 5'-CACGTG-3' for DNA-binding (Blackwell *et al.*, 1990; Prendergast and Ziff, 1991; Kerkhoff *et al.*, 1991; Papoulas *et al.*, 1992). This sequence is identical to the core site bound by USF (Gregor *et al.*, 1990; Marcu *et al.*, 1992). Further support for this sequence as the core Myc DNA-binding site is

the observation that the methylation of the core cytosine, 5'-CA^{m}CGTG-3', disrupts DNA-binding (Prendergast and Ziff, 1991). Systematic variation of the flanking sequences for the USF binding site determined that the sequence 5'-GACCACGTGGTC-3' is a high-affinity Myc binding site (Halazonetis and Kandil, 1991).

The HLH and LZ regions, both dimerization domains, are indirectly required for DNA-binding activity of the BR. Both play a role in positioning the BR in the proper orientation for DNA-binding. When LZ helices are assembled in parallel, the DNA-binding BR of each polypeptide is positioned to form a contact surface that can bind symmetric DNA sequences (Vinson *et al.*, 1989; Lüscher and Eisenman 1990). Similarly, it has been suggested for proteins containing the HLH motif, that the interaction of two of these domains may determine the affinity and possibly the specificity of DNA binding (Lüscher and Eisenman, 1990). Recent experiments demonstrated that the entire LZ interaction can be replaced by an S-S bond without loss of the BR DNA-binding specificity (O'Neil *et al.*, 1990; Talanian *et al.*, 1990). This suggests that these dimerization functions are necessary for the close association of the BR of protein partners. Notably, the specific binding of the synthetically linked BRs to DNA induced the BR to assume a helical conformation (O'Neil *et al.*, 1990; Talanian *et al.*, 1990; Lüscher and Eisenman 1990).

Dimerization Domains

Two general classes of transcription factors have been defined on the basis of their dimerization domains. The Jun and Fos proteins, along with GCN4, the ATF family of proteins and several other transcription factors all possess the leucine zipper dimerization domain (Vinson *et al.*, 1989; Lüscher and Eisenman, 1990). In general, the LZ consists of a 20-30 amino acid

segment with 4-5 leucine residues spaced every 7 residues. This arrangement produces an amphipathic α-helix with a leucine residue situated at every second helical turn (Landschulz *et al.*, 1988) The hydrophobic interface created on one side of the helix is responsible for the formation of a coiled-coil dimerization structure with an opposing LZ arranged in parallel (O'Shea *et al.*, 1989). Dimerization is believed to be stabilized by hydrophobic interactions between leucines of opposing α-helices and inter-helical salt bridging between charged residues positioned more laterally (Landschultz *et al.*, 1988, O'Shea *et al.*, 1989). This dimerization interface is linked to a basic region spaced 6 amino acids amino-terminal to the LZ (Landschulz *et al.*, 1988).

Proteins which contain the helix-loop-helix dimerization domain include MyoD, myogenin, two immunoglobulin κ enhancer-binding proteins and a variety of Drosophila differentiation factors (Davis *et al.*, 1987, Edmunson and Olson, 1989; Murré *et al.*, 1989a; Villares and Cacrera 1987; Alonso and Cabrera 1988; Caudy *et al.*, 1988; Thisse *et al.*, 1988; DePinho *et al.*, 1991). The general structure of the HLH motif consists of 2 amphipathic α-helices of 15 amino acids which are separated by an intervening loop of varying length. This dimerization structure is also preceded by a basic region of approximately 13 amino acids long and rich in arginines (Murré *et al.*, 1989a; Prendergast and Ziff, 1989).

Both the LZ and HLH dimerization motifs are indirectly required for DNA-binding due to their ability to juxtapose two BRs from dimeric protein complexes to form a DNA-binding domain (Landschultz *et al.*, 1988; Murré *et al.*, 1989b). A small subset of proteins having both dimerization motifs includes the Myc family of proteins (Lüscher and Eisenman, 1990), two proteins that bind the immunoglobulin μE3 motif, TFE3 and TFEB,

(Beckmann *et al.*, 1990; Carr and Sharp 1990), USF (Gregor *et al.*, 1990), AP4 (Hu *et al.*, 1990) and more recently the proteins Max (Blackwood and Eisenman, 1991), Myn (Prendergast *et al.*, 1991), Mxi1 (Zervos *et al.*, 1993) and Mad (Ayer *et al.*, 1993). The most extensive sequence homology within the *myc* gene family lies at the carboxy terminus. This region contains the two different dimerization domains and a basic region required for DNA-binding. In the Myc family of proteins the dimerization motifs and the specific DNA-binding domain are aligned; BR-HLH-LZ-COOH (Figure 12). The carboxy-terminal 72 amino acids of c-Myc contain both the HLH and LZ domains. The HLH region is located within amino acids 364-407, while the LZ region is enclosed by amino acid residues 406-439 (Lüscher and Eisenman, 1990; Dang, 1991; Marcu *et al.*, 1992). The HLH and the LZ domains were both required for c-Myc homodimer formation *in vitro* however, these same dimers were not detected *in vivo* (Dang, 1991; Littlewood *et al.*, 1992) and were biologically inactive. These studies suggested that c-Myc function requires heterodimer formation with a partner protein.

Protein Interactions with Myc

The determination that both the HLH and LZ domains were required for homodimer formation induced speculation that the Myc proteins may dimerize with both the HLH and LZ families of transcription factors. Studies of Myc dimerization with the BR-HLH proteins MyoD and E12 determined that Myc was unable to form dimers with either of these proteins. Coincident was the discovery that the LZ proteins Jun and Fos were also unable to dimerize with Myc. Interestingly, Myc can not heterodimerize with the HLH-LZ protein USF, even though the core DNA-binding site, 5'-CACGTG-3', is the same (Murré *et al.*, 1989b; Lüscher and Eisenman, 1990; Marcu *et al.*, 1992). This dimerization specificity is likely due to differences

in the spacing of the HLH and LZ domains of Myc, as compared to other HLH-LZ proteins, which maintains the hydrophobic residue heptad repeat across both the HLH helix II and the LZ to form a single α-helix (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991; Cole, 1991). In light of these observations, the recent discovery of a dimerization partner for Myc was an important break through.

Four Max proteins have been described, encoded by different messenger RNA splice variants. These forms are defined by the absence or presence of a 9 amino acid insert preceding the basic domain (Max1 and Max2, respectively) (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991), and by a carboxy-terminal truncation downstream of the DNA and dimerization domain (ΔMax1 and ΔMax2) (Mäkelä *et al.*, 1992). The human protein Max (Blackwood and Eisenman, 1991) and the mouse equivalent Myn (Prendergast *et al.*, 1991) contain HLH and LZ domains which are spaced identical to the equivalent domains in Myc and are capable of heterodimer formation. The structure of Max is considerably different from Myc in that it does not contain an extensive amino-terminal *trans*-activation domain (Figure 12). Max consists of 160 amino acids, 50% of which is composed of the BR-HLH-LZ regions. A short amino terminal domain of 15 to 24 amino acids precedes the BR and dimerization domains. These regions are followed by a 58 amino acid domain with acidic and serine plus threonine rich stretches (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991; Cole, 1991). There are three potential casein kinase II sites in the amino-terminal domain, and a cluster of five sites exist within the acidic stretch carboxy-terminal to the LZ (Berberich and Cole, 1992). A nuclear localization signal with the consensus sequence PQSRKKLR, has also been localized within the carboxy-terminal 23 amino acids of Max (Kato *et al.*, 1992a).

Max can form heterodimers with c-Myc, N-Myc and L-Myc, and these heterodimers bind 5'-CACGTG-3' 100-fold more efficiently than Myc alone (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991; Wenzel *et al.*, 1991; Kato *et al.*, 1992a). As well, both Myc and Max are nuclear proteins, and heterodimers have been detected in the nucleus (Blackwood *et al.*, 1992). Max is unable to form heterodimers with USF and AP4, two other HLH-LZ proteins (Blackwood and Eisenman, 1991) although Max is able to form homodimers *in vivo* (Amati *et al.*, 1992). Despite having an acidic sequence within the Max carboxy-terminal domain, this region does not seem to have any transcriptional activity (Kato *et al.*, 1992a). Using yeast as an *in vivo* model system *trans*-activation function of Myc-Max heterodimers has been localized to the amino-terminal domain of Myc (Amati *et al.*, 1992).

The BR-HLH-LZ motifs were determined to be necessary and sufficient for association and specific binding of Max with the Myc proteins *in vitro* (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991; Halazonetis and Kandil, 1991) as well as *in vivo* (Amati *et al.*, 1993). Deletions in the Max LZ domain that left the HLH domain intact were sufficient to disrupt binding to c-Myc, and eliminate c-Myc transformation activity. Conversely, deletion of the potential phosphorylation sequences in Max resulted in increased specific DNA-binding activity (Prendergast *et al.*, 1991). Notably, Max is phosphorylated by casein kinase II *in vivo* (Blackwood *et al.*, 1992; Berberich and Cole, 1992) and this post-translational modification inhibits the DNA-binding activity of Max homodimers, but not Myc-Max heterodimers (Berberich and Cole, 1992). Together these results suggest that both the HLH and LZ dimerization domains are required for biological activity of Myc-Max heterodimers and that phosphorylation of Max determines dimer formation.

There are two other proteins which have been described to interact with Myc. The first is a 500 kDa protein called Myc-associated protein or MYAP (Gillespie and Eisenman, 1989). This protein was detected by chemical cross-linking prior to v-Myc protein isolation. The second is the retinoblastoma gene product, p105 Rb (Rustgi *et al.*, 1991). The carboxy-terminal region of p105 Rb has been shown to interact with the amino-terminal *trans*-activation domain of c-Myc and N-Myc. Deletion of a 35 amino acid region in the carboxy-terminus of p105 Rb prevents its association with Myc.

Function of Myc

The c-*myc* proto-oncogene is a master regulator of cell growth and differentiation. The c-Myc protein has been classed as a competence factor (Kelly *et al.*, 1983; Campisi *et al.*, 1984), although more recently other functions have been ascribed to it. The expression of c-*myc* is necessary for cellular proliferation and apoptosis, while the down regulation of *myc* appears necessary for growth arrest and differentiation. The c-*myc* proto-oncogene would therefore appear to be an essential regulator for determining which of these paths a cell will take. The functional characterization of the Myc proteins has been enhanced by the discovery of a dimerization partner. Mutations which alter the ability of c-Myc to contribute to co-transformation, inhibition of differentiation and autoregulation co-localize to the dimerization motifs responsible for interaction with Max (Blackwood *et al.*, 1992; Ohmori *et al.*, 1993). It is therefore likely that an association with Max is necessary for the biological activity of c-Myc and the other Myc proteins.

The Myc family of proteins are differentially expressed with respect to development, differentiation, and neoplasia, while Max is expressed in a variety of cells and tissues (Blackwood and Eisenman, 1991). Although c-*myc*

has been shown to be stimulated by a number of mitogens, *max* mRNA and protein expression are unaffected or only slightly increased with serum stimulation (Blackwood *et al.*, 1992; Berberich *et al.*, 1992; Wagner *et al.*, 1992). Neither Max nor the Myc proteins vary during the cell cycle, however Max is expressed in quiescent cells that do not express c-Myc (Cole, 1991; Blackwood *et al.*, 1992; Wagner *et al.*, 1992). Additionally, the Myc proteins are rapidly degraded, while Max is a highly stable protein (Blackwood *et al.*, 1992; Wagner *et al.*, 1992). Together these observations suggest that the Myc family of proteins will be the rate limiting partners in heterodimer formation with Max. Mitogen stimulation would therefore be expected to induce a shift in the equilibrium from Max homodimers to Myc-Max heterodimers, while the withdrawal of stimulation would influence the opposite shift in the equilibrium (Amin *et al.*, 1993).

It is possible that Max could act as a positively acting co-factor for Myc function and a potentially negatively acting transcription factor in the absence of Myc. The complexity of Myc regulation and the differential regulation of the various members of the Myc family suggest the relative concentrations of Myc could be an important determinant of Myc-Max associations and ultimately, their influence on cell proliferation and behavior. Alternatively, the discovery of two additional proteins, which interact with Max (Ayer *et al.*, 1993; Zervos *et al.*, 1993), suggests a large family of proteins, which interact with Max and/or Myc, are involved in the eventual physiological function of Myc (Figure 13). Both Mxi1 (Zervos *et al.*, 1993) and Mad (Ayer *et al.*, 1993) have been described from the human system and are members of the BR-HLH-LZ family. Like Myc, Mxi1 and Mad form heterodimers with Max, but do not readily form homodimers nor do they interact with other BR-HLH-LZ, BR-HLH, or BR-LZ proteins. Additionally,

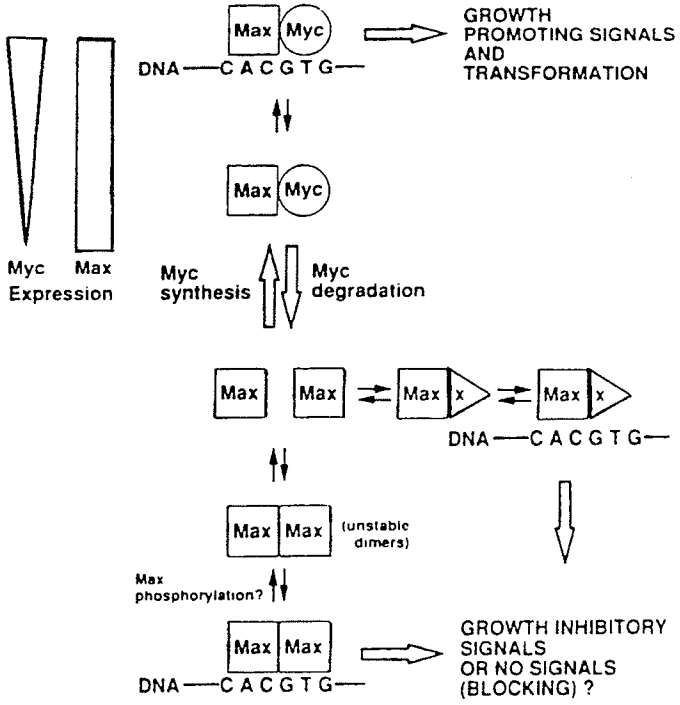


Figure 13. Model for Myc and Max Function. The biological effects of dimer formation between Max and Myc or Max and the repressor proteins (x) Mad and Mxi1 are schematically represented. The Myc/Max interaction produces growth promoting signals, whereas the formation of Mad/Max or Mxi1/Max heterodimers produces growth inhibitory signals. The intracellular concentration of Myc, as determined by its rate of synthesis and degradation, is proposed to be a primary level of regulation for Myc/Max function. The steady state levels of Max appear to be constant, but are as yet undetermined for Mad and Mxi1. This figure was modified from Amati et al., 1993 using information derived from Ayer et al., 1993.

like Myc-Max heterodimers, Mxi1-Max and Mad-Max heterodimers have an increased affinity for the 5'-CACGTG-3' consensus sequence over Max-Max homodimers. Unlike Max homodimers (Berberich and Cole, 1992), the DNA-binding ability of Mad-Max heterodimers is similar to Myc-Max binding (Ayre, *et al.*, 1993) and is unaffected by Max phosphorylation. When complexed with Max, both Mxi1 and Mad oppose the action of Myc in that they repress transcriptional activation. Finally, while the basal expression of Myc decreases during differentiation, the expression of Mxi1 and Max was found to be unchanged (Zervos, *et al.*, 1993). Together, these observations suggest a model where Max acts as a co-factor for Myc function and its availability is regulated by its interaction with Mxi1 and Mad. The subsequent association of Mxi1-Max or Mad-Max heterodimers with Myc-Max binding sites act to further block Myc function. A complete understanding of the function of the Myc family of proteins will therefore require the identification and characterization of all the members of this interacting family of proteins and a detailed analysis of their expression patterns.

Developmental Expression of Myc

The developmental patterns of expression of the *myc* genes to some degree correlate with their expression patterns in various tumors. The c-*myc* gene is expressed in a wide range of developing tissues, and is found associated with malignancies from a number of different tissues types. Conversely, the expression patterns of the N-*myc* and L-*myc* genes are more specific both developmentally and oncogenically (Zimmerman *et al.*, 1986). Although c-*myc* exhibits a generalized pattern of expression, in embryonic development preferential expression is seen in endodermal and mesodermal tissues (Marcu *et al.*, 1992; Pfeifer-Ohlsson *et al.*, 1985; Schmid *et al.*, 1989).

In contrast, the N-*myc* and L-*myc* genes are highly expressed in the early stages of a restricted number of differentiating cell lineages, suggesting their expression correlates more strongly with differentiative rather than proliferative processes (DePinho *et al.*, 1991; Marcu *et al.*, 1992). In general, high c-*myc* expression strongly correlates with proliferation and folding of partially differentiated epithelium. The exception is the primitive ectoderm which is highly proliferative but has very low levels of c-*myc* expression (Downs *et al.*, 1989).

Changing patterns of c-*myc* and N-*myc* expression coincide with the emergence and differentiation of major embryonic structures (Downs *et al.*, 1989). For example, there is high c-*myc* expression in the emerging primitive streak but barely detectable N-*myc* expression. This expression pattern reverses as the primitive streak matures. As well, there is differential expression of the *myc* family members in the developing neuronal, lymphoid and renal systems (Zimmerman *et al.*, 1986; DePinho *et al.*, 1991; Mugrauer and Ekblom, 1991). The c-*myc* gene is expressed in all stages of murine B and T cell development but N-*myc* is only found in the immature stages. In addition, during early neural development N-*myc* and L-*myc* have overlapping patterns of expression in the primitive brain and neural tube. However, as these tissues differentiate into the developing central and peripheral nervous systems, these *myc* genes display differential expression. At later stages of differentiation N-*myc* and L-*myc* are expressed in different regions of the brain (DePinho *et al.*, 1991). Finally, a recent report on the developmental expression of *Xenopus max* found expression of mRNA and protein in the egg, 32-cell stage, and mid-blastula transition (King *et al.*, 1993), however no reports have been published on expression patterns at later stages of development.

The functional role for the differential developmental expression of the *myc*-family of genes is gradually becoming apparent. Preventing the down-regulation of c-*myc* as proliferating myoblasts prepare to differentiate into myocytes inhibits muscle-cell differentiation (Miner and Wold, 1991). As well, embryonic lethality and significant developmental defects result when the N-*myc* (Charron *et al.*, 1992; Stanton *et al.*, 1992) or c-*myc* (DePinho *et al.*, 1991) genes are homozygously disrupted. An extensive comparison of c-*myc* and N-*myc* gene expression during gastrulation, neurulation, organogenesis and late fetal development is given in Stanton *et al.*, 1992.

Myc and the Cell Cycle

The cell cycle is defined as the interval between the midpoints of two sequential mitotic episodes (Howard and Pelc, 1951). It is subdivided into at least four consecutive phases; G_1, S, G_2, and M. The G_1, S and G_2 phases are all part of interphase, while M phase can be further subdivided into prophase, metaphase, anaphase and telophase. The G_1 phase represents the gap in time between mitosis (M phase) and the start of DNA replication (S phase). The G_2 phase represents the interval between the end of S phase when the DNA content has doubled and the onset of cell division which begins with mitosis (Darzynkiewicz, 1983). Late in G_1 is a restriction point R at which point cells are committed to completing the S, G_2 and M phases of the cell cycle (Pardee, 1974). Cells, which have withdrawn from the cell cycle, are said to be quiescent and arrested in a phase called G_0 (Baserga, 1978). Expression of the c-*myc* gene is necessary for transition from this resting stage to the G_1 phase and continued progression through the cell cycle.

The c-*myc* proto-oncogene has been shown previously to be a cell cycle competence gene intimately involved in the proliferative response (Kaczmarek *et al.*, 1985). Often regulation of c-*myc* expression has been

shown to be associated with mitogen stimulated growth (Kelly *et al.*, 1983). An increase in c-*myc* expression is observed during the mitogen stimulated transition of growth arrested cells from G_0 to the G_1 phase of the cell cycle (Kelly, *et al.*, 1983). This stimulation induces a rapid rise in c-*myc* mRNA and protein, peaking at 2 to 4 hours and then decaying to the lower steady state level found in exponentially growing cells (Dean *et al.*, 1986; Dubik *et al.*, 1987; Waters, *et al.*, 1991; Watson *et al.*, 1991). Microinjection of c-Myc protein into quiescent cells elicits a partial stimulus to enter the cell cycle (Armelin *et al.*, 1984), while estrogen activation of a Myc-estrogen receptor chimera is sufficient to trigger reentry into and progression through the cell cycle (Eilers *et al.*, 1991).

In proliferating cells the level of c-*myc* mRNA and protein are invariant throughout the cell cycle (Hann *et al.*, 1985; Thompson *et al.*, 1985), and the steady state level of c-*myc* expression correlates closely with the rate of cell proliferation. Cells, which constitutively express high levels of c-Myc, are unable to become quiescent (Kohl and Ruley, 1987), or to differentiate (Freytag, 1988), spend less time in G_1 (Karn *et al.*, 1989) and have reduced growth factor requirements (Kaczmarek *et al.*, 1985; Sorrentino *et al.*, 1986; Stern *et al.*, 1986). Conversely, cells, which express reduced levels of c-Myc, display a slower growth rate, which correlates with a 3 to 4 hour delay in entry into S phase (Shichiri *et al.*, 1993).

Candidate genes for controlling cell cycle progression include the cyclins, cyclin-dependent kinases and two tumor suppressor genes, p53 and pRb (Kirschner, 1992). Myc could be included in this listing with the recent observation that deregulated expression of c-Myc alters cyclin gene expression (Jansen-Dürr *et al.*, 1993). In particular, cyclin D_1 levels were repressed early in G_1 in Myc-transformed cells. Conversely, constitutive

expression or activation of Myc-ER chimeras led to increased levels of cyclins A and E. The activation of cyclin A expression was concurrent with a growth factor independent association of cyclin A and cdk2 with the cell cycle regulated transcription factor E2F and increased E2F activity.

During S phase E2F is found complexed with cyclin A (Mudryj *et al.*, 1991). In G_1, E2F is associated with pRb, an interaction which represses E2F activity (Chellappan *et al.*, 1991; Bagchi *et al.*, 1991; Heibert *et al.*, 1992). The Rb protein also associates with E1A (Whyte *et al.*, 1988) which forces pRb to dissociate from E2F (Chellappan *et al.*, 1991; Bandara and La Thangue, 1991). As well, Myc and pRb have also been reported to interact (Rustgi *et al.*, 1991). Significantly, the phosphorylation of c-Myc has been observed during the induction of growth arrest, prior to the dephosphorylation of pRb which is associated with quiescence (Maheswaran *et al.*, 1991). Additionally, antisense c-*myc* oligonucleotides repress c-*myc* transcription and block cell proliferation (Yokoyama and Imamoto, 1987; Cooney *et al.*, 1988; Watson *et al*, 1991), while overexpression of pRb results in a block in cell cycle progression from G1 to S phase (Heikkila *et al.*, 1987). Coincidentally, the co-injection of c-Myc but not EJ-Ras, c-Fos or c-Jun, inhibits the ability of pRb to arrest the cell cycle (Goodrich and Lee, 1992). However, Myc does not inhibit the activity of another tumor suppressor, p53 (Levine *et al.*, 1991). Taken together these observations suggest that Myc and pRb may specifically antagonize each other in their efforts to regulate the cell cycle.

A Role in DNA Replication

DNA replication occurs during the synthesis phase of the cell cycle. c-Myc has been suggested to associate with the nuclear matrix during this stage (Waitz and Loid, 1991). The c-Myc protein has been shown to bind to cloned autonomously replicating sequences (ARS) and promote their

replication in transfected cells (Iguchi-Ariga *et al.*, 1987). Notably, an initiation site for chromosomal replication has been mapped 1.6 to 2 kb upstream of the c-*myc* gene (Iguchi-Ariga *et al.*, 1988; Vassilev and Johnson, 1990) and this region can serve as an ARS element (Iguchi-Ariga *et al.*, 1988; McWhinney and Leffak, 1990). Sequences within this region were determined to be responsible for c-Myc replication activity and have been proposed to form a hairpin loop reminiscent of origins of replication (Iguchi-Ariga *et al.*, 1988). This region has also been shown to have transcriptional enhancer activity (Iguchi-Ariga *et al.*, 1988; Ariga *et al.*, 1989). The sequence 5'-TCTCTTA-3', was determined to be essential for c-Myc binding and was found to be necessary for both replication and enhancer activities (Ariga *et al.*, 1989). This sequence differs from the high affinity binding site, 5'-CACGTG-3', described more recently for the transcriptional activity of the Myc proteins (Blackwell *et al.*, 1990; Prendergast and Ziff, 1991; Kerkhoff *et al.*, 1991; Papoulas *et al.*, 1991; Halazonetis and Kandil, 1991). Significantly, another research group was unable to confirm c-Myc binding to this replication/enhancer region and suggested other factors were responsible for the binding activity (Saksela *et al.*, 1991). Alternatively, the binding of a factor to the purine rich element 5'-GGNNGAGGGAGARRRR-3' near the center of the c-*myc* origin of replication has been suggested to induce DNA bending (Bergemann and Johnson, 1992). Together these data suggest that controversey on the role of c-Myc in DNA replication still exists.

Regulation of Gene Expression by Myc

In addition to regulating cyclins A, E and D1 (Jansen-Dürr *et al.*, 1993), Myc has been reported to regulate the expression of a number of other genes, including itself (Penn *et al.*, 1990c). The α-prothymosin gene, which codes for an acidic nuclear protein that may play a role in cell proliferation, is

induced by a Myc-ER chimera in the absence of protein synthesis (Eilers *et al.*, 1991). Additionally, the ornithine decarboxylase gene is *trans*-activated by c-Myc (Bello-Fernandez *et al.*, 1993, Wagner *et al.*, 1993a). Similarly, the transcript and protein levels of two isoforms of Protein Kinase C, α and ß, were both induced by expression of an exogenous c-*myc* gene, but not by activated *ras* (Barr *et al.*, 1991). Conversely, Myc has been shown to suppress the expression of the major histocompatibility complex class I antigen gene (Bernards *et al.*, 1986; Lenardo *et al*, 1989), LFA-1 adhesion receptor genes (Inghirami *et al.*, 1990), class I HLA genes (Versteeg *et al.*, 1988) and several collagen genes (Yang *et al.*, 1991). One mechanism of suppression has been proposed to involve Myc-induced alterations in CTF/NF1 phosphorylation (Yang *et al.*, 1993). The 12 base pair c-Myc consensus binding site has been located within regulatory regions of the c-*sis* (5' enhancer), *H19* (3' enhancer), *nucleolin* (first intron), and *U3B RNA* (5'-flanking) genes (Halazonetis and Kandil, 1991), although a direct role of c-Myc in the regulation of these genes has not been shown. However, the *p53* promoter contains the 5'-CACGTG-3' motif, requires this sequence for full activity and is bound by Myc-Max heterodimers (Reisman *et al.*, 1993).

Myc overexpression, in the presence of Max, results in the transcriptional activation of a reporter gene containing the nucleotide sequence 5'-CACGTG-3'. This activity requires the dimerization and *trans*-activation domains of Myc to be intact. In contrast, the overexpression of Max, in the presence of Myc, represses transcriptional activity and requires the DNA-binding domain of Max (Kretzner *et al.*, 1992; Amin *et al.*, 1993; Gupta, *et al.*, 1993). These data suggest a model in which high levels of Myc in proliferating or stimulated cells form heterodimers with constitutively expressed Max and together these proteins bind 5'-CACGTG-3' containing

elements to activate transcription (Figure 13). Conversely, in growth arrested and differentiated cells the steady state level of Myc is low, leaving Max to form homodimers, which then repress transcription. Furthermore, phosphorylation of Max inhibits its ability to bind DNA, suggesting an additional mode of regulation (Berberich and Cole, 1992).

Myc and Apoptosis

Mammalian apoptosis, or programmed cell death, is an active cellular process requiring gene activation and protein synthesis. Opposing proliferation, apoptosis is a normal phenomenon of development. Apoptosis is characterized by distinctive morphological changes such as nuclear condensation and DNA fragmentation, and cells undergoing this process are disposed of by phagocytosis (Wyllie *et al.*, 1980; Williams, 1991). The removal of apoptotic cells occurs within one hour of the onset of this process (Ellis *et al.*, 1991), therefore few dying cells are ever visible even in the presence of extensive cell death (Perry *et al.*, 1983).

The involvement of c-Myc in apoptosis suggests that there is a tight association between cell division and cell death. This association seems to be regulated by the presence or absence of what are called survival factors (Ffrench-Constant, 1992). In the presence of serum, Myc-expressing fibroblasts survive and divide normally. However, upon serum withdrawal, metabolite depletion or treatment with antiproliferative drugs, these same cells die by apoptosis, with the extent of death correlating with the level of Myc expression (Evan *et al.*, 1992). Furthermore, Myc induces a rapid initiation of apoptosis in cells arrested either in G_0, G_1 or S phase, suggesting that the point in the cell cycle is irrelevant to the activation of the programmed cell death pathway.

The deregulated expression of c-*myc* is extremely common in tumor cells suggesting its acquisition may be an essential step towards oncogenesis (Hann and Eisenman, 1984; Yokota *et al.*, 1986). However, since uncontrolled expression of c-*myc* by itself would be lethal in any situation where growth conditions are limiting, successful proliferation likely requires the active suppression of apoptosis. Thus apoptosis normally provides a mechanism for preventing uncontrolled proliferation. Support for this model was found with the observations that activated Bcl-2 or ectopic expression of Bcl-2, a putative G protein (Haldar *et al.*, 1989), can suppress c-Myc induced apoptosis but does not effect the proliferative effects of c-Myc (Bissonnette *et al.*, 1992; Fanidi *et al.*, 1992; Wagner *et al.*, 1993b). Supporting the transcriptional regulatory function of Myc are the observations that the DNA-binding, dimerization and *trans*-activation domains of c-Myc are also required for induction of apoptosis (Evan *et al.*, 1992).

Transforming Activity

The transforming activity of some proto-oncogenes requires specific mutations that alter the normal function of the protein. One such example is the *ras* oncogene whose product, p21, has altered GTPase activity when point mutations occur within the guanine nucleotide binding site (DeVos *et al.*, 1988). The transforming activity of c-*myc* does not require such alterations since aberrant expression of a normal coding sequence is sufficient for tumorigenesis (Lee *et al.*, 1985; Kato and Dang, 1992). None of the Myc proteins are able to transform primary cells alone but c-Myc, N-Myc and L-Myc have all been shown to cooperate with an activated *ras* oncogene in transformation (Land *et al.*, 1983; Ruley, 1983; Yancopoulos *et al.*, 1985; Schwab *et al.*, 1985; Birrer *et al.*, 1988). Conversely, overexpression of c-Myc from a retroviral vector or Myc-ER chimera induces transformation of

immortalized fibroblasts (Small *et al.*, 1987; Eilers *et al.*, 1989). The transforming activity of the Myc proteins requires the integrity of the DNA-binding, dimerization and transcriptional activation domains (Dang, *et al.*, 1989; Min and Taparowsky, 1992). Interestingly, B-Myc was recently reported to inhibit both the transcriptional and transforming activities of c-Myc (Resar *et al.*, 1993). It has therefore been suggested that c-Myc's transforming activity is dependent on its ability to act as a transcriptional regulator.

Although Max has no transforming activity itself, co-expression of Max and Myc with activated Ras increased Myc transforming activity (Prendergast *et al.*, 1991; Blackwood and Eisenman, 1991; Lüscher and Eisenman, 1990). Furthermore, the ratio of Myc-Max heterodimers and Max homodimers was found to be important to the biological function of Myc (Prendergast *et al.*, 1992). Co-transfection of low levels of Max stimulated Myc transformation activity, while high levels suppressed this function, suggesting a conversion from Myc-Max heterodimers to Max homodimers. As well, the action of Max could be regulated by mutations in the phosphorylation sites of both the amino and carboxy-terminal sequences (Prendergast *et al.*, 1992). In contrast, mutation at residues threonine-58 or serine-62 of Myc did not alter transformation efficiency, although phosphorylation at these sites was required for transcriptional activation of gene expression (Gupta *et al.*, 1993). Another study (Mäkelä *et al.*, 1992) concluded that Max suppressed Myc co-transformation, while a truncated Max (ΔMax) lacking 62 carboxy-terminal amino acids, including the putative nuclear localization signal, enhanced transformation. Finally, although Myc and Max mutants alone do not cooperate with *ras*, pairs of complementary Myc and Max mutants induce large numbers of transformed foci (Amati *et al.*,

1993). Some of these studies seem to contradict each other and suggest that the functional regions of Max need to be more finely defined.

The Myc family of proteins therefore have a number of important cellular functions, including a role in regulating development, apoptosis and the cell cycle. These nuclear proteins most likely exert these functions through the transcriptional activation of a variety of responsive genes by binding the 5'-CACGTG-3' element. The activity of c-Myc is regulated in part by heterodimer formation with the related protein, Max. The formation of these dimers is regulated both by the level of c-Myc in the cell and the presence of repressor proteins capable of forming heterodimers with Max. The overexpression of c-*myc* would result in a shift in the equilibrium of these nuclear proteins which would favor Myc-Max heterodimer formation. This would lead to inappropriate gene activation, which is the essence of the transforming activity of the Myc proteins.

Mechanisms of Oncogenic Activation

There are a number of alterations which have been described for the activation of proto-oncogenes into oncogenes. These include point mutations, proviral insertion, chromosomal translocation and gene amplification. The transforming mechanisms of proviral insertion, chromosomal translocation and gene amplification were first discovered through studies of the c-*myc* locus (reviewed in Cole, 1986). Each of these rearrangements leads to constitutive or elevated c-*myc* expression, most often due to the loss of the appropriate controls which tightly regulate *myc* expression. The mechanism of activation of N-*myc* and L-*myc* is usually through gene amplification (DePinho *et al.*, 1992). The inappropriate expression of the normal *myc* genes is therefore capable of contributing to neoplasia (Marcu *et al.*, 1992).

The c-*myc* oncogene has been associated with a variety of malignancies of various cell types (Hann and Eisenman, 1984; Yokota *et al.*, 1986), while the N-*myc* and L-*myc* oncogenes have a more limited expression. N-*myc* activation is restricted to tumors of neural origin or early differentiated phenotypes such as neuroblastoma, retinoblastoma, embryonal carcinoma and small cell lung carcinoma (Kohl *et al.*, 1983; Schwab, *et al.*, 1983; Lee *et al.*, 1984; Jakobovits, *et al.*, 1985; Nau *et al.*, 1986; DePinho *et al.*, 1991). L-*myc* has been implicated only in the development of small cell lung carcinoma (Nau *et al.*, 1985; Zimmerman *et al.*, 1986; Marcu *et al.*, 1992).

Provirus Insertions

The first example of insertional mutagenesis in malignancy was the disruption of the c-*myc* locus by the avian leukosis virus (Hayward *et al.*, 1981). This virus is a replication competent retrovirus which lacks any oncogenic sequences (Kato and Dang, 1992). It has been found to insert upstream of the c-*myc* gene coding sequences in either orientation or downstream of the coding sequences creating a c-*myc* provirus hybrid transcript (Payne *et al.*, 1982). This insertional mutagenesis results in increased steady state levels of c-*myc* gene expression (reviewed in DePinho *et al.*, 1991; Kato and Dang, 1992). When the proviral LTR is inserted in a positive orientation upstream of c-*myc*, it is speculated that the viral LTR disrupts normal c-*myc* regulation and enhances expression. Insertion of the strong viral enhancer in the opposite orientation or downstream of the c-*myc* coding sequence may also disrupt or override the endogenous regulatory regions (Marcu *et al.*, 1992).

Amplification

The first demonstration of gene amplification in tumor cells was discovered through studies of the c-*myc* locus (reviewed in Cole, 1986). Amplification involves an increased gene copy number of a normal *myc* gene, rather than structural alterations. Elevated expression of the amplified gene is therefore due to the multiple copies rather than deregulation of expression. Consistent with this hypothesis is the observation that amplified c-*myc* in HL60 cells can be suppressed by inducers of differentiation suggesting its expression is not deregulated (Cole, 1986). The amplification of *myc* genes correlates well with the more aggressive forms of malignancies (Yokota *et al*, 1986). The amplification of c-*myc* is associated with breast carcinoma, small cell lung carcinoma, gastric adenocarcinoma, colon carcinoma, glioblastoma and a variety of leukemias (Yokota *et al.*, 1986; Escot *et al.*, 1986; Marcu *et al.*, 1992). The amplification of N-*myc* is detected in the most undifferentiated neuroblastomas (Kohl *et al.*, 1983, Schwab *et al.*, 1983), retinoblastoma (Lee *et al.*, 1984), small cell lung carcinoma (Nau *et al.*, 1986) and other tumors of neuronal and neuroendocrine origin (DePinho *et al.*, 1991). In contrast to the other *myc* genes, L-*myc* amplification has been observed only in small cell lung carcinomas (Nau *et al.*, 1985; Zimmerman *et al.*, 1986; Marcu *et al.*, 1992).

Translocations

The c-*myc* oncogene was the first activated gene to be associated with chromosomal translocations (reviewed in Cole, 1986). Reciprocal translocations between the c-*myc* locus and the immunoglobulin (Ig) loci were seen in B and T cell leukemias, Burkitt's lymphomas and various other lymphoid malignancies (reviewed in Cole, 1986; DePinho *et al.*, 1991; Spencer

and Groudine, 1991; Marcu *et al.*, 1992). Most of these translocations involve the Ig heavy chain switch regions or regions within the Ig heavy (H) or light (L) chain variable regions. For example, the c-*myc* gene is translocated into the JH (variable region) locus in early B cell tumors, the μ heavy chain switch regions in B cell tumors and the downstream switch regions in plasma cell tumors. The translocation events in these tumors correlate well with the actively rearranging locus associated with the particular B cell stage (DePinho *et al.*, 1991; Spencer and Groudine, 1991).

In Burkitt's lymphoma (BL) reciprocal translocation of the human c-*myc* locus on chromosome 8 with one of the Ig loci on chromosome 14, 2 or 22 is consistently seen (Spencer and Groudine, 1991). The reciprocal translocation t(8;14) is the most commonly described for BL (Pelicci *et al.*, 1986). This rearrangement translocates the c-*myc* coding sequences to one of the various switch regions of the IgH locus in a head to head orientation. In addition, this translocation predisposes transgenic mice to develop lymphomas (Adams *et al.*, 1985). The less common t(2;8) and t(8;22) variant translocations involve rearrangement of the c-*myc* gene, with a breakpoint 3' to the coding sequences, in a head to tail orientation with the IgL chain loci, κ and λ, respectively (Spencer and Groudine, 1991). Analogous translocations to the T-cell receptor locus have also been described for T-cell malignancies (Erikson *et al.*, 1986). With respect to the c-*myc* gene, the translocations can be categorized into three distinct classes; truncations or mutations that involve the 5' coding region, the 5' proximal region, or regions distal to the c-*myc* locus (DePinho *et al.*, 1991). The majority of translocations that have breakpoints 3' of the c-*myc* coding sequences have point mutations, insertions or deletions within 5' regulatory regions, exon 1, or intron 1 (Rabbitts *et al.*, 1983; 1984; Cesarman *et al.*, 1987; Showe and Croce, 1987; Zajac-Kaye *et al.*,

1988; Spencer *et al.*, 1990; Spencer and Groudine, 1991; Marcu *et al.*, 1992). In particular, mutations or truncation of exon 1 is a characteristic feature of translocated c-*myc* alleles in BL which alters or removes DH sites (Rabbitts *et al.*, 1983, 1984; Siebenlist *et al.*, 1984; Dyson and Rabbitts, 1985; Cesarman *et al.*, 1987; Spencer and Groudine, 1991).

The non-translocated c-*myc* allele is transcriptionally silent in most Burkitt's lymphoma cells (Leder *et al.*, 1983; ar-Rushdi *et al.*, 1983), with the Raji and Manca cell lines being two of the exceptions (Rabbitts *et al.*, 1984; Bentley and Groudine, 1986b; Eick *et al.*, 1990). In the Manca cell line the DNase I hypersensitive site II_2 is maintained in the unrearranged allele and transcripts initiated from the associated P_0 promoter are detected in the absence of P_1 and P_2 initiated transcription (Bentley and Groudine, 1986b). In the Raji cell line P_0 and P_3 initiated transcripts can be detected from both the translocated and the normal c-*myc* allele, while P_1 and P_2 RNAs are expressed only from the aberrant chromosome (Eick *et al.*, 1990). Conversely, in all other BL cells, transcripts from the translocated allele are elevated or constitutively synthesized. Significantly, three of the c-*myc* DNase I hypersensitive sites (II_1, II_2, III_1) are lost in the non-translocated allele but maintained in the translocated allele (Siebenlist *et al.*, 1984, 1988; Kakkis *et al.*, 1986; Dyson and Rabbitts, 1985). Additionally, these same DH sites are maintained in quiescent cells which have little transcription but are still capable of being induced (Siebenlist *et al.*, 1988; Spencer and Groudine, 1991). Similarly, Ig variable regions are insensitive to DNase I and transcriptionally silent until rearranged next to the constant region where they become nuclease sensitive and transcriptionally active (Spencer and Groudine, 1991). Together these observations suggest the non-translocated c-*myc* allele is normally repressed in these cells, while the translocated allele

is activated by its association with a transcriptionally active chromatin region. Somatic cell hybrid experiments support the notion that the cellular physiology and state of differentiation when the translocation event occurs are important factors in determining allele expression (Spencer and Groudine, 1991). Conversely, protein synthesis is not required for repression of the non-translocated c-*myc* allele suggesting autorepression is not occurring (Nishikura and Murray, 1988).

It is speculated that mutations arising in translocated c-*myc* alleles are a result of the close association with the Ig loci (Rabbitts *et al.*, 1983). These mutations have been shown to cause shifts in promoter usage (Bentley and Groudine, 1986b; Spencer and Groudine, 1991) and loss of the block to transcriptional elongation (Cesarman *et al.*, 1982; Spencer *et al.*, 1990). In normal immature B cells both c-*myc* alleles are in an active chromatin domain and initiate transcription. Additionally, the elongation block at exon 1 is active and the promoter ratio of P_1 to P_2 is normal, with P_2 initiating transcription 5 to 10 times more often than P_1. However, in translocated alleles retaining exon 1 P_1 initiated transcripts occur up to 4 times as often as those initiating from P_2 (Spencer *et al.*, 1990). With B cell maturation both c-*myc* alleles become repressed at the levels of chromatin and transcription initiation while both Ig alleles remain active. Translocation events, which truncate c-*myc* exon 1 or have breakpoints upstream of P_1, have been suggested to disrupt binding of proteins to regulatory sequences (Siebenlist *et al.*, 1984; Dyson and Rabbitts, 1985). Loss of c-*myc* regulatory function is further supported by the variety of mutations detected in the 5' flanking (Szajnert *et al.*, 1987) and exon 1 sequences (Battey *et al.*, 1983; Rabbitts *et al.*, 1983; 1984; Pelicci *et al.*, 1986; Cesarman *et al.*, 1987; Szajnert *et al.*, 1987). In particular, mutations near the 3' border of exon 1 are

consistently detected (Battey *et al.*, 1983; Rabbitts *et al.*, 1984; Pelicci *et al.*, 1986; Cesarman *et al.*, 1987; Szajnert *et al.*, 1987). Additionally, the 5' region of intron 1 is also subject to mutation and has been determined to be a site for protein binding (Zajac-Kaye *et al.*, 1988). It has been suggested that mutations in the 3' region of exon 1 and the 5' region of intron 1 may be involved in the disruption of the transcriptional elongation block (Cesarman *et al.*, 1987).

The co-localization of the c-*myc* coding sequences with one of the described IgH enhancer regions occurs in nearly all of the BLs described (Pettersson *et al.*, 1990). Additionally, other Ig enhancers and locus control regions are likely associated with the translocated c-*myc* gene (Spencer and Groudine, 1991). In the absence of these regulatory elements truncation of exon 1 and 5' regulatory sequences is insufficient for tumor development in transgenic mice (Adams *et al.*, 1985). Similarly, movement of translocated c-*myc* alleles lacking Ig regulatory sequences to a different cellular environment restores the elongation block and the normal promoter ratio (Spencer *et al.*, 1990; Richman and Hayday, 1989). Conversely, the addition of Ig enhancer sequences (Ck or Cμ) to truncated c-*myc* genes results in the development of lymphosarcomas in transgenic mice (Adams *et al.*, 1985) and high level expression in plasmacytoma cells (Feo *et al.*, 1986).

The constitutive or overexpression of the normal c-Myc protein is sufficiently oncogenic to contribute to the transformation of cells (Battey *et al.*, 1983; Lee *et al.*, 1985; Kato and Dang, 1992). Although protein coding regions are usually wild-type, occasionally mutations in the c-*myc* protein-coding sequence have been detected in BL cells (Rabbitts *et al.*, 1983). It is supposed that constitutive expression of c-*myc* is necessary for transformation thus the coding sequences accumulate few mutations, while

regulatory sequences, which bind negative effectors, become targets for mutations (Spencer and Groudine, 1991). Expression of c-Myc from a non-AUG codon near the 3' end of exon 1 is absent in some BL cells (Hann *et al.*, 1988). Gene rearrangements and mutations, which disrupt this p67 translation initiation site, alter the p64/p67 ratio. The significance of abnormal protein ratios is still unknown (Spencer and Groudine, 1991).

Some of the common features of translocation therefore include; specific translocations with immunoglobulin loci, the expression of specific B cell phenotypes, exon 1 mutations, repression of the non-translocated c-*myc* allele, promoter shifts and disruption of elongation block in those alleles retaining exon 1 (Spencer and Groudine, 1991). Translocation of c-*myc* into an active chromatin domain is considered to be essential but it is not sufficient for transformation. The gene must also be non-regulatable through the loss or disruption of regulatory sequences and the Ig enhanced initiation of transcripts from P_1, which is not subject to elongation block. It is therefore supposed that translocation leads to disruption of regulatory regions which activate normally silent c-*myc* alleles or alter the protein's normal expression pattern. Although there have been observations of protein coding mutations, the overwhelming majority of mutations interfere with the genetic control of c-*myc* expression rather than altering protein function. The two major hypotheses for the altered regulation of c-*myc*, which have emerged from these studies, encompass the disruption of positive or negative regulatory regions of the c-*myc* gene itself and the imposition of regulatory control onto c-*myc* by the Ig locus. The predominant difference in models for regulatory disruption by c-*myc* translocation are whether negative or positive control elements have been altered. In the final analysis it is likely that alterations of both positive and negative regulatory sequences complement the action of

Ig control regions to yield a final mechanism for translocation induced c-*myc* deregulation.

The transforming mutations of c-*myc*, which occur through proviral insertion, amplification or translocation, ultimately lead to the inappropriate expression of an otherwise normal transcript. Disruptions to the regulatory regions of c-*myc* and rarely the coding sequences of the gene are the result of some of these transforming mechanisms. The oncogenic activation of the c-*myc* gene therefore disrupts the delicate balance between the c-Myc protein, Max and the other nuclear proteins involved in directing the formation of Myc-Max heterodimers. The observation that DH site II_2 is maintained in the unrearranged c-*myc* chromatin of some BL cell lines and that initiation from the P_0 promoter of this allele produces normal c-*myc* transcripts suggests this region may be involved in regulating the oncogenic activity of c-*myc*.

Regulation of c-*myc* Expression

Gene expression can be regulated at transcriptional as well as post-transcriptional and post-translational levels (reviewed in Darnell, 1982; Falvey and Schibler, 1991). Mechanisms for transcriptional regulation include the control of transcript initiation and transcript elongation. Post-transcriptional gene regulation includes mechanisms for controlling pre-mRNA processing, mRNA transport to the cytoplasm, mRNA stability and translation initiation and elongation. The regulation of protein modifications, protein dimerization and protein stability are all mechanisms for controlling gene expression post-translationally. Dimerization and phosphorylation are post-translational regulatory mechanisms which have been previously discussed for the Myc proteins (see section on c-Myc Protein).

Transcriptional and post-transcriptional regulation of c-*myc* gene expression has also been demonstrated. Post-transcriptional mechanisms have been implicated in the rapid depletion of c-*myc* mRNAs in the later stages of growth factor response, while transcript initiation and elongation have been respectively implicated in the induction and repression of c-*myc* expression (reviewed in Spencer and Groudine, 1991; Marcu *et al.*, 1992).

Post-transcriptional Regulation

Control of mRNA Stability

The steady-state level of c-*myc* mRNA reflects both its rate of synthesis and rate of decay (reviewed in Piechaczyk *et al.*, 1987). Therefore, increased mRNA lability will result in a faster response to a change in transcription initiation. Both increased and decreased mRNA stability contribute to changes in c-*myc* levels in response to alterations in the cellular environment (reviewed in Spencer and Groudine). The normal half-life of c-*myc* transcripts ranges from 10 to 60 minutes. Although transcripts initiated from the P_1 and P_2 promoters appeared to have similar half-lives (Broome *et al.*, 1987), those lacking polyadenylate tails were more stable, with the half-life of P_1 longer than P_2 (Swartwout and Kinniburgh, 1989). Conversely, full length transcripts initiating at P_0 were more stable than those initiating from P_1 or P_2, although most P_0 initiated transcription results in premature termination and these mRNAs were unstable (Bentley and Groudine, 1986b). Additionally, transcripts initiating at P_3 in cells lacking exon 1 were more stable than normal c-*myc* mRNAs (Eick *et al.*, 1990).

Exon 1 of the c-*myc* gene was hypothesized to contribute to the instability of c-*myc* mRNAs (Pei and Calame, 1988). Significantly, truncated transcripts lacking the first exon accumulated to higher steady state levels

(Piechaczyk *et al.*, 1987). Conversely, an adenine plus uridine (A+U) rich region in the 3' untranslated portion of c-*myc* transcripts has been proposed to confer transcript stability (Jones and Cole, 1987). Although mutations introduced here did not alter transcript half-life (Bonnieu *et al.*, 1990), a ribonucleoprotein found to bind in this region contributed to mRNA stability (Brewer, 1991). More recently, an RNA fragment encompassing the carboxy-terminal coding strand of c-*myc* has been described to bind a 75 kDa protein which appeared to protect the transcript from degradation (Bernstein *et al.*, 1992).

Control of Protein Synthesis

The c-Myc protein is very unstable, with a half-life of about 25 minutes (reviewed in Spencer and Groudine, 1991). It has been observed that c-Myc protein levels generally reflect mRNA levels and that transient increases after stimulation to enter the cell cycle are followed by uniform translation, stability and modification efficiencies. In spite of these observations, the presence of alternative translation initiation sites, which are highly conserved, suggests there may be regulation of translation initiation (Hann *et al.*, 1988). Significantly, the ratio of the two c-Myc proteins p64 and p67 has been shown to vary between cell lines and was often altered in cells with translocated c-*myc* sequences (Hann and Eisenman, 1984; Hann *et al.*, 1988). Similarly, alterations in the post-transcriptional regulation of c-Myc stability have been observed after stimulation of cells to differentiate (Wingrove *et al.*, 1988; Spotts and Hann, 1990).

Modulation of c-*myc* mRNA translation efficiency by 5' untranslated sequences was proposed to occur through the formation of a stem-loop structure within intron 1 (Saito *et al.*, 1983). This was supported by the observation that translation efficiency could be modulated *in vitro* by the

deletion of 5' and 3' noncoding sequences (Darveau *et al.*, 1985). Conversely, translocations of c-*myc* which have a break point within intron 1 (Hann and Eisenman, 1984; Ramsay *et al.*, 1984) and *in vivo* transfection studies (Butnick *et al.*, 1985) have not demonstrated any control over c-*myc* translation.

Transcriptional Regulation

Control of Transcript Elongation

The major regulatory mechanism for the downregulation of c-*myc* associated with induced differentiation is a block to transcript elongation (reviewed in Spencer and Groudine, 1991; Marcu *et al.*, 1992). This mechanism is flexible in that it allows for the modulation of transcription without requiring the assembly and disassembly of the transcriptional apparatus. The c-*myc* gene was the first example of an eukaryotic cellular gene regulated by premature transcript termination (Bentley and Groudine, 1986a; Eick and Bornkamm, 1986; Nepveu and Marcu, 1986). Elongation block of human c-*myc* has been associated with P_2 initiated transcription, as P_1 initiated transcripts have not been demonstrated to succumb to premature termination (Spencer *et al.*, 1990). A c-*myc* transcription attenuation site has been mapped using heterologous promoters to a 95 bp sequence located 35 bp 5' of the exon 1 and intron 1 boundary (Bentley and Groudine, 1988). Pausing of transcription at this site upon cellular differentiation is associated with an increased sensitivity of DH site IVb to nucleases relative to the other DH sites (Bentley and Groudine, 1986a). Transcript initiation by RNA polymerase III is completely blocked at this region, while RNA polymerase II is only partially blocked and can transcribe past these sequences (Chung *et al.*, 1987). Conversely, a transcription attenuation site has been observed for

the P_1 promoter in the mouse c-*myc* gene and is located between the two promoters (Wright *et al.*, 1991).

Mutations within a 70 bp region located at the 3' border of the first exon were detected in all three human c-*myc* translocation types (8:14, 8:22, 8:2) and contributed to loss of the block to transcriptional elongation (Cesarman *et al.*, 1987). Significantly, a shift in promoter usage from P_2 to P_1 occurred when exon 1 sequences of translocated c-*myc* remained intact, suggesting exon 1 read through from P_1 (Spencer and Groudine, 1991; Marcu *et al.*, 1992). In the mouse, c-*myc* chromatin structure near the exon 1-intron 1 border is differentially accessible to restriction enzymes (Tourkine *et al.*, 1989). Accessibility is dependent on the integrity of the gene, active transcription and the existence of a strong block to transcriptional elongation. Additionally, a 20 bp region in intron 1 has been designated MIF and identified in human c-*myc* as a 138 kDa nuclear phosphoprotein binding site which is mutated in 5 out of 7 BL cell lines (Zajac-Kay *et al.*, 1988; Zajac-Kaye and Levens, 1990). This binding site is similar to one of the regions identified in the mouse c-*myc* sequence and corresponds to the human c-*myc* sequence, 5'-AGAGTAGTTATGGTAACTGG -3'.

In addition to signals near the 3' end of exon 1, promoter-specific effects on the premature termination of transcription have been reported (reviewed in Spencer and Groudine, 1991; Marcu *et al.*, 1992). The *in vivo* pausing of RNA polymerase has been mapped in human c-*myc* to sequences upstream of +47 relative to P_2 or +221 relative to P_1 (Krumm *et al.*, 1992). The region +3 to +25 relative to P_2 or +177 to +199 relative to P_1, which is completely conserved in the mouse gene, was predicted to form a stem loop structure. Additionally, the deletion or mutation of the ME1a1 sequence, 5'-GAGGGAGGG-3', located between the P_1 and P_2 transcription initiation

sites has been shown to prevent premature termination while also reducing P_2 initiated transcripts (Miller *et al.*, 1989b; Dufort *et al.*, 1993). Furthermore, the addition of this protein binding site to a heterologous promoter was shown to increase transcriptional block (Dufort *et al.*, 1993).

Control of Transcript Initiation

Induction of c-*myc* mRNA upon mitogen stimulation is largely regulated by an increase in transcript initiation (Marcu *et al.*, 1992). The c-*myc* gene supports transcript initiation from the same site by both RNA polymerase II and III, although initiation by RNA polymerase III does not require elements distal or proximal to the TATA box for maximal activity (Chung *et al.*, 1987; Sussman *et al.*, 1991). It has been suggested that the four c-*myc* promoters (P_0, P_1, P_2, and P_3) are independently regulated and differentially affected by c-*myc* deregulation in BL (Marcu *et al.*, 1992; Yang *et al.*, 1986). In particular, differential use of the P_1 and P_2 promoters, which was mitogen dependent, has been observed from the normal c-*myc* gene (Broome *et al.*, 1987). Additionally, transcripts initiating from the P_0 and P_3 promoters in unrearranged alleles were detected in the absence of P_1 and P_2 activity, while the reverse was true of the rearranged allele (Bentley and Groudine, 1986b; Eick *et al.*, 1990). Transcript initiation from P_1 and P_2 in the unrearranged allele was specifically downregulated in these cells. Additionally, when c-*myc* was truncated by breakpoints within intron 1, transcripts initiated from cryptic promoters in intron 1 (ar-Rushdi *et al.*, 1983; Bernard *et al.*, 1983; Hayday *et al.*, 1984), which was consistent with use of the P_3 promoter (Eick *et al.*, 1990). Significantly, the sequence, 5'-GGGTGGG-3', flanked by two Sp1 sites near the c-*myc* intron 1 and exon 2 boundary has been suggested to regulate transcript initiation from P_3 (Eick *et al.*, 1990).

The expression of c-*myc* can be potentiated by protein synthesis inhibitors suggesting c-*myc* is controlled by labile negative factors (Kelly *et al.*, 1983; Dean *et al.*, 1986). In particular, the possibility of labile proteins involved in regulating transcript initiation from the P_0 promoter is suggested by the increased sensitivity to cycloheximide, an inhibitor of protein synthesis, relative to the other three promoters (Eick *et al.*, 1990). Removal of 46 bp from the 3' end of the first exon results in a decrease of *myc* expression and P_2 activity (Yang *et al.*, 1986). Additionally, a sequence, 5'-GTTGGAAA-3', located 6-14 bp 5' of the exon 1 and intron 1 junction (Stanton *et al.*, 1983), bears some resemblance to an Sp1 site, however the strong position and orientation dependence of this site in c-*myc* suggest it does not bind Sp1. Although this region has been described as a positive modulator, its effect is mostly on regulating the elongation of initiated transcripts from P_2, since attenuation of this process has been mapped to this region (Bentley and Groudine, 1988) and mutations in this region have been shown to disrupt elongation (Cesarman *et al.*, 1987).

The disappearance of the promoter specific DH sites is indicative of the repression of transcriptional activity (Bentley and Groudine, 1986a; Siebenlist *et al.*, 1988). In particular the DH site II_2 significantly decreases after 24 hours of exposure of HL60 cells to differentiating agents, while decreases in sites III_1 and III_2 were not seen for another 24 hours. These changes temporally correlated with a loss of transcription initiation (Siebenlist *et al.*, 1988). The repression of c-*myc* has been proposed to be an autoregulatory phenomenon which is mediated by a repressor interacting within the first exon (Leder *et al.*, 1983; Rabbitts *et al.*, 1984). This autosuppression is proportional to the concentration of c-Myc protein and occurs at the level of transcript initiation (Penn *et al.*, 1990c). This negative

autoregulatory mechanism is active in non-tumorigenic cells but inactive in tumor cell lines, including the breast cancer cell line MCF 7 (Leder *et al.*, 1983; Grignani *et al.*, 1990). Conversely, autoregulatory activation of the c-*myc* gene has also been reported (Kitaura *et al.*, 1991). Activation was observed only with low levels of c-Myc in cells cultured without serum, while autosuppression was seen with cells grown in the presence of serum and with serum free growth conditions combined with increased amounts of c-Myc. In this study it was determined that a previously identified c-Myc binding site in the 5' flanking region of the c-*myc* gene (Iguchi-Ariga *et al.*, 1988; Ariga *et al.*, 1989) was necessary for both autoactivation and autosuppression (Kitaura *et al.*, 1991). Perhaps the self activation observed in growth limiting conditions reflects the role of c-Myc in apoptosis (Evan *et al.*, 1992).

Regulation of c-*myc* expression and c-Myc activity therefore occurs at a variety of levels. In addition to protein phosphorylation, the activity of c-Myc is controlled by the cellular levels of c-Myc and other Max associated proteins. The varying amounts of these proteins direct heterodimer formation with Max and determine if activation or repression of Myc responsive genes occurs. The quantity of c-Myc is partly regulated by the stability of c-*myc* transcripts and the rate of c-Myc synthesis. Similarly, the initiation and elongation of c-*myc* transcripts determines the availability of templates for translation. The tight control of all of these regulatory mechanisms appears necessary for maintaining the normal physiological functions of c-Myc. In particular, a complex arrangement of regulatory regions appears to direct the initiation of c-*myc* transcription.

Regulatory Elements of the c-*myc* Gene

Maximal expression from both the P_1 and P_2 c-*myc* promoters occurs under the influence of multiple regulatory regions found within the promoter proximal sequences and distal regions located within 2.3 kb upstream of the P_1 promoter (Hay *et al.*, 1987; Dubik and Shiu, 1992). A number of the regulatory factors, which associate with these regions, have been mapped within the human and mouse c-*myc* genes (reviewed in Spencer and Groudine, 1991; Marcu *et al.*, 1992). Some of these *trans*-activators interact at sites which are located in close proximity to hypersensitive sites associated with the transcriptional regulation of c-*myc*, while others have been localized to sequences which have been characterized as positive or negative elements by transient transfection assays (Figures 14 and 15). In addition to having a general role in the regulation of c-*myc* gene expression, some of these regions and their associated transcriptional regulators have been demonstrated to exert a promoter specific effect on transcription. Together these studies suggest independent transcriptional regulation of the c-*myc* promoters and the existence of multiple positive and negative elements which antagonize and synergize to yield appropriate c-*myc* expression.

Both positive and negative control elements have been described for the human c-*myc* gene (reviewed in Marcu *et al.*, 1992). Some of these reports are inconsistent, since almost identical c-*myc* fragments studied in different cellular backgrounds have been reported to exert both positive and negative effects. These discrepancies suggest there are cell type specific effects involved in determining the activity of each element. In essence, variations in the cellular environment reflect important differences in the availability of transcriptional regulators. Additionally, these observations support the

characterization of c-*myc* regulatory regions in the cell type of interest to ensure the appropriate cellular effect.

Large portions of 5' flanking sequence have been reported to have both positive; -2329 to -1257 and -1257 to -300 (Hay *et al.*, 1987), -2319 to -1980 (Lang *et al.*, 1988; Whitelaw *et al.*, 1988), and negative; -1527 to -1246 and -1052 to -607 (Lang *et al.*, 1988; Whitelaw *et al.*, 1988), -1052 to -511 (Lipp *et al.*, 1987) effects on the transcription initiation of the human c-*myc* gene (Figure 14). Additionally, fragments which encompass c-*myc* sequences -353 to +513 relative to P_1 have been reported to have a positive effect on transcription (Lipp *et al.*, 1987), but can be further fragmented into positive; -408 to -294 and -160 to -118 (Postel *et al.*, 1989), -353 to -101 and +66 to +513 (Lipp *et al.*, 1987), -293 to-101 and +112 to +121 (Hay *et al.*, 1987; DesJardins and Hay, 1993), -60 to -37 and +95 to +105 (Nishikura, 1986), +95 to +141 (Thalmeier *et al.*, 1989), -101 to +25 and +67 to +141 (Dubik and Shiu, 1992), -200 to +250 (Lang *et al.*, 1988; Whitelaw *et al.*, 1988) and negative; -353 to -293 and +112 to +121 (Hay *et al.*, 1987; DesJardins and Hay, 1993), -101 to +66 (Lipp *et al.*, 1987), -400 to -293 and +47 to +513 (Chung *et al.*, 1986), +141 to +202 (Dubik and Shiu, 1992) elements. In particular, the region -353 to -101 has been described as positive (Lipp *et al.*, 1987) but can be separated into negative, -353 to -293, and positive, -293 to -101, regions (Hay *et al.*, 1987). Similarly, although the region -1300 to -300 was shown to exert a positive effect (Hay *et al.*, 1987) a number of negative elements; -608 to -407, -400 to -293 (Chung *et al.*, 1986), -1052 to -607 (Lang *et al.*, 1988; Whitelaw *et al.*, 1988), -1052 to -511 (Lipp *et al.*, 1987) -1060 to -923 and -615 to -424 (Marcu *et al.*, 1988) and -353 to -293 (Hay *et al.*, 1987) have also been described.

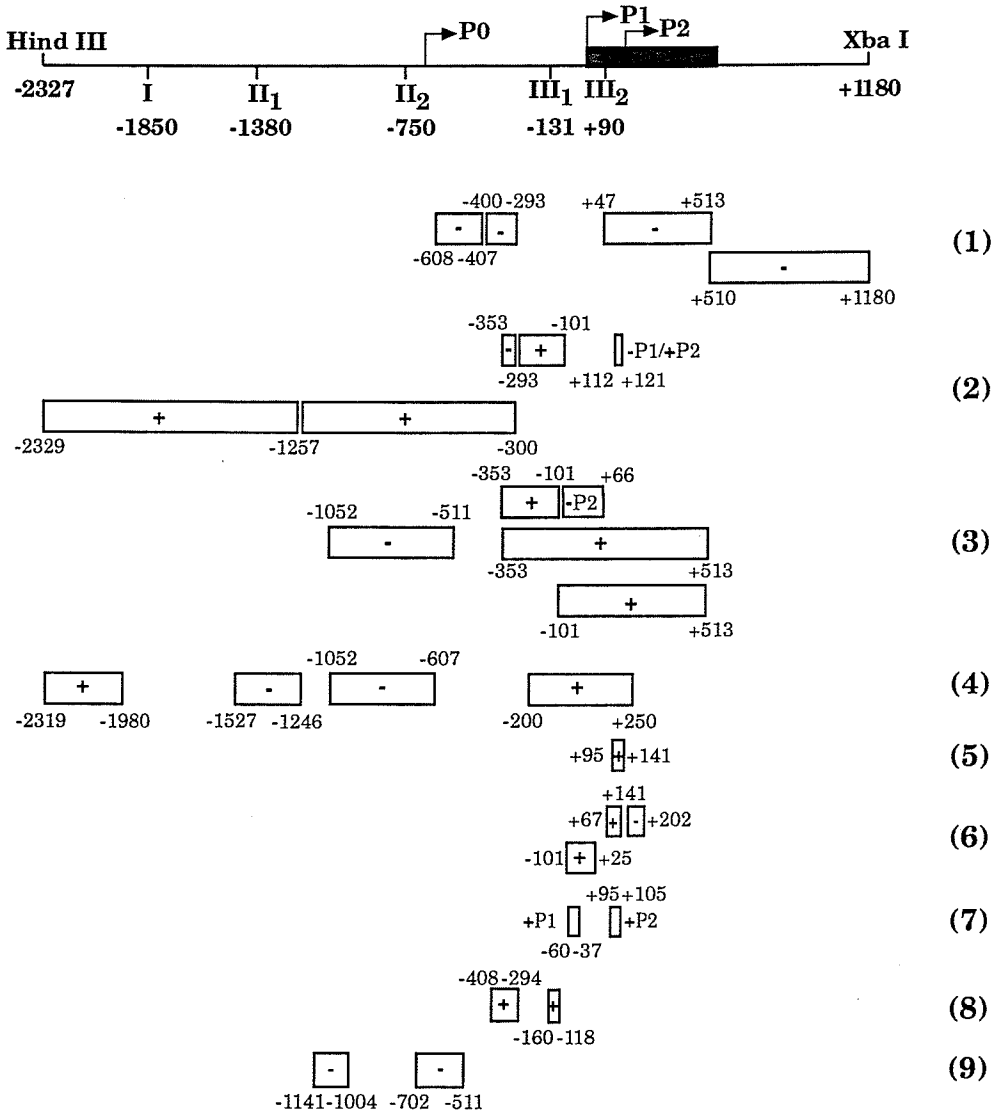


Figure 14. Positive and Negative Regulatory Regions of the c-*myc* Gene. A number of studies have identified positive and negative regulatory elements in the 5' untranslated regions of the c-*myc* gene. These regions and their associated references are diagrammed under the Hind III-Xba I portion of the c-*myc* gene containing the 5' flanking region, exon 1 and part of intron 1, the associated promoters (P0-P2) and DNase I hypersensitive sites (roman numerals). Numbers in brackets refer to (1) Chung *et al.*, 1986; (2) Hay *et al.*, 1987; DesJardins and Hay, 1993; (3) Lipp *et al.*, 1987; (4) Lang *et al.*, 1988; Whitelaw *et al.*, 1988; (5) Thalmeier *et al.*, 1989; (6) Dubik and Shiu, 1992; (7) Nishikura, 1986; (8) Postel *et al.*, 1989; (9) Remmers *et al.*, 1986; Weisinger *et al.*, 1988. With the exception of reference 9, all of these studies used the human c-*myc* gene. The sites characterized in reference 9 were identified with the mouse gene and the homologous human sequences are shown in this figure.

Some of these effects were also promoter specific, with sites -60 to -37 and -293 to -101 of the human c-*myc* gene specific for P_1 and sites +95 to +105 and -101 to +66 specific for P_2 (Nishikura, 1986; Lipp *et al.*, 1987; Hay *et al.*, 1987). Additionally, one site located at +112 to +121 relative to P_1 exerted a positive effect on P_2 and a negative effect on P_1 (DesJardins and Hay, 1993). Similarly, removal of the 3' terminal 60 bases of the mouse c-*myc* exon 1 disrupted transcript initiation from the P_2 promoter but not from P_1. This sequence was necessary, but not sufficient for P_2 activity (Yang *et al.*, 1986). Positive and negative sites have also been described for the mouse c-*myc* gene (Remmers *et al.*, 1986; Yang *et al.*, 1986; Kakkis *et al.*, 1989; Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992). In particular, similar promoter specific effects have been observed for P_1 and P_2 initiated transcripts (Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992). Upstream (Remmers *et al.*, 1986) and promoter proximal (Kakkis *et al.*, 1989) negative elements have also been characterized.

Distal Regulatory Elements and *Trans*-acting Factors

DNase I hypersensitivity of human c-*myc* chromatin near the DH I site (-1850) is often unaltered with differentiation (Siebenlist *et al.*, 1984; Dyson *et al.*, 1985) and has been suggested to contain a negative regulatory element (Siebenlist *et al.*, 1984). Conversely, one factor whose binding activity is altered upon differentiation has been identified as a positive *cis*-acting element (Avigan *et al.*, 1989). This factor, called FUSE (far upstream element), binds to the sequence, 5'-TCCCGAGGGA-3' within the region -1554 to -1526 relative to P_1 (Figure 15). The activity of this element disappeared when cells were induced to differentiate and was lost with the removal of the nucleotides CCGA from the inverted repeat.

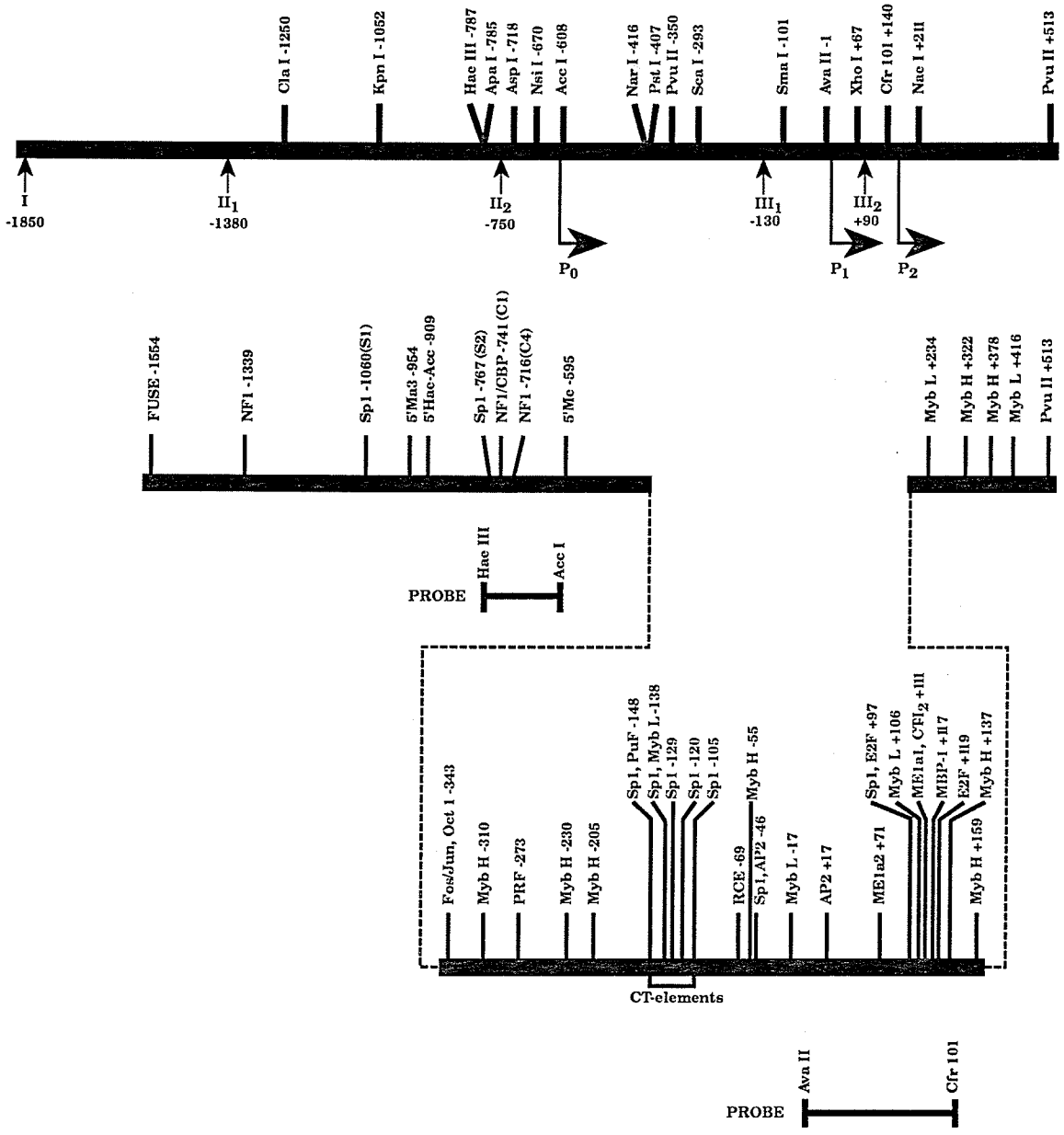


Figure 15. Transcriptional Regulators of the c-*myc* Gene. A number of *trans*-acting factors and their associated *cis*-acting elements have been identified for the c-*myc* gene and are discussed in the text. These sites and their approximate location are diagrammed relative to P1. The promoters (P0-P2), DNase I hypersensitive sites (Roman numerals) and various restriction sites are also labelled. The two regions of the c-*myc* gene studied using *in vitro* DNA binding assays, Hae III-Acc I and Ava II-Cfr101, are identified as probes below their location within the human c-*myc* gene. This figure was derived from references cited in the text.

Only two binding sites have been associated with the weak DH II_1 site found at -1380 of the human c-*myc* chromatin (Figure 15). They include an NF1 site, 5'-TGGAAGGCAGCCAA-3', located at -1339 relative to P_1 (Siebenlist *et al.*, 1984) and an Sp1 site, 5'-GCCCCTCCCA-3', found at -1060 relative to P_1 (Whitelaw *et al.*, 1988; Lang *et al.*, 1988; 1991). A binding site for Sp1 is also located near the DH II_2 site (-750) at -767 relative to P_1 (Whitelaw *et al.*, 1988; Lang *et al.*, 1988; 1990). Binding of factors to the sequence 5'-CAGGAGGGGCGG-3' is specifically inhibited with sequences which bind Sp1 at -1060 (Lang *et al.*, 1990). Similarly, an NF1 binding site, 5'-GTGGAAGG-3', is found at -716 relative to P_1 which is 50 bases 5' of the first P_0 start site (Bentley and Groudine, 1986b). This site has been further characterized and found to contain two NF1 sites within the sequence 5'-GGTGGAAGGTATCCAAT-3' (Whitelaw *et al.*, 1988; Lang *et al.*, 1988; 1990). Binding to this element can be competitively removed by an NF1 consensus sequence but not by other CCAAT-binding sequences (Lang *et al.*, 1990). These authors identified an additional NF1 site which is located in close proximity to a binding site for CCAAT-binding protein (CBP) further upstream but still within the DH II_2 region. This dual element is located at -741 relative to P_1 and contains the sequence 5'-TTTGGCAGCAAATTGG-3' (Figure 15). Binding of both factors occurs *in vitro*, although it was suggested that binding of one factor would preclude the interaction of the other. This binding site correlates with one published c-*myc* sequence (Gazin *et al.*, 1984), however binding of CBP is disrupted when the sequence 5'-TTTGGCAAGAATTGG-3', which is consistent with a second published c-*myc* sequence (Siebenlist *et al.*, 1984) is used. Both of these c-*myc* sequences were apparently derived from normal human DNA.

A negative response element has been described for the mouse c-*myc* gene between -1140 to -424 relative to P_1 which corresponds to the sequences -1221 to -511 relative to P_1 in the human gene (Remmers *et al.*, 1986; Weisinger *et al.*, 1988). The mouse c-*myc* negative element was further subdivided into two negative regions, -1060 to -923 and -615 to -424, (Weisinger *et al.*, 1988) which overlap with the negative elements, -1052 to -511 and -607 to -407 (Figure 14), described for the human gene (Chung *et al.*, 1986; Lipp *et al.*, 1987). Although six factor binding sites were described, only elements 5'Ma3 (8/10), 5'Hae-Acc (9/11) and 5'Me (11/15) were found to share sequence similarity with the human gene. Additionally the distance of these elements from the promoter was well conserved with sites 5'Ma3, 5'Hae-Acc and 5'Me respectively located at -954 to -945, -909 to -898 and -595 to -581 relative to human c-*myc* P_1(Figure 15). Sequence similarity with a known factor binding site was apparent for the 5'Ma1 site at -1128 to -1109 of the mouse gene. This site has the internal sequence 5'-GAGGAGGGG-3' which shares sequence similarity with the Sp1 consensus binding site (Kadonaga *et al.*, 1986). The remaining elements, 5'Ma2, 5'Ma3, 5'Hae-Acc, 5'Md, and 5'Me, shared sequence similarity with the SV40 enhancer, while 5'Md and 5'Me also shared sequence similarity with the SV40 promoter and the polyoma virus enhancer, respectively. The significance of these similarities is unclear.

An NF-κB like binding site, which has also been implicated in the negative control of the mouse c-*myc* gene, was identified at -1101 to -1081 relative to P_1 (Duyao *et al.*, 1990). This site would be located between the mouse c-*myc* 5'Ma1 and 5'Ma2 negative regulatory elements located at -1128 to -1109 and -1087 to -1084, respectively (Weisinger *et al.*, 1988). A significant sequence similarity, 9/10, exists between this NF-κB like site,

5'-GGGTTTTCCC-3', and a protein binding site located at -776 to -764 within a steroid dependent regulatory element of the chicken ovalbumin gene (Schweers and Sanders, 1991). This positive element is required for induction of the ovalbumin gene by estrogens, glucocorticoids, progestins and androgens. This *cis*-acting element does not contain consensus HREs nor does it appear to bind receptors for any of these steroid hormones. Interestingly, it is only active when linked to a negative element. An additional NF-κB like binding site, identified within exon 1 at +440 to +459, relative to P_1, mediated interleukin-1 induction of mouse c-*myc* expression (Kessler *et al.*, 1992).

Promoter Proximal Elements and *Trans*-acting Factors

A repressor element associated with the DH III_1 site (-131) of the human c-*myc* gene was originally located between -353 and -101 relative to P_1 (Hay *et al.*, 1987). This negative region was later subdivided into positive and negative elements (Hay *et al.*, 1987; DesJardins and Hay, 1993). An element localized to a 60 bp region at -353 to -293 relative to P_1 continued to have an inhibitory effect on both the P_1 and P_2 promoters (Hay *et al.*, 1987). The sequence, 5'-GCCTGCGATGATTTATACTCACAGGA-3', found at -343 to -318 (Figure 15) binds both Oct1 and AP1 proteins (Hay *et al.*, 1989; Takimoto *et al.*, 1989). The presence of Fos within the AP1 complex was confirmed by the ability of antibodies directed towards distinct regions of the Fos protein to specifically block complex formation and the immunoprecipitation of Fos protein from c-*myc* DNA-protein complexes (Hay *et al.*, 1989; Takimoto *et al.*, 1989). Although a number of Fos related antigens were detected in these complexes (Takimoto *et al.*, 1989), the interaction of *in vitro* translated Fos with this element was found to be dependent on the presence of Jun (Hay *et al.*, 1989). The additional

interaction of Octamer proteins with this element was confirmed by the ability of oligonucleotides containing an octamer motif to compete for binding and the association of purified Octamer protein with this negative element (Takimoto *et al.*, 1989). Significantly, although the degenerative binding sites for the Octamer and AP1 proteins overlap, the interaction of these factors to this c-*myc* element appears to be independent. Although it has not been determined which of these factors exerts the negative effect on c-*myc* transcription, it has been suggested that these factors may act at different points in the cell cycle (Takimoto *et al.*, 1989).

An element similar to the binding site for the plasmacytoma repressor protein, 5'-GAGAAAGGGAGAGGGTTT-3', falls within the human c-*myc* gene at -273 to -257 (Figure 15) relative to P_1 (Kakkis and Calame, 1987). This factor was characterized from mouse plasmacytomas which repress the nonrearranged mouse c-*myc* allele and was not present in human B cell lines representing early stages of differentiation when c-*myc* was still transcribed (Kakkis and Calame, 1987; Kakkis *et al.*, 1989). This factor is believed to displace or modulate the binding characteristics of a positive acting factor, CF1 (common factor 1), which interacts 3' of the plasmacytoma repressor protein site in the mouse c-*myc* gene (Kakkis *et al.*, 1989; Riggs *et al.*, 1991).

The region -293 to -101 relative to P_1 is associated with the DH III_1 site at -130 relative to P_1 in the human c-*myc* gene (Figure 15). Sequences located in this region have been described as necessary for transcript initiation from P_1 and stimulatory for P_2 activity (Hay *et al.*, 1987). Within this region is a sequence located between -157 and -97 relative to P_1 which contains 5 repeated CT-elements spaced every 10 bases (Hay *et al.*, 1987; DesJardins and Hay, 1993). This G+C rich sequence, 5'-TCCTCCCCACCTTCCCCACCCTCCCCACCCTCCCCATAAGCGCCCCTC

CCG-3', was characterized as a positive element for P_1. The integrity of all five of the repeated elements was required for basal initiation from P_1 and for maximal activity of P_2. The repeated elements in this sequence have significant similarity to Sp1 consensus sites on the opposing strand (Kadonaga *et al.*, 1986). Also found associated with sequences of the DH III_1 region was an Sp1 site located at -46 to -32 relative to P_1 (Figure 15), which was sufficient for P_1 initiated transcription from c-*myc* microinjected into *Xenopus* oocytes (Nishikura, 1986). This Sp1 site overlaps with one of two AP2 sites located at -43 and +17 relative to P_1 (Imagawa *et al.*, 1987). Additionally, a Rb control element (RCE), which binds Sp1, has been described for c-*myc* at -69 to -96 relative to P_1 (Kim *et al.*, 1992; Udvadia *et al.*, 1993).

A single copy of a related CT-element is associated with the human c-*myc* DH III_2 site, which is located at +90 relative to P_1. This element occurs in an inverted orientation 53 bp upstream of P_2 at +112 to +121 relative to P_1 (DesJardins and Hay, 1993). Designated as CT-I_2 (Figure 15), this protein binding site has the sequence 5'-GAGGGAGGG-3' and confers an inhibitory effect on transcription from P_1 while being required for initiation from P_2. Conversely, a region, +96 to +107 relative to P_1, encompassing an Sp1 site (Figure 15) was essential for P_2 transcription from human c-*myc* sequences microinjected into *Xenopus* oocytes (Nishikura, 1986). Removal of CT-I_2 element relieved the dependence of P_1 transcription on the integrity of the repeated elements while simultaneously decreasing P_2 transcription (DesJardins and Hay, 1993). Deletion of the five tandem repeats and the CT-I_2 element combined with the presence of a heterologous enhancer caused a shift in promoter usage analogous to that observed with the c-*myc* translocations found in BL cells.

Significant sequence similarity is observed between the mouse and human c-*myc* genes in this region and homologous G+C rich elements to human c-*myc* have been described within the 400 bp preceding P_1 and between P_1 and P_2 of the mouse c-*myc* gene (Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992). Four elements located upstream of mouse c-*myc* P_1 at -264 to -228, -188 to -156, -129 to -103 and -55 to -25, were respectively called 5'Mf1, 5'Mg1, 5'Mg2, 5'Mg3 (Asselin *et al.*, 1989). Contrary to the findings with human c-*myc*, the integrity of all four G+C rich elements 5' of P_1 was not required for P_1 transcription since the proximal 109 bp containing only the 5'Mg3 element and the TATA box was found to be sufficient. This was similar to what was observed when human c-*myc* sequences were microinjected into *Xenopus* oocytes (Nishikura, 1986).

The sites ME1a2 and ME1a1, located between the P_1 and P_2 promoters of the mouse c-*myc* gene at +58 to +78 and +97 to +118 relative to P_1, respectively, were also characterized (Lipp *et al.*, 1987; Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992). These sites were homologous to Sp1-like sequences present at +71 and +111 relative to P_1 of the human c-*myc* gene (Figure 15). Both of these elements were required for P_2 initiated transcription from the mouse c-*myc* gene, although optimal transcription relied heavily on the ME1a2 element (Moberg *et al.* 1991). Significantly, the ME1a1 element was analogous to human c-*myc* CT-I_2 in that it also had an inhibitory effect on P_1 transcription, while simultaneously enhancing initiation from P_2 (Asselin *et al.*, 1989). In addition to functioning as a regulator of transcription initiation, the ME1a1 element has also been described to regulate the attenuation of transcript elongation from both the P_1 and P_2 promoters (Miller *et al.*, 1989b; Wright *et al.*, 1991; Bossone *et al.*, 1992; Dufort *et al.*, 1993).

The interaction of Sp1 with the c-*myc* G+C rich elements has been demonstrated for both human and mouse genes (Asselin *et al.*, 1989; DesJardins and Hay, 1993). Using purified Sp1 protein, the mouse 5'Mg1 and 5'Mg3 elements were characterized as high affinity Sp1 sites, while 5'Mg2, ME1a2 and ME1a1 were determined to be low affinity sites (Asselin *et al.*, 1989). The interaction of additional factors to the low affinity sites was evident with nuclear protein extracts which also generated a DNase I hypersensitive site at the 5' border of the ME1a1 site which was not apparent in the presence of purified Sp1. Purified Sp1 protein was also able to form specific protein-DNA complexes with the five tandem CT elements and the single CT-I_2 element of the human c-*myc* gene (DesJardins and Hay, 1993). Confirmation of the interaction of Sp1 with these elements was acquired with the observations that the molecular weight and zinc requirement of the DNA binding activity resembled Sp1. Additionally, the presence of Sp1 antibodies and Sp1 consensus oligonucleotides disrupted protein interactions with these elements.

The antibiotic, mithramycin, appears to block transcriptional initiation of both the P_1 and P_2 promoters through the disruption of Sp1 binding (Snyder *et al.*, 1991). Partially purified Sp1 and mithramycin were specifically found to interact with an overlapping site, -46 to -32, in the human c-*myc* gene (Nishikura, 1986; Snyder *et al.*, 1991). Additional protein interactions within the regions -101 to +121 relative to P_1, which were not specified to be due to Sp1 binding, were also sensitive to the presence of mithramycin. Significantly, the region +67 to +121 contains the binding sites for ME1a2, Sp1 and ME1a1 or CT-I_2 (Nishikura, 1986; Asselin *et al.*, 1989; DesJardins and Hay, 1993), in addition to two E2F sites which are described below (Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Thalmeier *et al.*, 1989).

The recently cloned zinc finger protein ZF87 or MAZ (Pyrc *et al.*, 1992; Bossone *et al.*, 1992), which was able to bind the ME1a2 and ME1a1 sites of the mouse c-*myc* gene, was also able to interact with the homologous human c-*myc* elements. Although the interaction of these proteins with the human c-*myc* G+C rich elements generated a methylation interference pattern which differed from that characterized for nuclear extracts or purified Sp1 interactions (DesJardins and Hay, 1993), the data suggest that other proteins may also recognize these regulatory sequences. Significantly, a partially purified factor called PuF directly interacted with an inverted G+C rich sequence, 5'-GGGTGGG-3', which overlaps two of the G+C repeat elements found at -142 to -126 relative to P_1 (Postel *et al.*, 1989). This factor was recently identified to be homologous to nm23, a nucleoside diphosphate kinase involved in preventing metastasis (Postel *et al.*, 1993). Four of the repeat sites, -151 to -117, were also recognized by a ribonuclear protein (Davis *et al.*, 1989). This region was also capable of forming a triplex H-DNA structure which was capable of repressing c-*myc* transcription (Cooney *et al.*, 1988; Davis *et al.*, 1989; Postel *et al.*, 1989). Additionally, a factor, CTCF, has been described to interact with similar G+C repeat sequences located within the chicken c-*myc* gene upstream of the start of transcription (Lobanenkov *et al.*, 1986; 1990).

The DH III_2 region of human c-*myc* also contains two sequence elements, 5'-GCGGGAAA-3' and 5'-GATCGCGC-3', located at +98 to +105 and +121 to +129 relative to P_1, respectively (Figure 15). Both of these sites are recognized by the transcription factor E2F (Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Thalmeier *et al.*, 1989). The first site, located between the G+C rich elements ME1a2 and ME1a1, is conserved between mouse and human and is also bound by E2F in the mouse c-*myc* gene (Moberg *et al.*, 1991; 1992).

This site, which overlaps an Sp1 site identified for P_2 (Nishikura, 1986), was found to be necessary for P_2 initiation from both human (Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Thalmeir *et al.*, 1989) and mouse (Moberg *et al.*, 1991; 1992) c-*myc* genes. The presence of at least one E2F site was also important for the *trans*-activation of c-*myc* by the adenovirus E1a proteins (Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Thalmeir *et al.*, 1989). The second E2F site, which is located just 3' of the CT-I_2 or ME1a1 element, does not appear to be conserved between human and mouse c-*myc* (Hiebert *et al.*, 1989; Moberg *et al.*, 1991). Although binding to this site was weaker than to the conserved E2F site, its presence contributed to the overall activity of the P_2 promoter.

Adjacent to the 3' end of the nonconserved E2F site is a binding site for *myc* binding protein 1, MBP-1 (Figure 15). Interaction of this protein with the human c-*myc* gene occurred immediately 5' of the P_2 TATA box and repressed transcript initiation (Ray and Miller, 1991). It has therefore been suggested that the interaction of MBP-1 with its binding site may prevent the interaction of the TFIID proteins with the promoter. Additionally, a preliminary report suggests that protein interactions with the adjacent E2F and CT-I_2 elements of the human c-*myc* gene may be mutually exclusive (DesJardins and Hay, 1993). Significantly, the *Xenopus* c-*myc* I gene contains two tandem Sp1 sites plus one overlapping Sp1 and E2F site, all of which were required for transcriptional activity in oocytes (Modak *et al.*, 1993). The overlapping site mediated Sp1 but not E2F binding in these cells. Conversely, the binding of proteins to the ME1a2, E2F and ME1a1 elements in the chromatin of the mouse c-*myc* gene has been demonstrated by *in vivo* genomic footprinting (Plet *et al.*, 1992). Although protein interactions at the ME1a2 and E2F sites were lost in cells where the gene was silent or

truncated, these interactions were still apparent in cells induced to differentiate.

A vast number of Myb binding sites have been described for the human (Nakagoshi *et al.*, 1992; Zobel *et al.*, 1991; 1992) and mouse (Cogswell *et al.*, 1993) c-*myc* gene. Interestingly, the *trans*-activation ability of the Myb proteins has been characterized as both dependent and independent of the ability to interact with DNA (Klempnauer *et al.*, 1989). Initially, 6 Myb sites that were identified 1061 bp upstream of human c-*myc* P_1, were found to be unnecessary for *trans*-activation by c-Myb (Zobel *et al.*, 1992). Similarly, the deletion of 3 Myb binding sites downstream of P_2 (+615) and mutations in a fourth site next to the 3' side of the P_2 TATA box, had no significant effect on *trans*-activation. However, these authors found that the c-Myb DNA binding domain was required for this activity. Significantly, 15 additional Myb binding sites were later identified between -349 and +513 relative to P_1 (Nakagoshi *et al.*, 1992). Ten of these sites, +378 to +413 (2 sites), +322 to +348, +159 to +178, +137 to +152, -55 to -29, -205 to -191, -230 to -213 and -310 to -290 (2 sites), were determined to be high affinity Myb binding sites, while 5 sites, +416 to +439, +234 to +253, +122 to +106, -17 to -3 and -138 to -116, were characterized as low affinity sites (Figure 15). These authors found that the presence of Myb binding sites at specific positions was not necessary but that some number of sites was required for the Myb induced *trans*-activation function. Although not specifically addressed, the question of whether or not the Myb DNA binding domain was required for *trans*-activation was implied by the requirement for binding sites (Nakagoshi *et al.*, 1992).

Ten Myb binding sites upstream of P_1 (-238 to -1010) and six sites downstream of P_2 (+216 to +558 relative to P_1) were identified in the mouse

c-*myc* gene (Cogswell *et al.*, 1993). The integrity of sequences encompassing the 5 sites located between -868 and -424 and the 5 sites located between +148 and +515 was required for T-cell lines, but only the sequences 3' of P_2 were required in lymphoma and myeloid cell types. For both situations the DNA binding and *trans*-activating domains of c-Myb were required for c-*myc* activation. Although the sites characterized in the mouse c-*myc* gene appeared functionally equivalent to the human gene, only Myb site 15 identified in the mouse was identical to a site characterized in the human gene (Myb H +378), with both mapping approximately 400 bp 3' of P_1 (Cogswell *et al.*, 1993; Nakagoshi *et al.*, 1992). Some of the other Myb binding sites in the human c-*myc* gene overlap sequences previously identified to bind other *trans*-acting proteins. In particular, a Myb binding site located just 3' of the P_2 TATA box (Zobel *et al.*, 1992) may interfere with the MBP-1 site (Ray and Miller, 1991), while Myb sites +106 to +122 and -139 to -116, (Nakagoshi *et al.*, 1992) overlap the CT-I_2 or ME1a1 element and the CT or Sp1 repeats, respectively (DesJardins and Hay, 1993). In the mouse c-*myc* gene, one Myb binding site overlaps with the 5'Mf1 site (Asselin *et al.*, 1989; Cogswell *et al.*, 1993). Myb sites homologous to the promoter proximal sites identified in the human gene, -101 to +122 relative to P_1, have not yet been identified in the mouse.

There is a large body of literature which has focused on the interaction of *trans*-acting factors with regulatory regions of the c-*myc* gene. Most of these studies have concentrated on the regulatory regions proximal to the P_1 and P_2 promoters. Importantly, this work has revealed the complex network of regulatory proteins associated with the regulation of transcription from these promoters. In particular, multiple binding sites for Sp1-like proteins in this region are necessary for determining the appropriate promoter usage.

Conversely, very little research has dealt with distal regulatory regions. In particular, little attention has been given to the proteins associated with the DH II_2 region, even though strong hypersensitivity at this site in proliferating cells is lost upon cellular differentiation. The lack of interest in this region of the c-*myc* gene is in part due to the low percentage of transcription initiated from the associated P_0 promoter. However, the association of Sp1-like proteins with this region of the c-*myc* gene suggests that other factors interacting at DH site II_2 may be important in regulating the transcriptional activity of the P_1 and P_2 promoters.

Estrogen Regulation of c-*myc* Expression

The c-*myc* proto-oncogene has been shown to be stimulated by estradiol in the rat uterus (Murphy *et al*, 1987). Ovariectomy produced a substantial loss in uterine expression of c-*myc*, which was maximally increased, 8 fold, 3 hours after the administration of 17ß-estradiol. Similarly c-*myc* expression is responsive to estrogen in ER positive human breast cancer cell lines (Dubik *et al.*, 1987; Wilding *et al.*, 1988; van der Burg *et al.*, 1989). Growth of the ER positive breast cancer cell lines, MCF 7 and T47 D, in estrogen depleted conditions resulted in simultaneous decreases in c-*myc* expression and cell proliferation, both of which declined further with the addition of the anti-estrogen, tamoxifen (Dubik *et al.*, 1987; Wong and Murphy, 1991). Although tamoxifen has some agonist activity and may not have reduced c-*myc* expression to absolute basal levels, both growth inhibition and c-*myc* suppression were reversed by the addition of 100 nM 17ß-estradiol. Furthermore, surgical removal of estrogen sources from athymic nude mice growing MCF 7 breast carcinoma xenografts has been demonstrated to reduce tumor size and stimulate c-*myc* induced cell death (Kyprianou et. al., 1991). Conversely, the expression of c-*myc* in the ER

negative cell lines, MDA MB 231, HBL 100 and BT 20, was unaffected by the presence or absence of estrogens (Dubik *et al.*, 1987).

Stimulation of c-*myc* mRNA levels in MCF 7 human breast cancer cells to a level equivalent to that observed under control growth conditions occurred by 15 minutes, and was maximal, 10 fold, 60 minutes after the addition of estradiol (Dubik *et al.*, 1987). The level of c-*myc* mRNA was seen to decline over time but remained above the level observed in tamoxifen treated cells. The time frame for estrogen induced c-*myc* expression in T47 D cells was similar, peaking 2 hours after treatment. Both of these expression patterns were similar to the c-*myc* expression observed with *in vivo* estrogen induction in the rat uterus (Murphy *et al.*, 1987). In MCF 7 cells, c-*myc* accumulation in response to estrogen was enhanced in the presence of the protein synthesis inhibitor, cycloheximide, while treatment with actinomycin D, an inhibitor of transcription, suppressed estrogen stimulation (Dubik and Shiu, 1988; van der Burg *et al.*, 1989). In the absence of protein synthesis inhibitors a corresponding increase in c-Myc protein accumulation was observed in response to estrogen treatment of MCF 7 cells, and was maximal by 90 minutes (Watson *et al.*, 1991).

The effect of estrogen on c-*myc* in MCF 7 cells was determined to be due to an increase in transcription which was detectable using nuclear "run-on" assays within 5 minutes and maximal, 10 fold, by 20 minutes (Dubik and Shiu, 1988). Transcription remained above basal levels beyond 4 hours, but was significantly decreased from the maximal activity observed. Although another group reported no estrogen induced transcriptional induction of c-*myc* in this cell line (Santos *et al.*, 1988), their first time point was 60 minutes and therefore they would have missed the peak of transcriptional activity. These authors suggested post-transcriptional

modulation of c-*myc* by estrogen in MCF 7 cells, however, no effect of estrogen on the c-*myc* mRNA half-life, 18 minutes, was reported, and others have not observed such an effect (Dubik and Shiu, 1988). The transcription effect of estrogen on c-*myc* in ER^+ breast cancer cells has been confirmed recently (Davidson *et al.*, 1993), conversely, a more stable c-*myc* transcript with a half-life of 49 minutes has been detected in the ER^- cell line, MDA MB 231 (Dubik and Shiu, 1988).

Transcriptional activation by steroid hormones is known to occur through receptor complexes which interact with hormone responsive elements in the flanking regions of responsive genes (reviewed in Evans, 1988; Beato, 1989). With the characterization of estrogen regulated c-*myc* transcription one may expect to find DNA sequences related to the consensus ERE, 5'-AGGTCAnnnTGACCT-3', in regulatory regions of this gene. Significantly, an ERE-like half site 5'-AGGGCA-3' (Figure 16 B) has been identified between the P_1 and P_2 promoters of c-*myc* (Dubik and Shiu, 1992). These authors monitored estrogen induced chloramphenicol acetyl transferase (CAT) activity of c-*myc-CAT* chimeric genes transiently transfected into ER^+ MCF 7 cells and HeLa cells co-transfected with an ER expression vector. The c-*myc-CAT* constructs incorporated c-*myc* 5' flanking and exon 1 sequences from -2327 to +25 relative to the P_1 promoter (Figure 16 A). The minimal region required for estrogen induced activity, 6.2 fold, was found to be a 116 nucleotide region upstream of the P_2 promoter, between +25 and +141 relative to P_1. Significantly, the addition of 5' flanking sequences, -667 to +141, to this estrogen responsive region resulted in greater activity, 9.2 fold, which was reduced to 5.1 fold with the further addition of P_2 downstream sequences, +141 to +202. Additionally, removal of sequences near the P_1 promoter, -101 to +25, from the construct

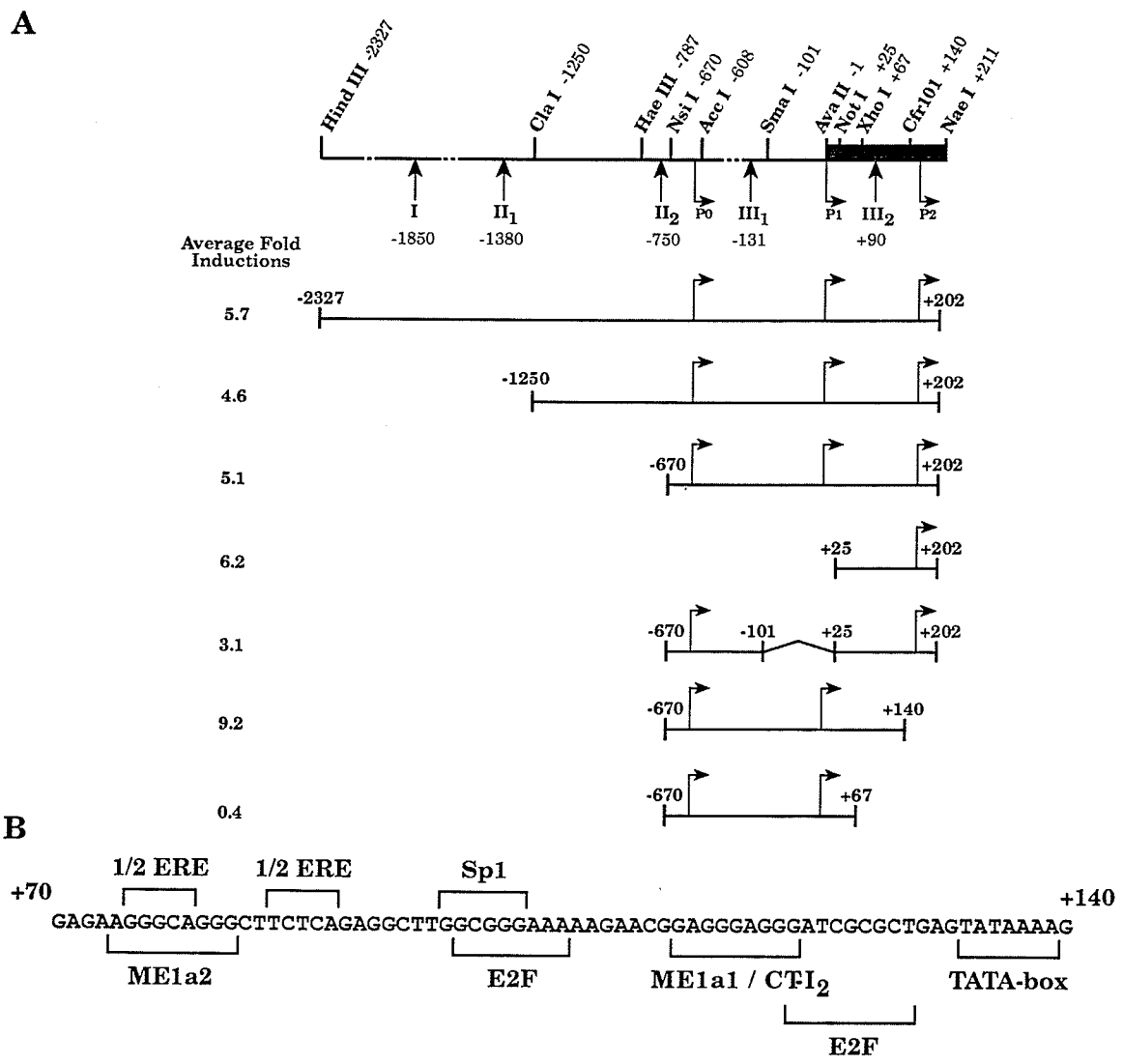


Figure 16. Estrogen Regulation of c-*myc* and the Identification of a Half-Site ERE. A. The regions of the c-*myc* gene used to identify an estrogen responsive element are diagrammed below a schematic of the c-*myc* 5' flanking sequences and the promoter region of exon 1. The fold induction observed using transient transfection and CAT assays is indicated on the left. These constructs were cotransfected with an ER expression vector into HeLa cells. This portion of the figure was reproduced from Dubik, 1991; Dubik and Shiu, 1992. **B.** The sequence of the human c-*myc* gene surrounding the putative estrogen response element (ERE). Binding sites for additional proteins identified to interact with this region of c-*myc* are labelled. The half site ERE identified by Dubik and Shiu, 1992 is based on the consensus sequence characterized in Figure 3C and is identical to the lower strand of the second half site of the uteroglobin ERE. This site overlaps the binding site for the Sp1-like factor, ME1a2. A second half site ERE based on the same consensus sequence of Figure 3C is also shown. Together these two half sites form the sequence 5'-GGGCA-n_5-TCTCA-3', which conforms to the lower strand of the consensus sequence shown in Figure 3C.

encompassing c-*myc* sequences -667 to +202 resulted in reduced estrogen responsiveness, 3.1 fold, while deletion of sequences near the P_2 promoter, +67 to +141, resulted in a complete loss of estrogen induction.

Although the authors identified only a half site ERE-like element in the +25 to +141 region (Figure 16 B), co-expression of ER mutants lacking the DNA binding domain were unable to elicit a response, suggesting the integrity of this domain was required for transcriptional activation (Dubik, 1991; Dubik and Shiu, 1992). The modulating capacity of the addition or deletion of various regulatory regions to the minimal region required for an estrogen response was suggestive of a role for other regulatory factors in the estrogen induced transcription of c-*myc* in these cells. Significantly, the half-site ERE identified in these studies overlaps a human c-*myc* element which is homologous to the ME1a2 site described for the mouse gene (Lipp *et al.*, 1987; Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992). A possible second half site ERE is present in this region (Figure 16 B) and is spaced 5 bp 3' of the half site identified by the above studies (Dubik and Shiu, 1992).

Although previous reports support monomeric binding of the ER to half site elements, the ability of the ER to directly interact with these putative c-*myc* half site EREs has not been determined. Additionally, the role of the surrounding c-*myc* regulatory regions in supporting the estrogen response is also unclear. For the maximal estrogen response the sequences from -670 to +140 relative to P_1, are required. In addition to retaining the integrity of the CT elements or Sp1 sites required for maintaining the proper ratio of P_1 and P_2, this region contains DNA sequences associated with the DH II_2, DH III_1 and DH III_2 sites. Therefore, elements associated with DH site II_2 may be necessary for the human c-*myc* gene to maximally respond to estrogen.

Hypothesis

The transcriptional regulation of c-*myc* expression in ER^+ hormonally dependent breast cancer cells involves a different complement and/or activity of nuclear *trans*-activating factors than ER^- hormonally independent breast cancer cells.

Significance and Rationale

The World Health Organization and the National Cancer Institutes of Canada and the United States have established that breast cancer is a significant health problem for women and a leading cause of death. Although the methods for detecting breast cancer have improved, there has been no significant change in breast cancer mortality. Advances have been made in the development of markers for predicting patient prognosis, however the treatment of breast cancer has largely remained unchanged over the last 20 years. Some of these treatments have focused on blocking estrogen action, however treatment failure often results in breast cancer recurrence, which threatens patient survival. Significantly, it is known that the ovarian influences dictated by the age of a woman at menarche, first pregnancy and menopause are important factors in establishing breast cancer risk. Research focused towards an understanding of the hormonal regulation of breast cancer is therefore important to an understanding of the general biology of this disease. In particular, an understanding of the molecular changes that select for the hormonally independent phenotype may lead to mechanisms to prevent breast cancer progression to this more aggressive stage.

Overexpression of the c-*myc* gene due to gene amplification has been associated with a poor patient prognosis which results in an earlier relapse of

the disease and an earlier death. These associations suggest the c-*myc* gene could be a significant marker for patient survival. Cellular proliferation, differentiation and apoptosis are all processes regulated by c-Myc, therefore alterations in the regulation of c-*myc* expression would be expected to have a substantial impact on these pathways. Significantly, the overexpression of c-*myc* has been shown experimentally to disrupt the relative concentrations of c-Myc in the cell, resulting in an increase in the formation of heterodimers with Max and increased gene activation by c-Myc. The estrogen regulation of c-*myc* in breast cancer cell lines implies that expression of this gene would be effected by hormonal therapies aimed at blocking estrogen action. This suggests that molecular changes, which result in a loss of estrogen regulation and subsequent constitutive expression of c-*myc*, may be involved in the progression of breast tumors from estrogen dependence to estrogen independence.

A significant body of research suggests that changes in the chromatin structure of a gene correspond with protein interactions at critical regulatory regions. In particular, chromatin alterations are generally the result of changes in nucleosome position and/or structure due to the interaction of *trans*-acting factors and other effector proteins at these sites. The chromatin organization of many promoter specific and distal regulatory regions of the human c-*myc* gene have previously been mapped and some of the transcriptional regulators associated with these sites have been characterized. An analysis of c-*myc* chromatin structure in ER^+ and ER^- breast cancer cell lines using DNase I hypersensitivity was therefore undertaken to initiate a study of molecular differences associated with estrogen regulated and constitutive c-*myc* expression in these cells. Differences in c-*myc* chromatin structure detected by this method were

further analyzed by *in vitro* DNA binding studies aimed at characterizing the specific protein interactions involved. Protein interactions in the vicinity of the putative c-*myc* ERE were also analyzed using these *in vitro* techniques. Additionally, an analysis of RNA levels, gene amplification, promoter usage and methylation status of the c-*myc* gene in these cells was performed to aid in the interpretation of the other findings.

Methods and Materials

Chemicals

Deoxyribonuclease I (DNase I) type II was obtained from Sigma (St. Louis, Missouri). All other DNase I was purchased from Gibco BRL Life Technologies (Burlington, Ontario), Promega Biocan (Mississauga, Ontario) or Stratagene (La Jolla, California). The Bgl II (50 unit/μl) restriction endonuclease was purchased from Gibco BRL Life Technologies. All other restriction endonucleases were purchased from Pharmacia LKB Biotechnology (Baie d'Urfe, Quebec), Boehringer Mannheim (Laval, Quebec), Gibco BRL Life Technologies, Stratagene or New England Biolabs (Mississauga, Ontario). The non-specific carrier DNA poly dI-dC was from Sigma, Pharmacia LKB Biotechnology or Boehringer Mannheim. Radiolabelled nucleotides were from ICN Biomedicals (Mississauga, Ontario) and DuPont New England Nuclear (Lachire, Quebec). Nitrocellulose used for RNA and DNA transfers was nitroplus 2000 purchased from Fisher Scientific (Nepean, Ontario). Fetal calf serum and Dulbecco's minimal essential medium (DMEM) powder were purchased from Gibco BRL Life Technologies. Unless otherwise stated, all of the other cell culture ingredients were purchased from Flow Labs (Mississauga, Ontario). Chemicals required for electrophoresis included agarose (Gibco BRL Life Technologies), low melting point agarose (FMC Bioproducts; Rockland, Maine), acrylamide (Boehringer Mannheim), bis-acrylamide (Boehringer Mannheim, ICN Biomedicals), urea (ICN Biomedicals), ammonium persulfate (Bio-Rad Laboratories; Richmond, California) and Temed (ICN Biomedicals). Unless otherwise stated, all of the other chemicals were obtained from Sigma, while supplies were from Fisher Scientific or Baxter Diagnostics, CanLab Division (Mississauga, Ontario).

Cell Lines

The T47 D5 cell line was kindly provided by Dr. R. L. Sutherland (Garvan Institute for Medical Research, Sydney, Australia) in whose laboratory it was originally isolated (Reddel *et al.*, 1988). The S30 cell line (Jiang and Jordan, 1992) was kindly provided by Dr. V. C. Jordan (University of Wisconsin Comprehensive Cancer Center, Madison, Wisconsin). The MCF 7/AdrR cell line and the parental cells, MCF 7 WT, from which they were originally isolated (Vickers *et al.*, 1988) were kindly provided by Dr. K. H. Cowan (National Cancer Institute, National Institutes of Health, Bethesda, Maryland). All other cell lines were obtained from sources described previously (Dubik *et al.*, 1987, see also publications for original descriptions as referred to in the Introduction).

Most of the cell lines were maintained at 37°C with 5% CO_2 in DMEM adjusted to pH 7.6 before filter sterilization and supplemented with 5% fetal bovine serum (FBS), 2 mM L-glutamine, 0.3% glucose, 100 IU/ml penicillin G sodium and 100 μg/ml streptomycin sulfate (penicillin/streptomycin; Gibco BRL Life Technologies). The MCF 7/AdrR cell line was maintained in the above medium in the presence or absence of 7 μM adriamycin (Sigma). The S30 cell line was maintained in phenol red free (PRF) DMEM supplemented with 10% twice dextran-charcoal stripped FBS (CS-FBS), 6 ng/ml bovine insulin, 25 mM Hepes, 26 mM $NaHCO_3$, 2 mM L-glutamine, 100 IU penicillin G sodium, 100 μg/ml streptomycin and 500 μg/ml Geneticin. Cells were harvested by scraping the cells off the monolayer with a rubber policeman. After centrifugation the cell pellet was frozen and stored at -70°C until RNA, DNA or nuclei were isolated. Nuclear extracts were always isolated from freshly harvested cells.

For experiments where cells were deprived of estrogens, 5% CS-FBS and PRF DMEM were used as the growth medium. To prepare the charcoal-coated dextran for treatment of the FBS, 5.0 g of acid-washed activated charcoal (Sigma) was washed with 100 ml sterile purified water at 4°C for 1 hour. The unsettled charcoal powder was decanted, the remaining charcoal was resuspended in 80 ml of sterile purified water and 0.5 g of Dextran T70 (Pharmacia LKB Biotechnology) was added slowly with stirring. This mixture was divided into two tubes and centrifuged in a Sorval SS34 rotor (DuPont New England Nuclear) at 10,000 rpm for 10 minutes. The pellet from one tube was added to 500 ml of FBS, stirred slowly for 1 hour at room temperature then placed at 55°C for 30 minutes. After centrifugation at 10,000 rpm for 20 minutes in a Sorval GSA rotor the FBS was decanted into a bottle, the second charcoal-coated dextran pellet was added and treatment was as previously described. Centrifugation to clarify the CS-FBS was followed by sequential filtration through Whatman #1 filters and 0.22 micron filters (Whatman International; Maidstone, England and Micron Separations; Westboro, Massachusetts, respectively) before filter sterilization.

The normal growth medium of the stock cells at 70-80% confluence was changed to PRF-DMEM supplemented with 5% CS-FBS, 2 mM L-glutamine, 0.3% glucose and penicillin-streptomycin (100 IU/ml and 100 μg/ml, respectively) for 2-3 days. These cells were passaged using PRF trypsin directly into fresh PRF-DMEM containing 5% CS-FBS and the previously described supplements. After 2-3 days, these cells were used to set up experiments by plating the cells at a density of $3\text{-}4\text{x}10^6$ cells per dish in 25 ml fresh PRF-DMEM containing 5% CS-FBS and the previously described supplements. After 24 hours the cells were treated with 25 μl absolute

ethanol (0 minute time point) or 25 µl $20x10^{-6}$ M 17-ß estradiol (Sigma) in absolute ethanol for various time periods.

Plasmid DNA Isolation

A Hae III-Acc I (-787 to -607 relative to P_1) c-*myc* DH site II_2 DNA fragment (Figures 15 and 27) was subcloned into the Sma I site of the multicloning region of pSp73 (Promega Biocan). This c-*myc* fragment could be excised using the Xho I and Bam H I restriction sites in the multicloning region of the plasmid. The Ava II-Cfr 101 (-1 to +140 relative to P_1) c-*myc* DH III_2 DNA fragment (Figures 15 and 27) was subcloned into the Sma I site of the multicloning region of pBluescript (Stratagene). This c-*myc* fragment could be excised using the Eco R I and Bam H I restriction sites in the multicloning region of the plasmid. The c-*myc* Pvu II-Pvu II 800 bp exon 1 fragment was also subcloned into the pSp73 vector while the c-*myc* Pst I-Pst I exon 2 fragment was a gift from Dr. R.P.C. Shiu (Dubik *et al.*, 1987) who obtained the clone from Dr. W.S. Hayward (Saito *et al.*, 1983). Clones for the thyroglobulin gene (Baas *et al.*, 1985) and 28s ribosomal RNA (Gonzalez *et al.*, 1985) were gifts, while the PRA/calcyclin clone was isolated by Dr. L.C. Murphy (Murphy *et al.*, 1988).

Plasmids were maintained and amplified in DH5α *E. coli* cells grown in Luria-Bertani medium (10 g Bacto-tryptone, 5 g Bacto-yeast extract, 10 g NaCl per liter, pH 7.5) in the presence of 200 µg/ml ampicillin (Sambrook *et al.*, 1989). Plasmid DNA was amplified at 37°C, overnight in the above medium in the presence of 20 µg/ml chloramphenicol (Clewell, 1972). After pelleting the bacterial cells by centrifugation in a Sorval GSA rotor at 5000 rpm for 15 minutes at 4°C, the plasmid DNA was isolated by alkaline lysis (Birnboim and Doly, 1979). This method consisted of the sequential addition (10 ml per 500 ml culture) of lysis buffer (25 mM Tris-HCl pH 8.0, 10 mM

EDTA, 50 mM glucose, 5 mg/ml lysozyme), alkaline-SDS (200 mM NaOH, 1% SDS) and acetate (either 3M NaAc pH 5.5 or 3 M potassium, 5 M acetate pH 4.8) with horizontal shaking and incubations on ice for 10 minutes, 10 minutes and 30 minutes, respectively. Centrifugation in a Sorval SS34 rotor at 15,000 rpm for 30 minutes was followed by room temperature precipitation of the resulting supernatant with 0.6 volumes of iso-propanol. After 30-60 minutes, the precipitated DNA was pelleted by room temperature centrifugation in the same rotor at 7500 rpm for 30 minutes, desiccated and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA pH 8.0).

The circular plasmid DNA was purified from nicked plasmid DNA and residual bacterial DNA by ultracentrifugation at 22°C for 18 hour at 55,000 rpm followed by 1 hour at 45,000 rpm in a Beckman Ti70.1 rotor (Beckman Instruments; Mississauga, Ontario) through a 1.1 gm/ml cesium chloride (Cabot; Revere, Pennsylvania) density gradient containing 833 μg/ml ethidium bromide (Radloff *et al.*, 1967). The plasmid DNA was illuminated with ultraviolet light and removed with a 22 gauge syringe. The ethidium bromide was removed by multiple extractions with iso-amyl alcohol (Davis *et al.*, 1986) and the plasmid DNA was precipitated overnight at -20°C with 200 mM NaAc and 2 volumes of absolute ethanol. After centrifugation in a Sorval SS34 rotor at 11,000 rpm for 30 minutes, the plasmid DNA was dried under vacuum and dissolved in TE buffer.

Oligonucleotides

Synthesis of the c-*myc* oligonucleotides (Regional DNA Synthesis Laboratory; Calgary, Alberta) was based on previously published human c-*myc* (Gazin *et al.*, 1984) and oligonucleotide sequences: S1, S2, C1, and C4 (Lang *et al.*, 1991); ME1a1, ME1a2 and E2F (Asselin, *et al.*, 1989; Hall, 1990; Hiebert, *et al.*, 1989; Lipp, *et al.*, 1989; Moberg *et al.*, 1991,1992; Plet, *et al.*,

1992; Thalmeier, *et al.*, 1989); and MBP1 (Ray and Miller, 1991). Although the ME1a1, ME1a1 and E2F elements were originally described for the mouse c-*myc* gene, the homologous sequences of the human c-*myc* gene were used in synthesizing the oligonucleotides. The consensus Sp1, NF1 and AP1 oligonucleotides were from the Hotfoot DNase I Footprinting Kit (Stratagene). The common names under which oligonucleotides O-II, O-II/III, O-III/IV and O-IV were synthesized and stored were U1-S3', U1, U2-E5' and U2, respectively.

Nucleotide sequences of the oligonucleotides were as follows:

S1 5'-TCGGGTACCCCCTGCCCCTCCCATATTCTCCC-3'

S2 5'-TGAGCAGGCGGGGCAGGAGGGGCGGTATCT-3'

C1 5'-TGCTGCTTTGGCAGCAAATTGGGGGACT-3'

C4 5'-CAGTCTGGGTGGAAGGTATCCAATCCAGAT-3'

O-II 5'-ACATAATGCATAATACATGACTCCCCC-3'

O-II/III 5'-ACATAATGCATAATACATGACTCCCCCCAACAAATGCAA3'

O-III/IV 5'-TCCCCCCAACAAATGCAATGGGAGTTTATTCATAACG-3'

O-IV 5'-GCAATGGGAGTTTATTCATAACGCGCTCTC-3'

ME1a1 5'-AGAACGGAGGGAGGGATCGCGCTGA-3'

ME1a2 5'-GCCTCGAGAAGGGCAGGGCTTCTC-3'

E2F 5'-AGGCTTGGCGGGAAAAA-3'

MBP-1 5'-AGGGATCGCGCTGAGTATAAAAGCCGGTTT-3'

Sp1 5'-GATCGATCGGGGCGGGGCGATC-3'

NF1 5'-ATTTTGGCTTGAAGCCAATATG-3'

AP1 5'-CTAGTGATGAGTCAGCCGGATC-3'

Labelling of DNA

The pSp73 and pBluescript plasmids, containing the DH site II_2 or DH site III_2 sequences described above, were linearized at the 5' or 3' end of

the c-*myc* insert DNA with one of the described restriction endonucleases in an overnight reaction, according to directions supplied by the manufacturer. The enzyme digestion was terminated with a 10 minute incubation at 65°C followed by sequential extractions with phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1). The linearized plasmids were then precipitated with 0.3 M NaAc and 2.5 volumes of 95% ethanol for a minimum of 10 minutes at -80°C and pelleted for 10 minutes in a microcentrifuge (International Equipment Company; Needham Hts., Massachusetts). The pellet was washed with 70% ethanol, the liquid removed and the DNA was dried in a speedvac (Savant Instruments, Hicksville, New York) for approximately 10 minutes. The DNA was end-labelled with the Klenow fragment of DNA polymerase I (Gibco BRL Life Technologies) and [α^{32}P]-TTP (Xho I site), [α^{32}P]-dGTP (Bam H I site) or [α^{32}P]-dATP (Eco R I) for 20 minutes at 37°C in nick translation buffer consisting of 50 mM Tris-HCl pH 7.5, 10 mM $MgSO_4$, 0.1 mM DTT and 50 µg/ml BSA (Boehringer Mannheim). The reaction was terminated by heating at 65°C followed by precipitation as described previously.

Alternatively, the pSp73 plasmid containing the c-*myc* DH site II_2 fragment was dephosphorylated twice with calf intestinal alkaline phosphatase (Boehringer Mannheim) in the presence of phosphatase buffer (20 mM Tris-HCl pH 8.0, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$) for 30 minutes at 37°C after being linearized with Bam H I. The reaction was terminated by heat inactivation for 10 minutes at 70°C before organic extraction and ethanol precipitation as described previously. The dephosphorylated DNA was end-labelled with T4 polynucleotide kinase (Pharmacia LKB Biotechnology) and [γ^{32}P]-ATP for 30 minutes at 37°C in kinase buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 50 µg/ml BSA). The

reaction was terminated with the addition of EDTA pH 8.0 to 15 mM followed by organic extraction and ethanol precipitation as described above.

All end-labelled DNA fragments were then digested at the opposing restriction enzyme site in the multicloning region of the plasmid or at an internal restriction site within the c-*myc* DNA for 2 to 4 hours according to the manufacturer's directions. For the smaller DH III_2 fragment, digestion at the internal Xho I restriction site was used (Figure 40). For the smaller DH site II_2 DNA fragments, digestion at the internal restriction sites Asp I, Alu I and Nsi I were used (Figure 29). The reactions were terminated with a 10 minute incubation at 65°C then loaded in the presence of 0.1% (w/v) bromophenol blue and 0.1% (w/v) xylene cyanol onto a nondenaturing 4 or 8% polyacrylamide gel (20% acrylamide, 1% bis-acrylamide) with 0.5x TBE (44.5 mM Tris-borate, 1 mM EDTA pH 8.5) as the running buffer. After electrophoresis at room temperature for 1 to 2 hours with 200-250 volts, the gel was exposed to Kodak (Eastman Kodak; Rochester, New York) XAR film for 5 to 10 minutes and the labelled DNA fragment was cut from the gel.

The DNA was passively eluted overnight at 37°C in 400 µl gel elution buffer (0.5 M NaCl, 20 mM Hepes pH 7.6, 1 mM EDTA). The next day the elution buffer was moved to a fresh microfuge tube, the gel pieces were washed with 100 µl of elution buffer and the combined buffer was precipitated with 2 volumes of 95% ethanol for 10 minutes at -80°C. The precipitated DNA was pelleted for 10 minutes in a microcentrifuge, resuspended in 100 µl TE buffer and reprecipitated with 0.3 M NaAc and 2 volumes of 95% ethanol as previously described. The pellet from this second precipitation was washed with 70% ethanol as described before, then air dried before redissolving in TE buffer.

Some of the c-*myc* fragments were isolated as insert DNA prior to any labelling reactions. The DNA fragments were excised from the vector DNA using the described restriction sites in the multicloning region of the plasmid. Electrophoresis through a 1% agarose gel with 1x TBE (89 mM Tris-borate pH 8.5, 2 mM EDTA) or 1x TAE (40 mM Tris-acetate, 1 mM EDTA) as the running buffer was used to size separate the insert DNA from the plasmid. DNA inserts separated by this method were cut from the gel and electroeluted (Girvitz *et al.*, 1980; Sambrook *et al.*, 1989) into a well lined with Spectra Por 3000 dialysis membrane (Spectrum; Houston, Texas) or a commercially available electroeluter (International Biotechnologies Inc.; New Haven, Connecticut). Alternatively, electrophoresis through a 1.5-2.0% low melting point agarose gel with 1x TAE as the running buffer was used. DNA inserts isolated in this manner were cut from the gel, melted at 65°C for 10 minutes then sequentially extracted three times with TE buffered phenol and twice with chloroform or ether. The DNA was subsequently precipitated with 0.3 M NaAc and 2.5 volumes of 95% ethanol as described before.

The Pvu II c-*myc* exon 1 and Pst I c-*myc* exon 2 inserts were nick translated (Amersham; Oakville, Ontario) or random prime labelled (Boehringer Mannheim) with [α^{32}P]-dCTP according to the manufacturer's directions. These radiolabelled c-*myc* fragments were separated from unincorporated [α^{32}P]-dCTP on a 5 ml G-50 sephadex (Pharmacia LKB) column using TE buffer as the carrier. The other c-*myc* fragments and cDNA inserts for PRA/calcyclin, thyroglobulin and 28s ribosomal RNA used in RNA or DNA hybridizations were also radiolabelled and purified in this manner. The c-*myc* inserts and the various oligonucleotides used in electrophoretic mobility shift assays were end-labelled with the appropriate radiolabelled dNTP and the Klenow fragment of DNA polymerase I as previously

described. These were separated from unincorporated radiolabelled dNTP using a G-25 Quick Spin column (Boehringer Mannheim) according to the supplied directions. The insert DNA representing the Nsi I-Acc I c-*myc* sequence and the oligonucleotides for this region (O-II, O-II/III, O-III/IV, O-IV) sometimes end-labelled poorly. This was possibly due to the formation of concatamers. To alleviate this problem these DNA fragments were heated at 95°C for 5 minutes then cooled to room temperature for 30 minutes prior to end-labelling as already described.

RNA Analysis

RNA was isolated by cell lysis in GITC lysis buffer (4.2 M guanidinium isothiocyanate, 0.5% (w/v) N-lauroyl sarcosine, 25 mM sodium citrate, 50 mM ß-mercaptoethanol, pH 7.0). The frozen cell pellet was vigorously mixed in 7-8 ml of GITC buffer, followed by approximately 10 passes through a 22 gauge needle to shear the DNA. After bringing the volume of GITC buffer to 25 ml, the RNA was purified through a 14 ml cesium chloride cushion (5.7 M CsCl, 100 mM EDTA, pH 7.5) at 22°C in a Beckman Ti70.1 rotor at 25,000 rpm for 22 hours (Chirgwin *et al.*, 1979). The GITC buffer was removed before decanting the remaining liquid and the RNA pellet was resuspended in 1 ml TE buffer heated to 65°C. The resuspended RNA was incubated at 65°C for 15 minutes, cooled on ice, then precipitated overnight at -20°C with 0.2 M NaAc and 2.5 volumes of absolute ethanol. The precipitated RNA was collected by centrifugation at 11,000 rpm in a Sorval SS34 rotor for 30 minutes at 4°C, followed by desiccation of the pellet and resuspension in TE buffer.

Poly-adenylate (poly A^+) RNA was isolated by oligo deoxythymidine (dT) cellulose (Pharmacia LKB Biotechnology) chromatography (Aviv and Leder, 1972). Oligo dT cellulose (0.1 g) was allowed to swell in 10 ml of

sterile 1x loading buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM EDTA), packed into a column and sequentially equilibrated with sterile washing buffer (100 mM NaOH, 5 mM EDTA), sterile purified water, and sterile 1x loading buffer. Each RNA sample in 1 ml of sterile TE buffer was boiled for 4 minutes then stored on ice for 3 minutes before adding an equal volume of sterile 2x loading buffer (40 mM Tris-HCl pH 7.5, 1 M NaCl, 2 mM EDTA). The RNA sample was loaded onto the column which was subsequently washed with sterile 1x loading buffer. The poly A^+ RNA was eluted with 4 ml prewarmed (65°C) sterile elution buffer (1 mM EDTA pH 7.5), the sample was collected on ice and immediately precipitated at -80°C with 200 mM NaAc and 2.2 volumes absolute ethanol. The column was reequilibrated prior to loading the next RNA sample. After an overnight precipitation, the poly A^+ RNA samples were centrifuged at 4°C for 60 minutes at 11,000 rpm in a Sorval SS34 rotor. The RNA pellets were desiccated and resuspended in 200 µl TE then precipitated a second time as already described.

Five to 10 µg of poly A^+ RNA was denatured in 50% (v/v) deionized formamide (Aldrich; Milwaukee, Wisconsin) and 2.2 M formaldehyde then size separated by overnight electrophoresis at 35 volts in 1.2% (w/v) agarose gels containing 2.2 M formaldehyde (Lehrach *et al.*, 1977) and 1x gel running buffer consisting of 200 mM MOPS (United States Biochemical; Cleveland, Ohio) pH 7.0, 50 mM NaAc and 5 mM EDTA. The RNA was transferred to nitrocellulose (Thomas, 1980) overnight using 20x SSC (see below) as a carrier for passive diffusion. Filters were baked for 2 hours at 80°C under vacuum and then prehybridized in hybridization solution (see below) for at least 3 hours. The filters were then hybridized with a ^{32}P-labelled human c-*myc* exon 2 cDNA (see Figure 19) or with a ^{32}P-labelled 28s ribosomal RNA

probe (Gonzalez *et al.*, 1985). Hybridizations, usually for 48 h, were performed at 42°C in the presence of 50% (v/v) deionized formamide, 0.1% (w/v) each of BSA, Ficoll and polyvinylpyrrolidine (Denhardt, 1966), 5x SSPE (1x SSPE: 1.15 M NaCl, 10 mM NaH_2P0_4, 1 mM EDTA pH 7.4), 250 μg/ml salmon sperm DNA (sheared and heat denatured at 100°C for 10 minutes) and 0.1% (w/v) SDS. At the end of the hybridization period the filters were washed twice in 2 x SSC, 0.1% SDS (1x SSC: 150 mM NaCl, 15 mM Na-citrate pH 7.0) for 15-30 minutes at room temperature, followed by three 20 minute washes in 0.1 x SSC, 0.1% SDS at 65°C. Filters were exposed to Kodak XAR film at -80°C with an intensifying screen.

RNase Protection Assay

This was essentially performed as described previously (Dotzlaw *et al.*, 1990) with some minor modifications. A radiolabelled antisense c-*myc* Pvu II-Pvu II exon 1 RNA (see Figure 24) was synthesized from 1μg of Hind III linearized template DNA with 10 units of T7 RNA polymerase (Pharmacia) in the presence of 40 units RNAsin (Promega/Biocan), 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 6.25 μM [α^{32}P]-CTP and 400 μM each of ATP, GTP, UTP (all unlabelled ribonucleotides were from Boehringer Mannheim and stored in 50 μM EDTA). Synthesis occurred for 60 minutes at 37°C followed by incubation for an additional 15 minutes in the presence of 10 units RNase-free DNase I (Promega Biocan). The antisense RNA was extracted once with phenol-chloroform-isoamyl alcohol (25:24:1) in the presence of 20 μg yeast tRNA then precipitated three times with 2 M NH_4Ac and 3 volumes of absolute ethanol for 15 minutes at -80°C.

The radiolabelled antisense RNA was resuspended in 100 μl of 1x hybridization buffer (80% (v/v) deionized formamide, 40 mM Pipes pH 6.4, 0.4 M NaCl, 1 mM EDTA). Fifteen μg of each sample RNA, isolated as described

previously but without poly A^+ selection, was precipitated with 200 mM NaAc and 2 volumes of absolute ethanol. The pellets were dissolved in 29 µl of 1x hybridization buffer and 1 µl of the antisense RNA probe was added to each sample prior to denaturation at 80°C. The hybridization proceeded at 50°C for 16-18 hours. Single stranded RNA was removed from each reaction by the addition of 350 µl of buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM EDTA) containing 40 µg/ml RNase A and 2 µg/ml RNase T1 (both ribonucleases were from Boehringer Mannheim) and incubation for 60 minutes at 37°C. The reaction was terminated by the addition of 10 µl 20% SDS and 2.5 µl 20 mg/ml proteinase K (Boehringer Mannheim) followed by a further 15 minute incubation at 37°C. The samples were extracted once with 400 µl phenol-chloroform-iso-amyl alcohol (25:24:1) and precipitated for 15 minutes at -80°C in the presence of 25 µg tRNA and 1 ml absolute ethanol. Precipitated RNA was collected by microcentrifugation for 15 minutes at 4°C, desiccated, then dissolved in 10 µl loading buffer (80% (v/v) deionized formamide, 1 mM EDTA, 0.1 % (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol). The samples were denatured for 5 minutes at 85°C prior to electrophoresis (5% polyacrylamide, 7 M urea sequencing gel) for 2 hours at 1700 volts with 1x TBE as the running buffer. When electrophoresis was complete the gels were dried on Whatman paper and exposed to Kodak XAR film at -80°C with an intensifying screen.

DNA Analysis and Endogenous Methylation

Genomic DNA was isolated from ER^- BT 20, MDA 468, MDA MB 231 and ER^+ T47 D5, ZR 75-1, MCF 7 breast cancer cell lines stored at -80°C as cell pellets. The cell pellets were thawed on ice then resuspended in 3 ml digestion buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS). Proteinase K was added to a final concentration of 100 µg/ml and

allowed to incubate for 20 hours at 37°C. Another 3 ml of digestion buffer and 100 μg/ml proteinase K were added and the incubation was continued for 2 more hours at 50°C. The mixture was extracted twice with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) followed by two extractions with chloroform-isoamyl alcohol (24:1). The DNA was precipitated overnight at -20°C with 200 mM NaCl and 2 volumes of absolute ethanol. DNA was pelleted in a Sorval SS34 rotor at 10,000 rpm for 30 minutes at 4°C, washed with 70% ethanol and dried under vacuum for 5-10 minutes. The DNA was resuspended in 1.5 ml TE buffer, incubated for 2 hours at 65°C then stored overnight at 4°C. Residual RNA was removed by incubation for 60 minutes at 37°C in 0.1% (w/v) SDS and 0.1 mg/ml RNase A. The samples were extracted once with phenol-chloroform-isoamyl alcohol (25:24:1) and twice with chloroform-isoamyl alcohol (24:1) then precipitated as previously described and resuspended in 500 μl TE buffer.

Fifteen μg of each DNA sample was digested overnight with 30 units of Bgl II according to the manufacturer's directions and then incubated for 60 minutes with an additional 20 units of enzyme the next day. DNA was size separated by electrophoresis for 19 hours at 35 volts on 1% agarose with 1x TBE (89mM Tris-borate, 2mM EDTA) as the running buffer. The DNA was transferred to nitrocellulose by passive diffusion (Southern, 1975) using an alkaline buffer (Smith and Summers, 1980) as a carrier. Prior to transfer the DNA was denatured and neutralized by soaking the agarose gel at room temperature for 60 minutes in denaturation buffer (1.5 M NaCl, 500 mM NaOH), followed by an additional 60 minutes in neutralization buffer (1 M NH_4Ac, 20 mM NaOH). Neutralization buffer was used as the alkaline carrier for the overnight transfer of the fractionated DNA to nitrocellulose. The nitrocellulose filter was baked, hybridized and washed as described for

RNA analysis except the ^{32}P-labelled DNA probes were the Pst I-Pst I c-*myc* exon 2 fragment (Dubik et. al., 1987), the PRA/calcyclin cDNA (Murphy *et al.*, 1988) and the thyroglobulin cDNA (Baas *et al.*, 1985).

Twenty five µg aliquots of MCF 7 (ER^+) and MDA MB 231 (ER^-) genomic DNA were digested overnight with 50 units of the restriction enzymes Pst I, Pst I / Hpa II, Pst I / Msp I, and Pst I / Hha I according to the manufacturer's directions. DNA was size separated by electrophoresis and transferred to nitrocellulose as described above. The blots were hybridized with either a ^{32}P-labelled Pst I-Taq I c-*myc* fragment, which abuts the 5' end of the 1700 bp Pst I-Pst I c-*myc* P_0 fragment, or a ^{32}P-labelled Apa I-Pst I c-*myc* fragment, which abuts the 3' end of the same Pst I-Pst I fragment (Figure 26). Buffer and temperature conditions for hybridizations and washes were as previously described for RNA analysis.

DNase I Hypersensitivity Assay

All steps were performed on ice unless otherwise stated. Five ml of ice cold cell lysis buffer (CLB: 10 mM Tris-HCl pH 7.5, 100 mM KCl, 10 mM sodium butyrate, 5 mM $MgCl_2$, 0.1% (v/v) NP40) was added to each frozen cell pellet which contained 50-100 x 10^6 cells. To this was added 100 µl iodoacetamide (0.1 g/ml), 50 µl PMSF (100 mM) and 50 µl aprotinin (0.1 mg/ml) and the cell pellet was thawed and resuspended gently in the above solution using a cut-off pasteur pipette. The suspension was then transferred to a loose-fitting (B) glass/glass Dounce homogenizer and hand homogenized (8-10 strokes). The homogenate was centrifuged at 2,000 x g for 10 minutes at 4^{o}C, and the supernatant was discarded. Five ml of ice cold CLB without NP40 was added to the pellet together with the above mix of protease inhibitors, and the pellet was resuspended gently with a cut-off pasteur pipette. The suspension was homogenized as before (3 strokes) and the

homogenate centrifuged as described above. The supernatant was discarded and the pellet was resuspended and homogenized as before in 5 ml of ice cold nuclei preparation buffer (NPB: 50 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM sodium butyrate, 5 mM $MgCl_2$) to which was added the previously described mix of protease inhibitors (iodoacetamide, PMSF, aprotinin). The homogenate was centrifuged at 2000 x g for 10 minutes at 4°C. The supernatant was discarded and the pellet containing the intact nuclei was resuspended gently in 2 ml per 50 x 10^6 cells of sucrose cushion solution (SCS: 1.8 M sucrose, 50 mM Tris-HCl pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 10 mM sodium butyrate) to which was added the above mix of protease inhibitors. The nuclei in SCS (2 ml) were then layered carefully onto a 9 ml cushion of SCS and ultracentrifuged in a Beckman SW40 rotor at 64,000 x g for 30 minutes at 4°C. The pellet was gently resuspended in 5 ml of NPB and centrifuged at 2000 x g for 10 minutes at 4°C. The resulting pellet was resuspended in 2 ml of NPB and this preparation was then subjected to DNase I digestion (DNase I type II; Sigma D-4263).

An aliquot of the preparation was taken and the DNA concentration determined using the fluorescent dye (Hoescht 33258) method and a TKO mini-fluorometer (Hoeffer Scientific Instruments; San Francisco, California) according to the manufacturers instructions (modified from Cesarono *et al.*, 1979). The nuclei were diluted in NPB to 2500 μg DNA per ml and $CaCl_2$ was added to 1 mM. Half ml aliquots were added to 1.5 ml microcentrifuge tubes maintained on ice and buffer or concentrated DNase I solutions were added to achieve a final concentration of 0-20 units DNase I per ml of nuclei. Each tube was treated individually, with a 1 minute preincubation at 37°C followed by addition of the appropriate DNase I solution and incubation with gentle mixing for a further 3 minutes at 37°C. The reactions were stopped by

the addition of EDTA to 10 mM and stored on ice until all tubes had been treated. The contents of each tube were transferred to 15 ml polypropylene centrifuge tubes and the volume of each made to 1.54 ml with purified water. To each tube was added 80 µl 250 mM EDTA, 80 µl 4 M NaCl, 200 µl 22%, (w/v) sarcosyl and 100 µl 10 mg/ml pronase (Boehringer Mannheim) and the contents incubated at 37°C overnight in a shaking waterbath.

The next day 40 µl of RNase A (5 mg/ml) was added to each sample and the incubation at 37°C was continued for 2 hours. Each sample was extracted three times with phenol-chloroform-isoamyl alcohol (25:24:1) at room temperature for 60 minutes followed by two extractions with chloroform-isoamyl alcohol (24:1). The DNA was precipitated overnight at -20°C with 0.2 M NaAc pH 5.5 and 2 volumes of ice cold ethanol. The precipitated DNA was recovered by centrifugation, desiccated briefly and dissolved in 100 µl TE buffer. Thirty µg of DNA was digested at 37°C overnight with 60 units of Bgl II as previously described. An additional 60 units of Bgl II was added the next morning and the incubation was allowed to continue at 37°C for a further 2 hours. The restriction enzyme digested DNA was size separated by agarose gel electophoresis and transferred to nitrocellulose as described previously. Nitrocellulose filters were hybridized to the Pst I-Pst I c-*myc* exon 2 DNA fragment (see Figure 19) which was labelled with [α^{32}P]-dCTP by random priming. Hybridization conditions were as previously described for RNA analysis.

Nuclear Protein Isolation

Nuclear extracts were isolated from MCF 7 and MDA MB 231 breast cancer cell lines by a modification of a previously described method (Dignam *et al.*, 1983). All manipulations were performed on ice and with reagents chilled to 4°C, unless otherwise noted. Cells were harvested at subconfluence

by scraping with a rubber policeman then pelleted at 4°C and washed with Isoton (Coulter Electronics; Surrey, British Columbia). The cells were resuspended in 5 packed cell volumes (PCV) of buffer A (10 mM Hepes pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$), the proteinase inhibitors 1 mM PMSF, 2 μg/ml aprotinin, 1 μg/ml leupeptin and 40 μg/ml bestatin were freshly added along with 0.5 mM DTT, and the mixture was incubated for 10 minutes on ice. NP40 was added to 0.25% (v/v) and the cells were lysed with 4-5 strokes of a glass B pestle in a glass Dounce homogenizer. After centrifugation for 10 minutes at 2000 rpm in a Sorval SS34 rotor, the nuclear pellet was washed with 5 PCV then 2 PCV of buffer A with all previous additions except for the NP40.

The nuclei were resuspended to 200 A_{260} units per ml in buffer C (20 mM Hepes pH 7.9, 25% glycerol, 0.1 M KCl, 0.2 mM EDTA) with PMSF, aprotinin, leupeptin, bestatin and DTT added as before, then KCl was added to a final concentration of 0.42 M. The nuclei were extracted by gentle mixing for 30 minutes at 4°C. The extracted nuclei were pelleted at 12,000 x g for 30 minutes at 4°C and the supernatant was dialyzed against buffer D (20 mM Hepes pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA with freshly added 0.5 mM DTT and 1 mM PMSF). Dialysis was for 60 minutes at 4°C using 6000-8000 MW cut off Spectra Por dialysis membrane (Spectrum; Houston, Texas) in a microdialyzer (Integrated Separation Systems; Hyde Park, Massachusetts). The dialyzed nuclear extract was centrifuged for 10 minutes at 12,000 x g to remove any precipitates, aliquoted on ice, snap frozen in a dry ice-ethanol bath or liquid nitrogen and stored at -80°C. Nuclear protein concentrations were measured spectrophotometrically (Bradford, 1976) using reagents which were commercially available (Bio-Rad Protein Assay, catalogue # 500-0006).

Electrophoretic Mobility Shift Assay

Nuclear extracts (2-30 μg per 20 μl reaction) from MCF 7 or MDA MB 231 breast cancer cells were preincubated on ice for 5-10 minutes with 2 μg poly dI-dC in binding buffer (20 mM Hepes pH 7.6, 0.1 M KCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 20% glycerol, 0.05% NP40, 0.1 mM $ZnCl_2$). For competition assays the competitor DNA was also added at the preincubation step. Radiolabelled DNA (0.25 to 1.0 ng) was then added to the reaction (10 μl final volume) and the incubation was continued on ice for 30 minutes. This order of probe and nuclear extract addition has been suggested to minimize nonspecific protein-DNA interactions (Demczuk *et al.*, 1991). The DNA-protein complexes were resolved on 4% polyacrylamide gels which were prerun for 60 minutes at 200 volts using 0.5x TBE (45 mM Tris-borate, 1mM EDTA) as the running buffer. The gels were run at 4°C with 200-250 volts until the labelled probe was near the bottom of the gel (1-2 hours). The gels were then dried on Whatman paper and exposed to Kodak XAR film at -80°C with an intensifying screen.

Stairway Assay

This assay was based on a previous modification of the electrophoretic mobility shift assay (van Wijnen *et al.*, 1992). In brief, the stairway assay differs from the electrophoretic mobility shift assay in that it involves a comparison of DNA fragments that have sbeen equentially deleted from one end of a larger DNA fragment. This deletion analysis assists in mapping protein binding sites to a particular region of the larger DNA fragment. The Hae III-Acc I (-787 / -607 relative to P_1) c-*myc* DH site II_2 DNA fragment (Figure 29) was end-labelled at the Xho I or Bam H I sites of the pSp73 vector as previously described, then restriction digested with the opposite endonuclease. After purification of the labelled DNA from a 4%

nondenaturing polyacrylamide gel, the c-*myc* fragments were shortened by endonuclease digestion at the internal restriction sites Asp I, Alu I and Nsi I and gel purified as previously described. These labelled fragments were used as probes for the stairway assay. Conditions for protein-DNA complex formation and analysis were as described for the electrophoretic mobility shift assay.

***In vitro* Footprinting**

Nuclear extracts were incubated as described for the electrophoretic mobility shift assay except the reaction contained 30 to 60 µg of crude nuclear extract and 2-5 ng of DNA radiolabelled at only one end of the fragment. Additionally, the 40 µl reaction was incubated at room temperature instead of on ice and samples were treated with 1 µl of 0.01 to 0.4 units of DNase I for 2 minutes. The DNase I digestion was terminated with 150 µl of stop buffer (1% SDS, 100 mM KCl, 0.02 mM EDTA) containing 10 µg/µl proteinase K and 5 µg/µl tRNA then incubated for 60 minutes at 55°C. The samples were extracted one time with phenol-chloroform-isoamyl alcohol (25:24:1) then precipitated with 2.5 volumes of ethanol. Pellets were resuspended in gel loading buffer (80% (v/v) deionized formamide, 1 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol) to 10,000 cpm per µl then denatured for 3-5 minutes at 95-99°C and briefly stored on ice prior to electrophoresis. Equal counts were loaded onto a prewarmed 6% polyacrylamide (43.5% acrylamide, 1.5% bis-acrylamide), 8 M urea sequencing gel run for at least 60 minutes at 50°C with 60-65 watts and 1x TBE as the running buffer. When electrophoresis was complete the gels were dried on Whatman paper and exposed to Kodak XAR film at -80°C with an intensifying screen.

The G and G+A chemical sequencing reactions (Maxam and Gilbert, 1980) were performed to assist in characterizing the DNase I protected

region. For the G+A reaction the radiolabelled DNA and 1 μg of poly dI-dC as a carrier DNA were brought to 30 μl with purified water. One μl of 1 M formic acid was added to the DNA and the reaction was incubated for 25 minutes at 37°C. Piperidine was freshly diluted to 1 M, 150 μl was added to the G+A reaction and the reaction was then stored on ice. For the G reaction, radiolabelled DNA and 1 μg of poly dI-dC as a carrier DNA were brought to 7 μl with purified water, 7 μl of cacodylate buffer (50 mM sodium cacodylate, 1 mM EDTA pH 8.0) and 100 μl freshly made dimethyl-sulfate solution (1 ml cacodylate buffer, 10 μl DMS) was added in the fume hood. This reaction mix was incubated at room temperature for 1 minute, 50 μl of DMS stop buffer (1.5 M NaAc pH 7.0, 1.0 M ß-mercaptoethanol, 100 μg/ml tRNA) was added and the DNA was immediately precipitated with 750 μl absolute ethanol for 5 minutes on dry ice. The ethanol was removed with a pasteur pipette, the DNA was resuspended in 100 μl purified water and 1 ml n-butanol was added. After extensive mixing, the DNA was pelleted by briefly microcentrifuging at room temperature. The resuspension and n-butanol precipitation was repeated, 150 μl 1 M piperidine was added and the G reaction was stored on ice. Both the G and G+A reaction were incubated at 90°C for 30 minutes to cleave the DNA at the chemically modified nucleotides. The reactions were cooled on ice then precipitated by extensive mixing in the presence of 1.2 ml n-butanol as already described. The DNA was resuspended by mixing in 150 μl purified water, moved to a clean microfuge tube and reprecipitated with extensive mixing in the presence of 1 ml n-butanol. The sequencing reactions were resuspended in gel loading buffer and run alongside the DNase I footprinting reactions as described above.

Methylation Interference

End labelled DNA was methylated at G residues by performing the chemical modification reaction described for chemical cleavage sequencing without cleaving the DNA with piperidine (Maxam and Gilbert, 1980; Ausubel *et al.*, 1989). Radiolabelled DNA and 1 μg of poly dI-dC as a carrier DNA were brought to 7 μl with purified water and 7 μl of cacodylate buffer (50 mM sodium cacodylate, 1 mM EDTA pH 8.0) was added. In the fume hood 1 μl of undiluted DMS was added and the reaction allowed to incubate for 1 minute at room temperature. The reaction was stopped with 40 μl of DMS stop buffer then purified by repeated ethanol precipitations on dry ice. The radiolabelled methylated DNA was then used in the electrophoretic mobility shift assay (EMSA) as described previously.

Results

Different mechanisms regulate c-*myc* expression in hormonally dependent and independent human breast cancer cells (Dubik and Shiu, 1988). In particular, c-*myc* is stimulated by estrogens in responsive cells but constitutively expressed in nonresponsive cells. This gene therefore provides an excellent model for analyzing the molecular mechanisms responsible for these regulatory differences. An analysis of the chromatin structure of the c-*myc* gene revealed increased DNase I hypersensitivity of DH site II_2, which is associated with the P_0 promoter, in estrogen receptor negative breast cancer cells. This difference was not due to amplification, alterations in promoter usage or differences in DNA methylation. Instead, it was determined that a different level or complement of DNA binding proteins associated with DH sites II_2 and III_2 c-*myc* chromatin existed between the ER^+ MCF 7 and ER^- MDA MB 231 cells. The cooperative interactions of Sp1

proteins bound at multiple sites in the c-*myc* 5' flanking sequences may provide a mechanism for the association of these two regulatory regions.

c-*myc* Expression

Estrogens stimulate the expression of the c-*myc* proto-oncogene in both the rat uterus (Murphy *et al.*, 1987) and estrogen responsive human breast cancer cell lines (Dubik *et al.*, 1987). The estrogen stimulation of c-*myc* in the ER^+ MCF 7 and T47 D breast cancer cells occurred at the transcriptional level (Dubik and Shiu, 1988). The level of c-*myc* transcription increased within 5 minutes of estrogen stimulation and was maximal by 20 minutes. Concomitant increases in mRNA accumulation were observed within 15 minutes and were maximal by 60 to 120 minutes. This response was specific to estrogens, since the steroid hormones dexamethasone, dihydrotestosterone, medroxyprogesterone acetate and retinoic acid, had little effect on c-*myc* mRNA accumulation (Dubik and Shiu, 1992). Additionally, it was previously determined that expression of c-*myc* mRNA in ER^- breast cancer cell lines was constitutive and unaffected by the presence or absence of estrogens (Dubik and Shiu, 1988).

The growth conditions previously used to monitor estrogen activation of c-*myc* in breast cancer cell lines utilized DMEM containing the estrogenic contaminants present in the pH indicator phenol red (Berthois *et al.*, 1986), charcoal stripped FBS and the antiestrogen, tamoxifen (Dubik *et al.*, 1987; Dubik and Shiu, 1988), which has some intrinsic agonist activity. It was therefore important to determine if estrogen stimulation of c-*myc* followed a similar time course in the estrogen depleted growth conditions used in the studies described in this thesis. As described in the Methods and Materials, MCF 7 cells grown to 70-80% confluence under normal growth conditions were changed to phenol red free-DMEM containing 5% dextran-charcoal

stripped FBS for 2-3 days. The cells from each 150 cm^2 flask were passaged into three 150 cm^2 flasks containing fresh estrogen depleted medium using phenol red-free trypsin/EDTA. These cells were used to set up experiments in fresh estrogen depleted medium 2-3 days later. After treatment with ethanol (0 minute time point) or 10 nM 17-ß estradiol for 10, 15, 30, 60, 120 and 180 minutes, the MCF 7 cells were harvested, poly A+ enriched RNA was isolated and the level of c-*myc* determined by Northern blot analysis.

Estrogen stimulated a marked increase in the steady state levels of c-*myc* mRNA at the 30 minute time point (Figure 17). A maximal increase of 7 fold was observed at 60 minutes after treatment (Figure 17). Subsequently, the c-*myc* mRNA level decreased but remained substantially elevated above the untreated level until at least 180 minutes, which was the last time point measured. This time course of estrogen stimulation was similar to that previously reported (Dubik *et al.*, 1987; Dubik and Shiu, 1988).

The steady state level of c-*myc* mRNA was monitored in a number of exponentially growing human breast cancer cell lines (Figure 18). Under normal growth conditions c-*myc* expression was found to vary but did not correlate with ER status. Low steady state levels were detected in the ER^+ cell line T47 D and the ER^- cell lines MDA 468 and BT 20. Conversely, high levels of c-*myc* mRNA were observed in the ER^- cell line MDA MB 231, and the ER^+ cells ZR 75-1, T47 D5 and MCF 7. Interestingly, the steady state levels of c-*myc* mRNA roughly correlated inversely with the mean doubling times of the cell lines (given in days as the mean +/- standard error, n=3). The cells with the lowest mean doubling times were MDA MB 231 (1.45 +/- 0.01), ZR 75-1 (1.96 +/- 0.01), T47 D5 (1.33 +/- 0.01) and MCF 7 (1.27 +/- 0.01), while longer mean doubling times were detected in T47 D (2.09 +/- 0.04),

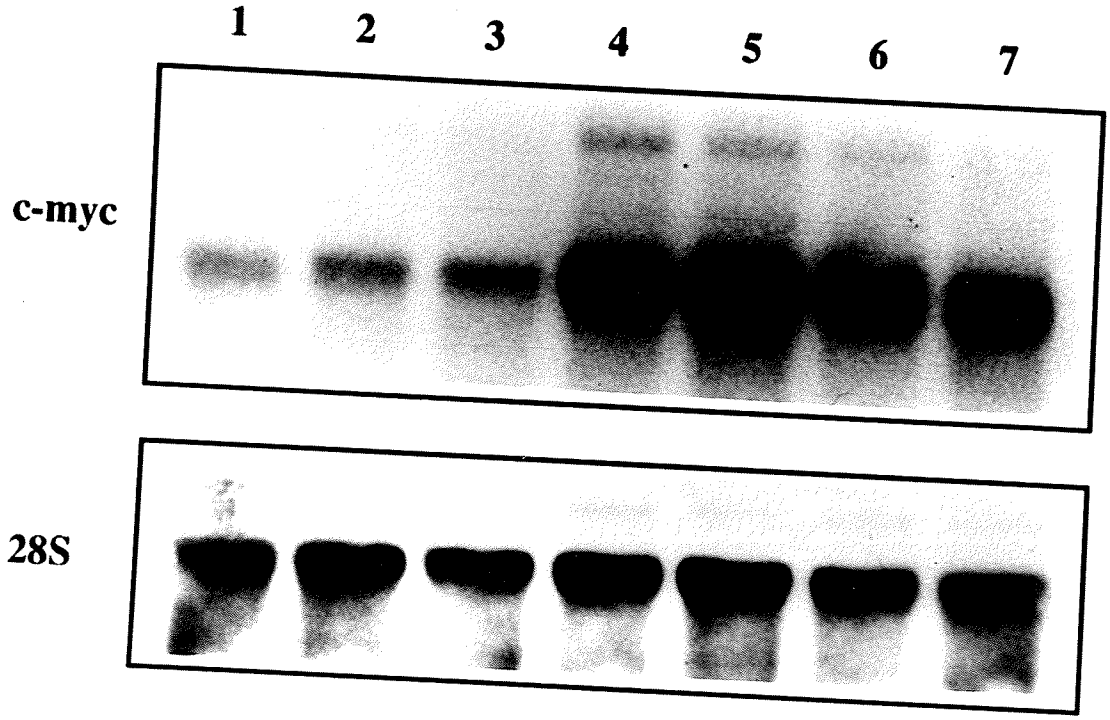


Figure 17. Estrogen Regulated Expression of c-*myc* mRNA. Messenger RNA was isolated from MCF 7 cells which had been grown in medium which was depleted of estrogen and left untreated or supplemented with 17 ß-estradiol for various time periods, as described in Methods and Materials. Eight μg of poly A^+ RNA from untreated cells (lane 1) or cells treated with 10 nM 17 ß-estradiol for 10 min (lane 2), 15 min (lane 3), 30 min (lane 4), 60 min (lane 5), 120 min (lane 6) and 180 min (lane 7) were separated by electrophoresis and transferred to nitrocellulose also as described. The top and bottom panels represent the pattern of hybridization obtained using a radiolabelled human c-*myc* exon 2 cDNA (7 hour exposure) and a 28S ribosomal RNA probe (2 hour exposure), respectively.

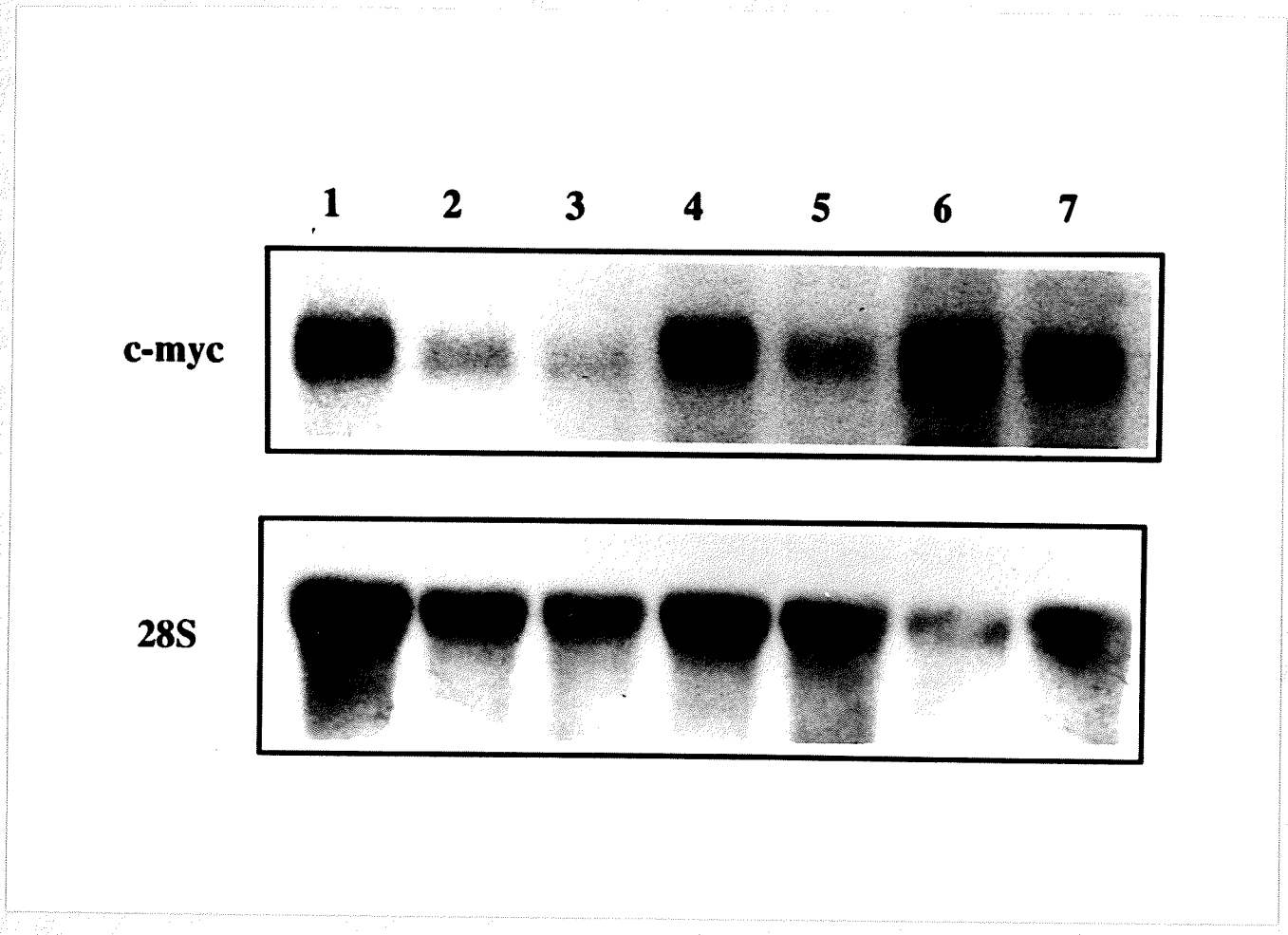


Figure 18. c-*myc* Expression in Human Breast Cancer Cell Lines. Steady state mRNA isolated from exponentially growing breast cancer cell lines was isolated and analyzed for levels of c-*myc* expression as described in Methods and Materials. Twelve μg of poly A^+ RNA from MDA MB 231 (lane 1), MDA 468 (lane 2), BT 20 (lane 3), ZR 75-1 (lane 4), T47 D (lane 5), T47 D5 (lane 6), and MCF 7 (lane 7) were hybridized to a radiolabelled c-*myc* exon 2 cDNA (top panel, 18 hour exposure) and 28S ribosomal RNA probes (bottom panel, 30 minute exposure).

MDA 468 (2.0 +/- 0.02) and BT 20 (4.09 +/- 0.23) (Murphy and Dotzlaw, 1989c).

Differential DNase I Hypersensitivity

Regulation of the c-*myc* gene is closely associated with regulation of cell growth. Significantly, estrogen stimulates the expression of c-*myc* and the proliferation of breast cancer cells. Breast cancer cells which are ER^+ or ER^- reflect the extreme ends of the spectrum of hormonally dependent and hormonally independent cells. A comparison of these cells may provide molecular clues to the transition from hormonal dependence to independence in human breast cancer. Initially, a detailed examination of the chromatin structure associated with the estrogen activated c-*myc* gene compared to that of the constitutively activated gene in hormonally independent human breast cancer cells was undertaken. To clarify the role of estrogen in the development of these molecular differences, the c-*myc* chromatin of estrogen stimulated breast cancer cells was also analyzed. Additionally, the c-*myc* chromatin of a cell line in which an ER expression vector was stably introduced and a cell line which had progressed to the ER^- phenotype was compared to their parental cells.

The chromatin structure of the c-*myc* gene, as assessed by DNase I hypersensitivity of the 5' flanking region, was examined in human breast cancer cells. The location and intensity of the DH sites were characterized and compared between ER^+ and ER^- cell lines. Initially, nuclei isolated from the ER^+ cell line, ZR 75-1, and the ER^- cell line, MDA MB 231, were subjected to digestion with increasing concentrations of DNase I. DNA purified from these digests was further fragmented with the restriction enzyme Bgl II, size separated by electrophoresis, transferred to nitrocellulose and hybridized to a c-*myc* exon 2 probe (Figure 19 A and B). Increasing the

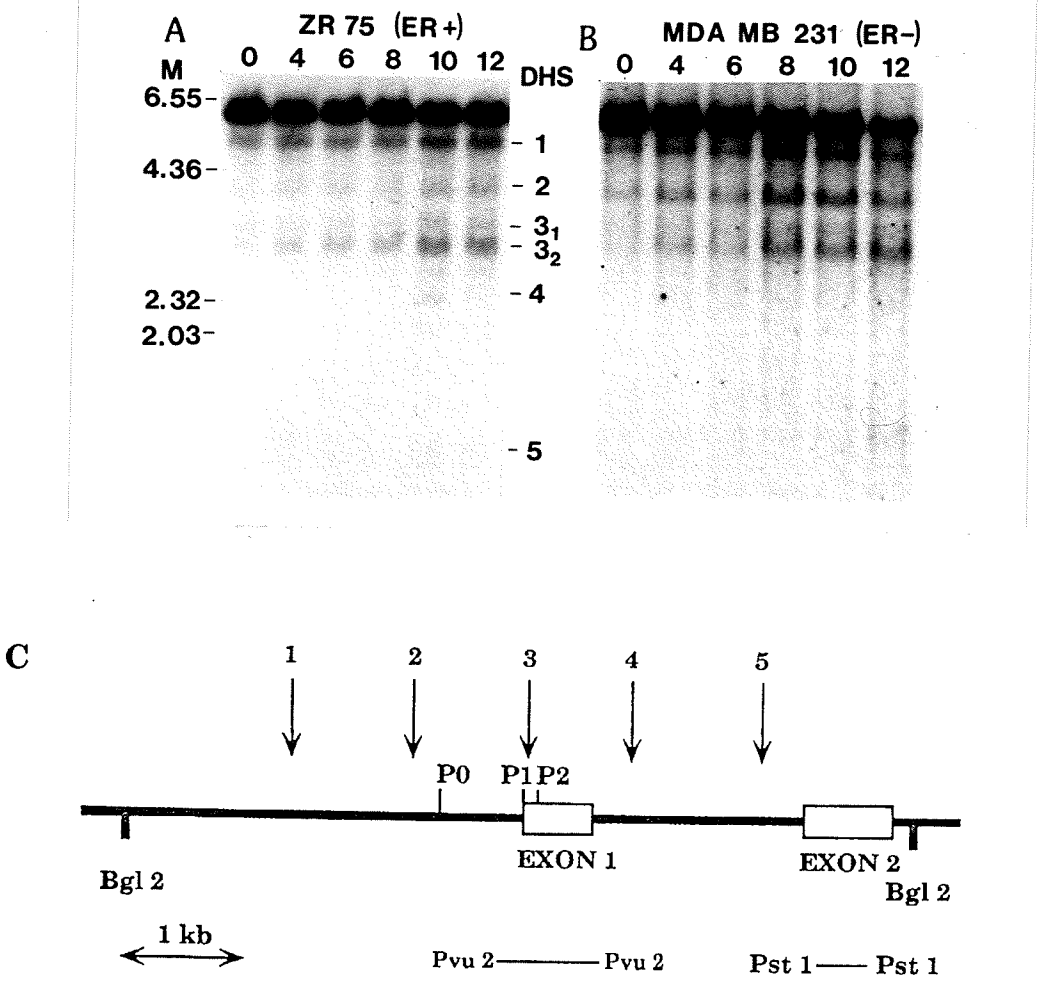


Figure 19. DNase I Hypersensitive Sites of c-*myc* Gene Chromatin of ZR 75-1 and MDA MB 231 Breast Cancer Cell Lines. Nuclei isolated from an ER^+ cell line, ZR 75-1 **(A)** and an ER^- cell line, MDA MB 231 **(B)**, were subjected to increasing concentrations of DNase I (0 to 12 units/ml) as indicated at the top of each lane. Following DNase I digestion the DNA was isolated and subjected to Bgl II digestion followed by DNA analysis with a c-*myc* exon 2 probe, as described in Methods and Materials. M refers to the DNAfragment length standards (lambda phage DNA digested with Hind III) in kilobase pairs (kbp). DHS refers to the DNase I hypersensitive sites (1 or I, 2 or II_2, 3_1 or III_1, 3_2 or III_2, 4 or IVb, and 5 or V) which are lower molecular weight hybridizing bands appearing below the 6.2 kbp parent band. **(C)** The positions of P0, P1 and P2 in the human c-*myc* gene as well as the positions of DNase I hypersensitive sites 1, 2, 3 (representing III_1 and III_2), 4, and 5. The Pst 1 fragment shown is the c-*myc* exon 2 probe used in these studies. These studies were done with technical assistance from Tim Salo (Miller *et al.*, 1993).

concentration of DNase I resulted in the detection of the 6.2 kbp parent c-*myc* fragment, as well as fragments of lower molecular mass, which were indicative of nuclease activity at DH sites (Figure 19 C). A low but detectable level of DNA fragmentation in the control digestion (no added DNase I) was consistent with the presence of an endogenous nuclease activity in these nuclei (Vanderbilt *et al.*, 1982). A number of DH sites were identified in both the ER^+ ZR 75-1 and the ER^- MDA MB 231 cell lines. None of these DNase I hypersensitive sites however were unique to breast cancer cell lines and have all been described in the HL60 cell line (Bentley and Groudine, 1986a; Siebenlist *et al.*, 1988). DNA fragments of 5.2, 4, 3.4, 3.1, 2.4, and 1.4 kbp were identified in both the ZR 75-1 and the MDA MB 231 breast cancer cells which corresponded to DH sites I, II_2, III_1, III_2, IVb and V, respectively, in HL60 cells (Bentley and Groudine, 1986a; Siebenlist *et al.*, 1988).

In both the ER^+ ZR 75-1 and ER^- MDA MB 231 breast cancer cells the intensity of DH III_2 was very much greater than DH III_1. Significantly, the intensity of DH II_2 was different between the two human breast cancer cell lines. As determined by the increased signal intensity relative to DH site I, DH site II_2 was considerably more sensitive to DNase I digestion in the c-*myc* chromatin of the ER^- MDA MB 231 cells than the ER^+ ZR 75-1 cell line. Densitometric scanning indicated the intensity of DH site II_2 of c-*myc* chromatin in MDA MB 231 cells approached that of DH site I, with the intensity of DH site II_2 being 0.72 of DH site I. In contrast, DH site II_2 of ZR 75-1 cells was approximately 0.32 the intensity of DH site I. Therefore, differential accessibility of DH site II_2 was observed between these two ER^- and ER^+ breast cancer cell lines.

To assess the generality of this phenomenon, the DH site pattern in other ER^+ and ER^- breast cancer cell lines was also investigated. A

comparison was made of two additional ER+ cell lines, MCF 7 and T47 D5, and two additional ER- cell lines, MDA 468 and BT 20, for nuclease accessibility of these DH sites (Figure 20 A). The prominence of DH II_2 relative to DH I was consistently observed to be less in all of the ER+ cell lines compared to the ER- cell lines. Relative to DH I, the intensities of DH II_2 of c-*myc* gene chromatin of MCF 7, T47 D5, MDA 468, and BT 20 cells were 0.33, 0.30, 0.48, and 0.43, respectively. This indicated that the observation of differential sensitivity of c-*myc* chromatin at DH II_2 was a general phenomenon which appeared to be associated with the ER status of breast cancer cell lines.

The relationship between ER status and DH II_2 sensitivity suggested there may be a specific effect of estrogen on the DH pattern in MCF 7 cells. Estrogen treatment of MCF 7 cells, which had been maintained under essentially estrogen free conditions (phenol red free, twice charcoal-dextran stripped FBS), had been shown previously to significantly increase c-*myc* mRNA accumulation (Figure 17). Using these same conditions the DH pattern in untreated cells and cells treated with estrogen for varying periods of time (0-30 minutes) was determined. A comparison between MCF 7 and MDA MB 231 cells grown in the absence of estrogen revealed that the overall sensitivity of MCF 7 cells to nuclease treatment was increased (Figure 21). This was evident from the decreased hybridization signal of the parent Bgl II c-*myc* fragment when nuclei from MCF 7 cells grown under estrogen depleted conditions were treated with 2 units/ml DNase I as compared to nuclei isolated from MCF 7 cells grown under normal conditions and treated with 5 units/ml of DNase I (Figures 20 and 21). Nuclease digestion, rather than a loading error, appeared to be responsible for this difference, since there was increased hybridization to smaller sized c-*myc* fragments (Figure 21). This

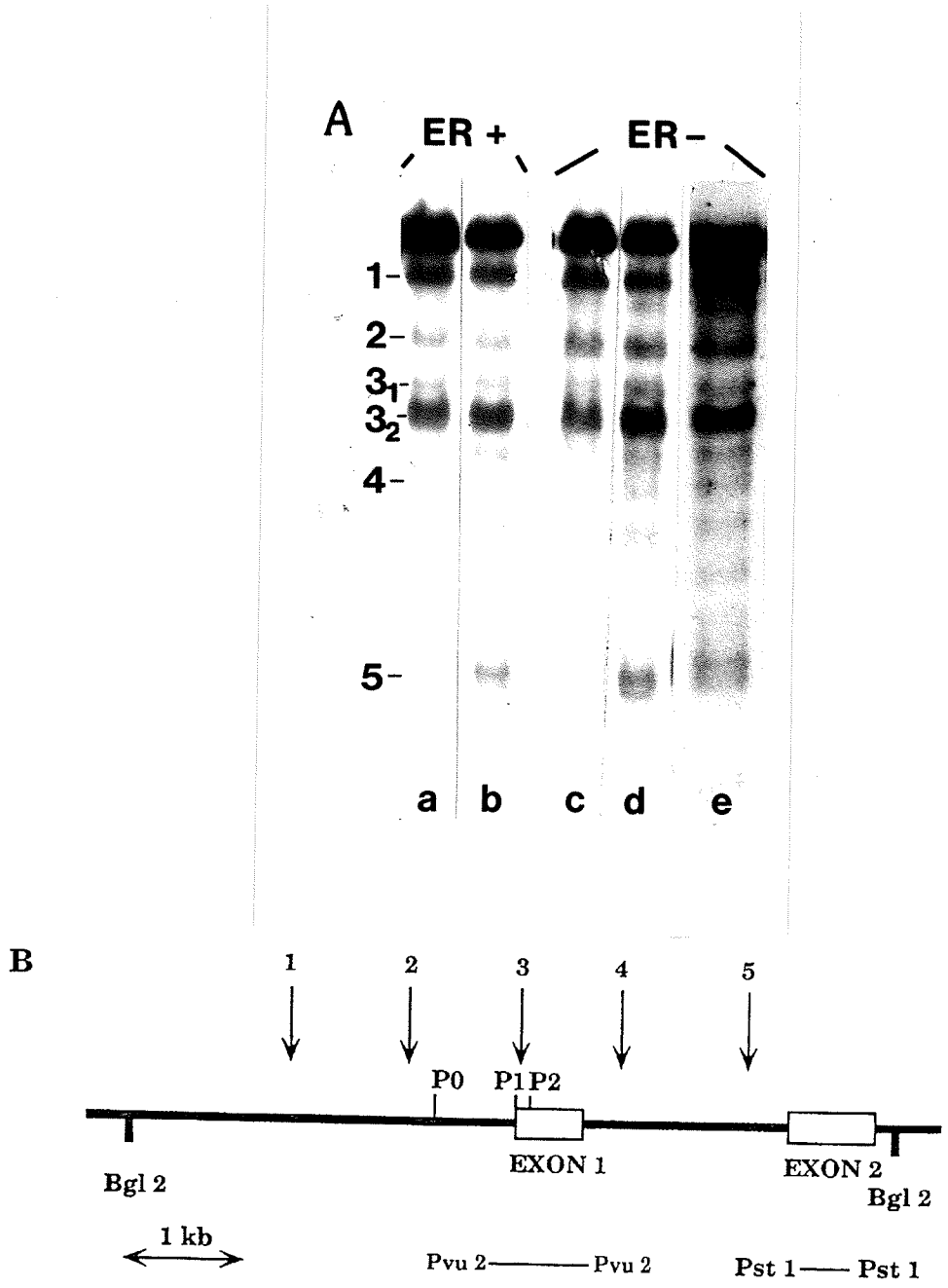


Figure 20. DNase I Hypersensitive Sites of c-*myc* Gene Chromatin in Other Breast Cancer Cell Lines. Nuclei isolated from ER^+ human breast cancer cell lines T47 D5 (5 units/ml; lane a) and MCF 7 (5 units/ml; lane b) and ER^- human breast cancer cell lines MDA 468 (3 and 5 units/ml DNase I; lanes c and d) and BT 20 (5 units/ml; lane e) were treated with DNase I and analyzed for c-*myc* DNase I hypersensitive sites **(A)**, as described in Methods and Materials. The DNase I hypersensitive sites are indicated at the left side of the figure. **(B)** A schematic diagram of the human c-*myc* gene, showing the positions of the promoters, the DNase I hypersensitive sites (as described in the legend to Figure 19) and the Pst 1 c-*myc* exon 2 fragment used in these studies. These studies were done with technical assistance from Tim Salo (Miller *et al.*, 1993).

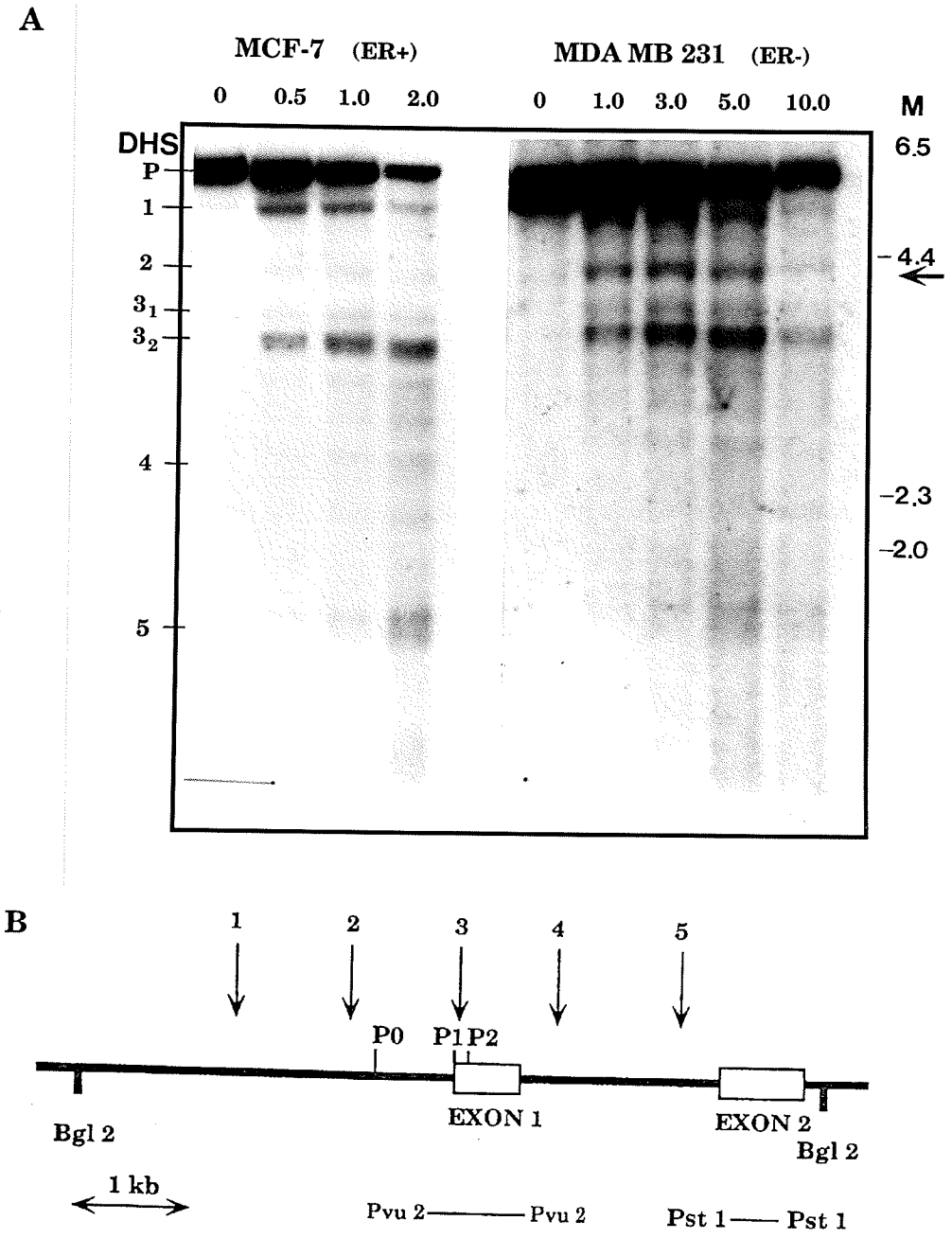


Figure 21. DNase I Hypersensitivity of c-*myc* in Breast Cancer Cells Grown Under Estrogen Free Conditions. Nuclei were isolated from ER^+ MCF 7 and ER^- MDA MB 231 breast cancer cell lines grown under estrogen depleted conditions and analyzed for DNase I hypersensitive sites as described in Methods and Materials (**A**). Units of DNase I are indicated above each lane and the DNase I hypersensitive sites are indicated (DHS). M refers to the DNA fragment length standards (lambda phage DNA digested with Hind III) in kilobase pairs (kbp). The arrow indicates DNase I hypersensitive site 2 or II_2. (**B**) A schematic diagram of the human c-*myc* gene, showing the positions of the promoters, the DNase I hypersensitive sites (as described in the legend to Figure 19) and the Pst 1 c-*myc* exon 2 fragment used in these studies. This study was done with technical assistance from Norm Huzel.

increased nuclease sensitivity was even more evident with 3 units/ml DNase I and was unaffected by estrogen addition (Figure 22). Conversely, the DNase I sensitivity of nuclei from MDA MB 231 cells appeared to be unchanged by these differences in growth conditions (Figures 19 and 21).

Regardless of this general increase in DNase I sensitivity, the position and relative intensity of the DH sites in the c-*myc* chromatin of MCF 7 cells maintained in phenol red free, charcoal stripped serum were essentially similar to cells maintained in normal medium, which contained estrogens and other growth factors (compare Figures 20 A and 21). In particular, the relative hypersensitivity of DH site II_2 in MCF 7 cells remained low compared to MDA MB 231 cells grown under the same conditions (Figure 21). A trend towards increased DNase I hypersensitivity associated with both DH II_2 and DH III_2 occurred with 17ß-estradiol stimulation (Figure 22), although the relative intensities of these DH sites remained the same.

Although the breast cancer cell lines MCF 7 and MDA MB 231 were both derived from human breast epithelial tumors, they were isolated from different women and therefore share no direct genetic relationship. The MCF 7 and MDA MB 231 cell lines are however, cellular models for the hormonally dependent and hormonally independent breast cancer phenotypes, respectively. These two cell lines therefore represent the extremes of breast cancer cell types with respect to hormonal status. Near the end of this study, new models aimed at altering the hormonal phenotype of these two cell lines became available. Continuous treatment with adriamycin produced the MCF 7/AdrR cell line, which no longer expresses the estrogen receptor (Vicker *et al.*, 1988). Conversely, transfection with an ER construct produced the S30 cell line, which is parentally derived from MDA MB 231 cells (Jiang and Jordan, 1992).

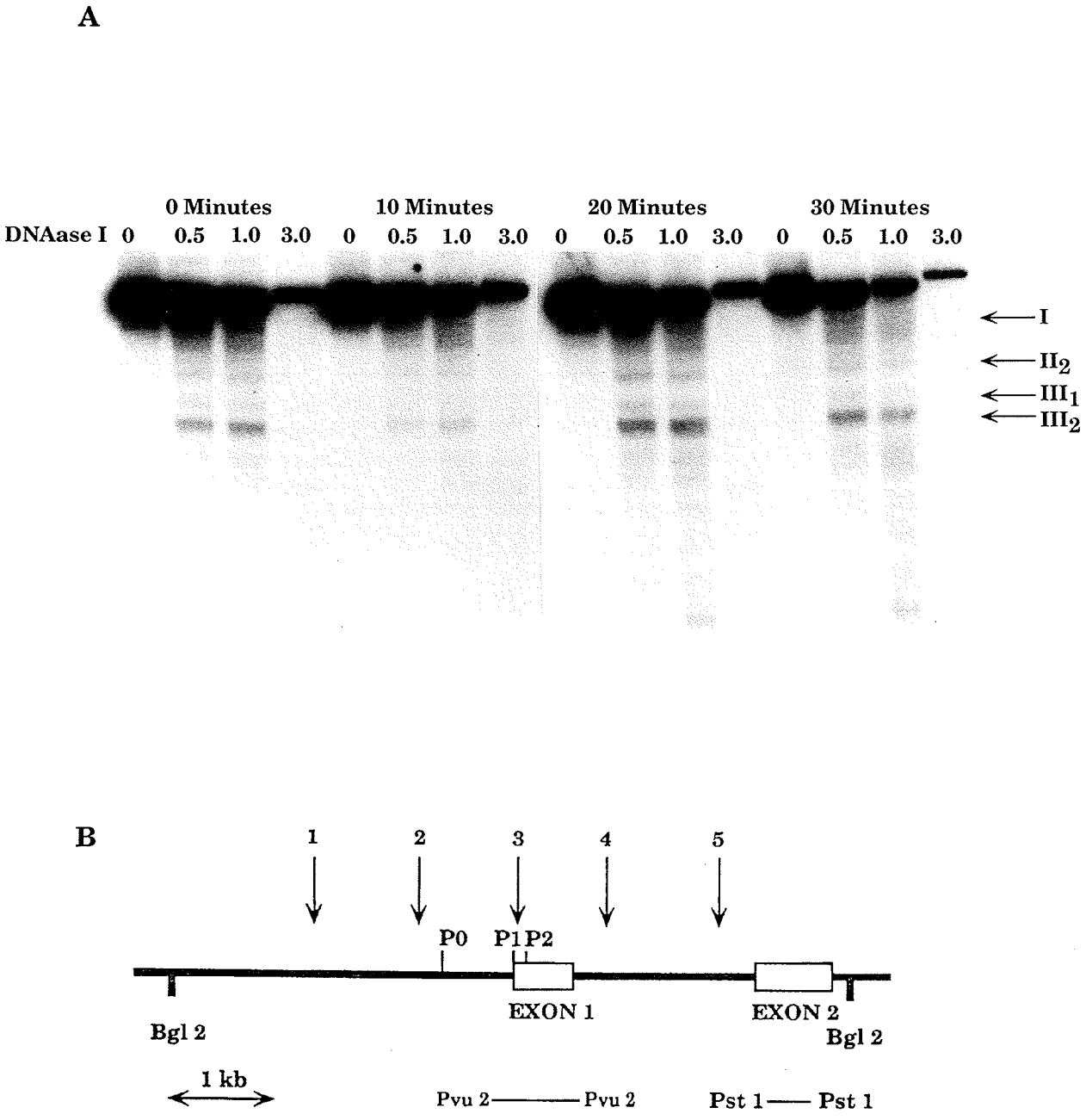


Figure 22. DNase I Hypersensitivity of c-*myc* in Estrogen Rescued MCF 7 Cells. Nuclei were isolated from MCF 7 breast cancer cells grown under estrogen depleted conditions then supplemented with 17 ß-estradiol for various indicated time periods as described in Methods and Materials **(A)**. The units of DNase I are indicated above each lane and arrows refer to DNase I hypersensitive sites I, II_2, III_1 and III_2. **(B)** A schematic diagram of the human c-*myc* gene, showing the positions of the promoters, the DNase I hypersensitive sites (as described in the legend to Figure 19) and the Pst 1 c-*myc* exon 2 fragment used in these studies. This study was done with technical assistance from Norm Huzel.

The endogenous PR gene and an ERE-*CAT* vector transiently transfected into the S30 cells responded normally to estrogen, however estrogen inhibited rather than stimulated cellular proliferation (Jiang and Jordan, 1992). This is consistent with other attempts to restore estrogen activated growth in ER^- cells (Kushner *et al.*, 1990; Maminta *et al.*, 1991) and suggests that the reintroduction of the ER does not completely restore the ER^+ phenotype. Conversely, the long term culture of MCF 7 cells in the presence of adriamycin resulted in their resistance to this drug and loss of ER expression (Vickers *et al.*, 1988). Therefore, in the MCF 7/AdrR cell line the development of adriamycin resistance coincides with the concomitant development of estrogen independence and tamoxifen resistance. A preliminary analysis of the c-*myc* chromatin of S30 cells (Figure 23) revealed that the relative accessibilities of DH site II_2 remained similar to the parental ER^- MDA MB 231 cells. Conversely, MCF 7/AdrR cells grown in the presence or absence of adriamycin have an altered chromatin structure. The DNase I hypersensitivity of the c-*myc* gene was increased at DH II_2 in MCF 7/AdrR as compared to the parental ER^+ MCF 7 cell line (Figure 23). These preliminary findings suggest that gradual progression to an ER^- phenotype can alter the chromatin structure of the c-*myc* gene, however stable transfection of the ER can not reconstitute an ER^+ c-*myc* chromatin structure.

Promoter Usage

The DH II_2 site maps about 100 bp upstream of the P_0 promoter of the c-*myc* gene (Bentley and Groudine, 1986b). It has therefore been suggested that hypersensitivity at DH site II_2 may be related to the function of the P_0 promoter (Bentley and Groudine, 1986a; 1986b). Additionally, differential

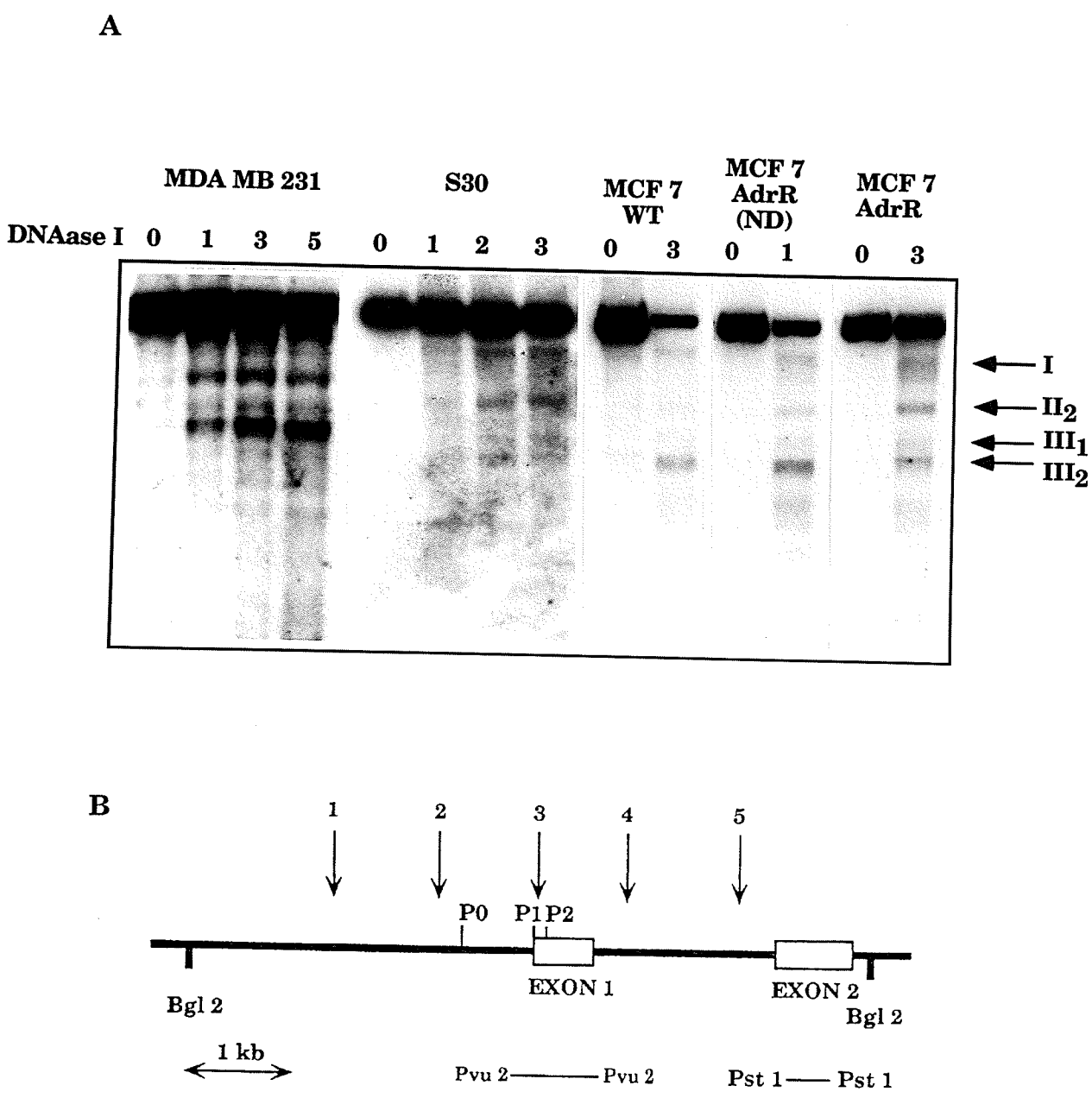


Figure 23. DNase I Hypersensitivity of c-*myc* in S30 and MCF 7 Adr R Cells. Nuclei were isolated from ER^+ S30 cells and ER^- MCF 7 AdrR cells grown in the presence and absence (ND) of adriamycin and treated with 0 to 3 units/ml of DNase I as described in Methods and Materials (A). The DNase I hypersensitivity of c-*myc* in these cells was compared to the parental cell lines MDA MB 231 and MCF 7 WT. DH sites I, II_2, III_1 and III_2 are indicated with arrows and the units of DNase I are listed above each lane. (B) A schematic diagram of the human c-*myc* gene, showing the positions of the promoters, the DNase I hypersensitive sites (as described in the legend to Figure 19) and the Pst 1 c-*myc* exon 2 fragment used in these studies. This study was done with technical assistance from Norm Huzel.

promoter usage has been previously observed in unrearranged and translocated c-*myc* sequences (Bentley and Groudine, 1986b; Broome *et al.*, 1987; Eick *et al.*, 1990). It was therefore possible that the differential sensitivity of DH site II_2 observed between ER^+ and ER^- breast cancer cells may have reflected differential promoter usage in the c-*myc* gene. The possibility that this was occurring at the P_0 promoter was investigated using an RNase protection assay.

RNA was isolated from several exponentially growing ER^+ and ER^- human breast cancer cell lines and then hybridized to antisense riboprobes complementary to the 784 bp Pvu II-Pvu II region (Figure 24 A) which began 3' to P_0 and spanned P_1 and P_2. Two major protected fragments of approximately 530 and 345 bp were seen in both ER^+ and ER^- cells (Figure 24 B). These data were consistent with both P_1 and P_2 being used in human breast cancer cell lines. Transcripts initiating from P_0 would be expected to hybridize to the entire Pvu II riboprobe, excluding any vector multicloning sequences remaining after linearization. However, protected fragments of this size were not detected in any cell line. The observation that in all cell lines the abundance of the 345 bp protected fragment was always higher than the 530 bp fragment, suggested that P_2 was generally the major promoter used. A possible exception was the ER^+ cell line ZR 75-1, where the intensity of each fragment was similar and suggested an equivalent use of P_1 and P_2.

Preliminary results also suggested that P_2 was the major promoter used for c-*myc* expression in both ER^+ and ER^- human breast cancer biopsy samples (data not shown). Significantly, preliminary studies indicated that estrogen treatment of ER^+ MCF 7 cells maintained under estrogen depleted conditions did not alter the relative ratio of P_1 and P_2 usage nor did it increase transcripts initiating from P_0 (data not shown). However, the

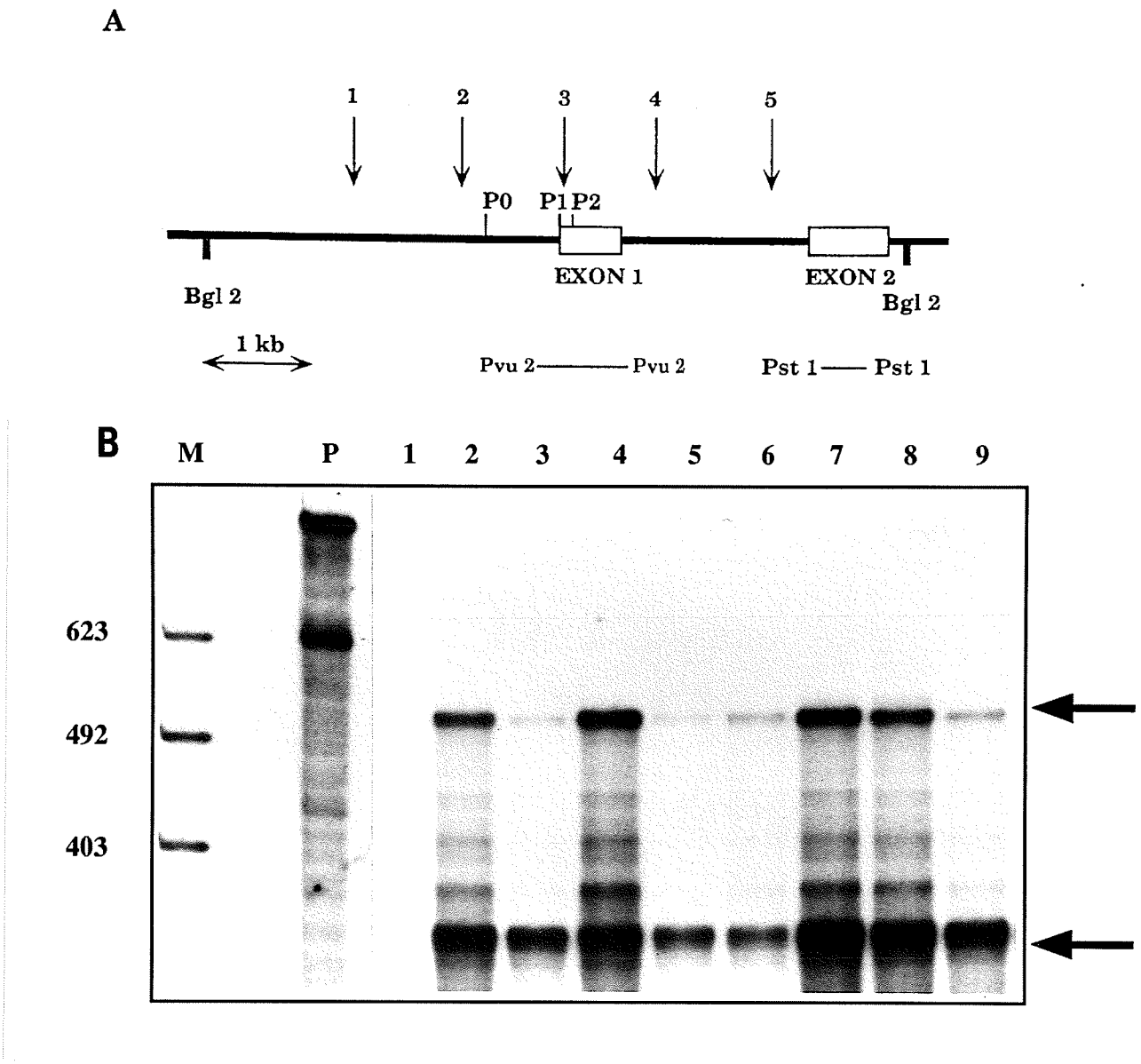


Figure 24. c-*myc* Promoter Usage in Estrogen Receptor Positive and Estrogen Receptor Negative Breast Cancer Cell Lines. An antisense riboprobe, illustrated in **(A)** as a c-*myc* exon 1 Pvu II-Pvu II fragment, was hybridized to yeast transfer RNA (lane 1) and RNA isolated from MDA MB 231 (lane 2), MDA 468 (lane 3), ZR 75-1 (lane 4), BT 20 (lane 5), T 47D (lane 6), T 47D 5 (lane 7), T 47D 5RP (lane 8), MCF 7 (lane 9) human breast cancer cells and subjected to RNase protection analysis as described in Methods and Materials. **(B)** The arrows indicate two major protected fragments of approximately 530 and 345 bp. P is the antisense riboprobe alone. The numbers to the left indicate the size in bp of the markers (M). This study was done with technical assistance from Helmut Dotzlaw (Miller *et al.*, 1993).

observed increase in P_2 protected fragments correlated with the estrogen stimulated mRNA accumulation reported earlier. Additionally, the data from the promoter usage study (Figure 24) also reflected the steady state mRNA levels of c-*myc* in the various human breast cancer cell lines (Figure 18), confirming the earlier observation that there was no apparent relationship between the level of c-*myc* mRNA and ER status.

Amplification

Translocation of the c-*myc* gene has been closely associated with Burkitt's lymphoma (reviewed in Cole, 1986; DePinho *et al.*, 1991; Spencer and Groudine, 1991; Marcu *et al.*, 1992). Although c-*myc* rearrangements have occasionally been observed in breast tumors (Escot *et al.*, 1986; Münzel *et al.*, 1991), rearrangements of the c-*myc* gene have not been detected in any of the breast cancer cell lines used in the above DNase I hypersensitivity studies (Dubik *et al*, 1987). Amplification of c-*myc* has been observed in a number of neoplasms (reviewed in Cole, 1986; DePinho *et al.*, 1991; Spencer and Groudine, 1991; Marcu *et al.*, 1992) including breast carcinomas (Escot *et al.*, 1986; Mariani-Constantini *et al.*, 1988; Münzel *et al.*, 1991; Berns *et al.*, 1992; Pavelic *et al.*, 1992). Significantly, in the non-tumorigenic cell line, HBL 100, 5 fold c-*myc* amplification has been observed and found to be responsible for the enhanced constitutive level of c-*myc* expression (Dubik *et al*, 1987). Similarly in this study it was determined that the c-*myc* gene was amplified 2 fold in the cell lines BT 20 and MCF 7 compared to normal human DNA and the other breast cancer cell lines analyzed.

The observation that amplification occurred in both ER^+ MCF 7 cells and ER^- BT 20 cells suggested that amplification of c-*myc* would not be responsible for the increased sensitivity of DH site II_2 observed in MDA MB 231 cells relative to MCF 7. To confirm this suggestion and since

human breast cancer cell lines in culture are sometimes genetically unstable (Reddel *et al.*, 1988), the ER^+ and ER^- breast cancer cell lines used for the DNase I hypersensitivity studies were analyzed for amplification of the c-*myc* gene (Figure 25). The c-*myc* locus is found on the long arm of human chromosome 8, at 8q24. To determine if this region of chromosome 8 was amplified, size separated DNA was hybridized to a c-*myc* exon 2 probe and compared to the intensity of hybridization to a thyroglobulin probe and a PRA/calcyclin probe. The thyroglobulin gene is also located at 8q24 in the human genome (Baas *et al.*, 1985), while the PRA/calcyclin gene has been localized to chromosome 1 (Murphy *et al.*, 1988). Amplification of the c-*myc* gene in MDA MB 231 cells was not observed, however, the previously observed 2 fold amplification of c-*myc* in BT 20 and MCF 7 cells was confirmed. Additionally, approximately 5 fold amplification of c-*myc* in T47 D5 cells was found, confirming another previous study (Wong and Murphy, 1991).

Endogenous Methylation

Proteins interacting with DNA form bonds with individual bases of a target sequence, hence methylation has the potential to disrupt sequence specific proteins from binding. In particular, the c-Myc protein itself was inhibited from binding DNA when its binding site was methylated (Prendergast and Ziff, 1991). Since DNA methylation at CpG islands has been implicated in the regulation of transcription (Bird, 1992), differences in the methylation pattern of c-*myc* DNA could explain differential regulation, by altering the accessibility of a transcription factor binding site in the 5' flanking region of the gene. G+C-rich or methylation free islands, MFIs, of active housekeeping genes were generally not methylated *in vivo* (Bianchi *et*

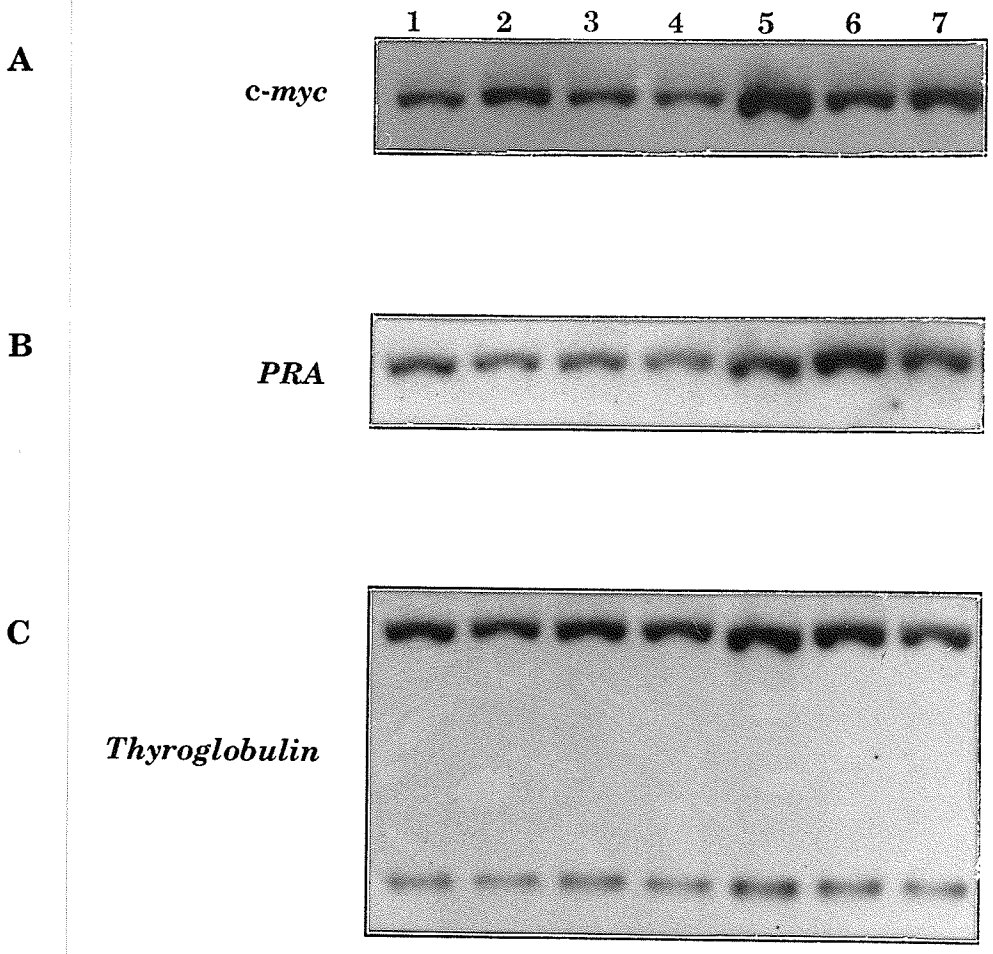


Figure 25. Comparison of Breast Cancer Cell Lines for c-*myc* Amplification. DNA was isolated from a variety of breast cancer cell lines and digested with the restriction enzyme Bgl II. Twelve μg of normal human DNA (lane 1) and DNA from ER^- BT 20 (lane 2), MDA 468 (lane 3), MDA MB 231 (lane 4) and ER^+ T 47 D5 (lane 5), ZR 75-1 (lane 6) MCF 7 (lane 7) cell lines were size separated and subjected to DNA analysis as described in Methods and Materials. A comparison of the hybridization signal to **(A)** a c-*myc* exon 2 cDNA, **(B)** a PRA or calcyclin probe (Murphy *et al.*,1988) and **(C)** a thyroglobulin probe (Baas *et al.*, 1985) was used to determine if the c-*myc* gene was amplified. This study was done with technical assistance from Norm Huzel.

al., 1989). An exception was the truncated c-*myc* gene which, though expressed, had a methylated G+C-rich island covering exons 1 and 2.

Differential methylation of the c-*myc* 5' flanking sequences of ER^+ and ER^- human breast cancer cell lines could be responsible for differential hypersensitivity at DH site II_2 if the methylation altered an important protein binding site or the general structure of the c-*myc* chromatin. An analysis of the nucleotide sequences from -2119 to -409, relative to P_1, (Figure 26) of the c-*myc* 5' flanking region using methylation sensitive restriction endonucleases was therefore undertaken. Genomic DNA from ER^+ MCF 7 and ER^- MDA MB 231 cell lines was isolated, digested with Pst I to generate the -2119 to -409 region, and then further digested with the methylation sensitive restriction endonucleases; Hpa II, Msp I and Hha I. Hpa II and Msp I both recognize the sequence CCGG, but Hpa II will not restrict the site if it is CmCGG, while Msp I will restrict CmCGG, and neither will restrict mCmCGG (Sambrook *et al.*, 1989). Similarly, Hha I recognizes the sequence GCGC, but will not restrict GmCGC or GCGmC (Sambrook *et al.*, 1989). Thedigested samples were analyzed after size separation and transfer to nitrocellulose by hybridization with either probe 1 (Pst I-Taq I c-*myc* fragment) that abuts the Pst I site at -2119 or probe 2 (Pst I-Apa I c-*myc* fragment) that abuts the Pst I site at -409 (Figure 26 A).

Digestion with Pst 1 alone revealed a 1710 bp c-*myc* fragment in both cell lines which was detected with either probe 1 or 2 (Figure 26 B and C, lanes 1 and 2). Using probe 1 a 937 bp fragment in both cell lines was detected after Pst I and Msp I digestion (Figure 26 B, lanes 5 and 6). The same size fragment was also detected after Pst I and Hpa II digestion (Figure 26 B, lanes 3 and 4), suggesting that the Hpa II-Msp I restriction site at -1182 was unmethylated. All other Hpa II-Msp I restriction sites 3' of -1182

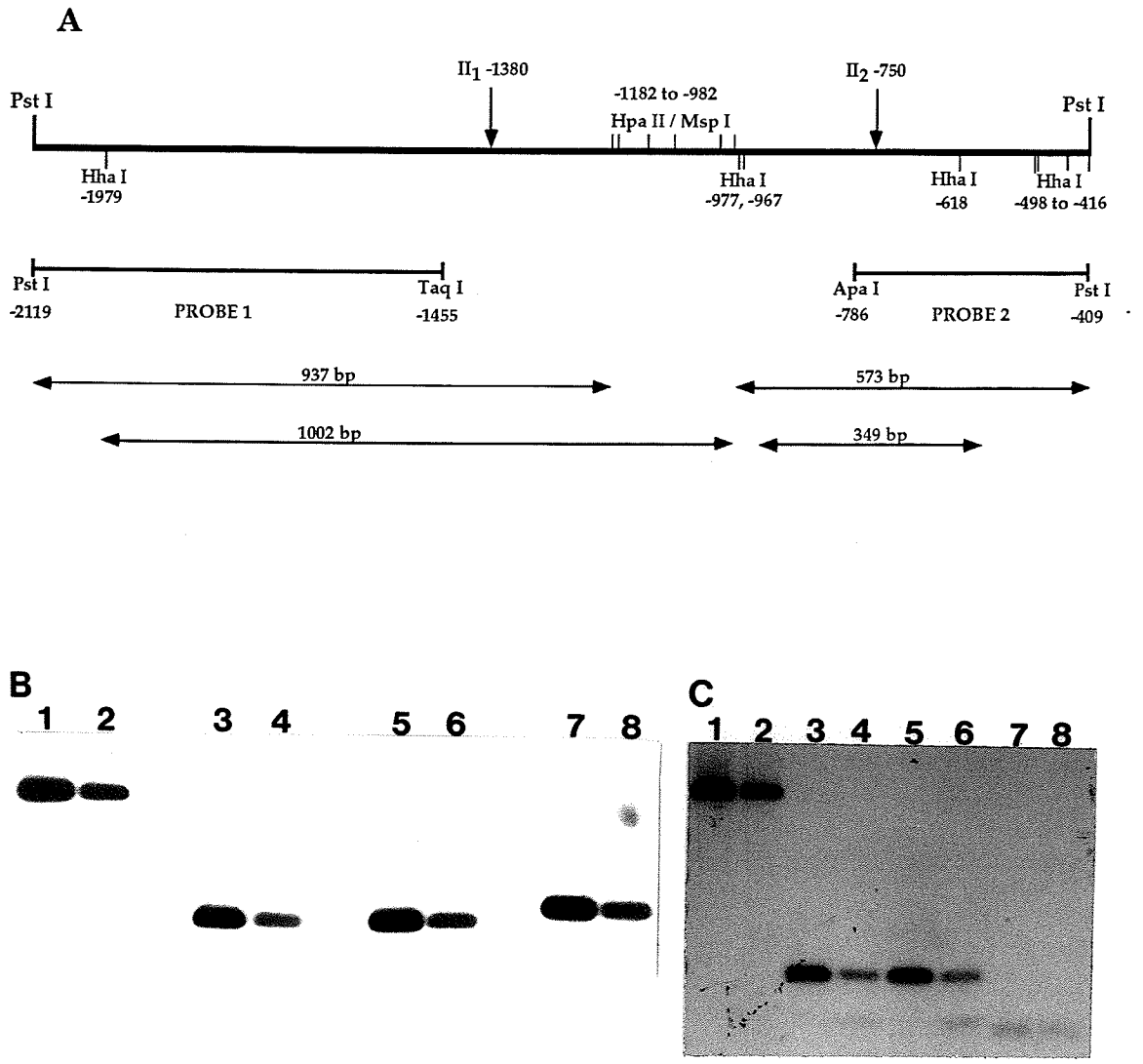


Figure 26. Methylation Status of the c-*myc* 5' Flanking Sequences in MFC 7 and MDA MB 231 Cells. **(A)** Diagram illustrating the two probes used for the DNA analysis and the expected c-*myc* gene chromatin restriction sites in the region -2119 to -409 relative to P_1. For both **(B)** and **(C)** the genomic DNA was digested with Pst I (lanes 1 and 2), Pst I and Hpa II (lanes 3 and 4), Pst I and Msp I (lanes 5 and 6) and Pst I and Hha I (lanes 7 and 8). **(B)** Size separated DNA fragments from MCF 7 (lanes 1,3,5,7) and MDA MB 231 (lanes 2,4,6,8) cells were hybridized to the Pst I-Taq I 664 bp probe. **(C)** The same DNA samples were hybridized to the Pst I-Apa I 377 bp probe. This study was done in collaboration with Dr. Yan Jin.

were thus rendered undetectable with probe 1. Using probe 2, a 573 bp fragment was detected in both Hpa II and Msp I digested DNA from both cell lines (Figure 26 C, lanes 3 to 6). These data suggested that the Hpa II-Msp I site at -982 was also unmethylated in both breast cancer cell lines. Further analysis of the methylation status of the Hpa II-Msp I sites between -1182 and -982 was therefore not possible with either probe.

Pst I and Hha I digested DNA generated a 1002 bp fragment that was detected with probe 1 (Figure 26 B, lanes 7 and 8). This suggested that Hha I restriction sites at -1979 and -977 were unmethylated in both cell lines. Therefore, Hha I restriction sites 3' to -977 were not detectable with probe 1. A 349 bp fragment was detected using probe 2 in Pst I and Hha I digested DNA of both cell lines (Figure 26 C, lanes 7 and 8). These data suggested that Hha I sites at -967 and -618 were unmethylated in the two cell lines The restriction fragments that would be generated if the Hha I sites 3' of -618 were unmethylated, would range in size from 120 to 2 bp. These fragments would be undetectable by this analysis due to their size and therefore their methylation status remained unresolved. In summary the data suggested this region of the c-*myc* gene was undermethylated in both MCF 7 ER+ and MDA MB 231 ER- human breast cancer cell lines.

Protein Associations with DH site II_2

No association of promoter usage, amplification or differential methylation with the differential DH of the chromatin structure surrounding the human c-*myc* DH II_2 site in breast cancer cell lines was apparent from the previously described studies. However, a significant body of research suggested that hypersensitivity to nucleases was indicative of displacement or disruption of the nucleosomal packaging of DNA by the association of *trans*-acting factors with *cis*-acting elements located near DH sites (reviewed

in Gross and Garrard, 1988; Bonifer *et al.*, 1991; Felsenfeld, 1992; Adams and Workman, 1993). Therefore, an alternative explanation for the differential nuclease activity at c-*myc* DH site II_2 in ER^+ and ER^- breast cancer cells was that *trans*-acting factors were differentially associated with this regulatory region in these two cell types. To investigate this possibility the electrophoretic mobility shift assay (EMSA) combined with DNase I footprinting were used to characterize proteins that could interact *in vitro* with the region spanning the DH II_2 site.

Initial studies used a Hae III-Acc I DNA fragment of the c-*myc* gene which encompassed the DH II_2 site from -787 to -607, relative to P_1 (Figure 27). Nuclear extracts from MCF 7 ER^+ and MDA MB 231 ER^- cell lines were incubated with the end labelled fragment and subjected to non-denaturing gel electrophoresis (Figure 28). Several DNA-protein complexes were seen when crude nuclear extracts from either cell line were incubated with the end-labelled Hae III-Acc I 203 bp fragment. The spectrum of the complexes formed were similar, but differences were noted (Figure 28, see triangles). Additionally, the possibility that there were cooperative interactions between the various proteins interacting with this region of c-*myc* was suggested by the inability of increased nuclear extract added to the reaction to increase the intensity of the individual DNA-protein complexes which differed between the two extracts.

The Stairway assay (van Wijnen *et al.*, 1992) was used to further delineate the potential protein-DNA interactions in the Hae III-Acc I region. This fragment was end-labelled independently at either end and shortened at internal restriction sites (Figure 29 A). The results of this assay demonstrated that proteins were binding to sites located along the entire length of the Hae III-Acc I fragment, accounting for the complicated nature of

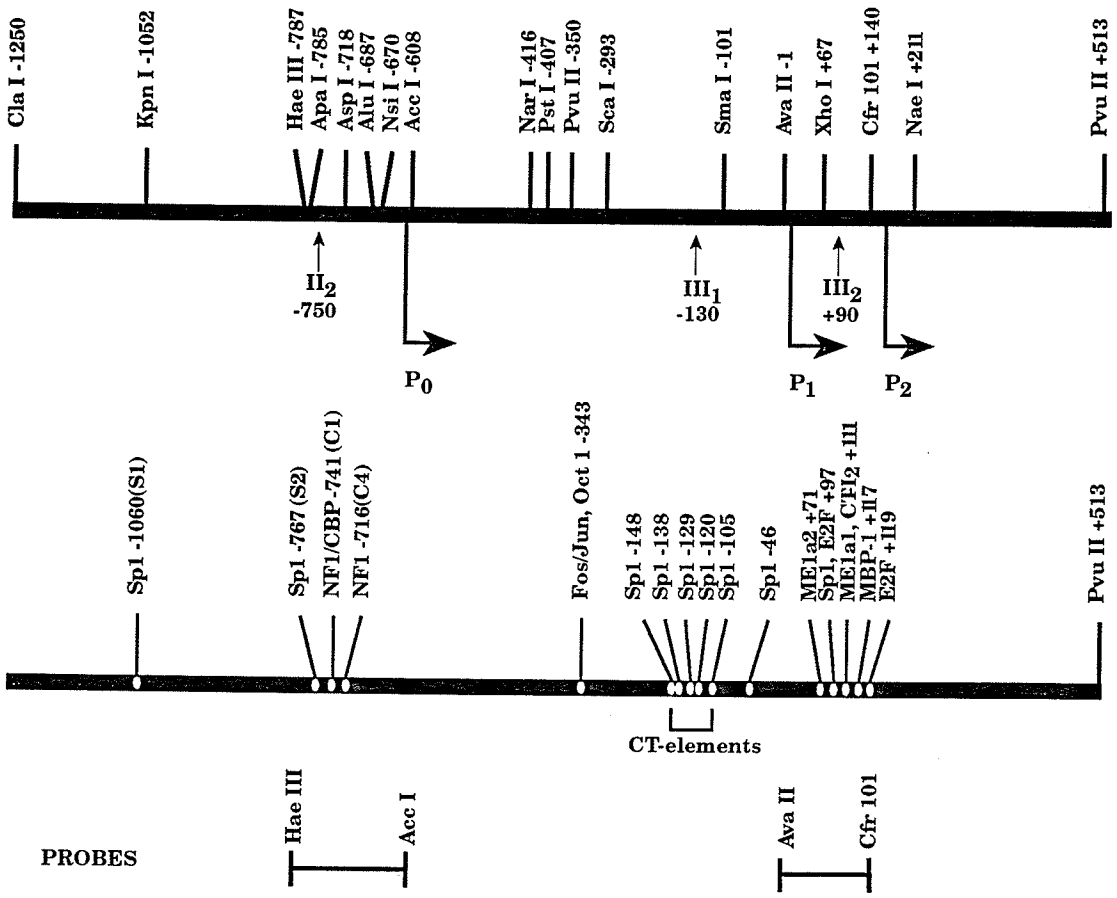


Figure 27. Summary of *in situ* DNAase I Hypersensitive Sites and Relevant *in vitro* *Trans*-acting Factor Binding Sites in the 5' Flanking Sequences of c-*myc*. A number of *trans*-acting factors have been found to associate with regulatory regions of the c-*myc* gene, including the DH II_2 and DH III_2 regions. In particular, Sp1, NF1 and NF1/CBP binding sites have been identified in the DH II_2 region, while sites for ME1a2, E2F, ME1a1 and MBP-1 have been identified in the DH III_2 region. These sites are located within the 5' portion of a c-*myc* Hae III-Acc I fragment and the 3' portion of a c-*myc* Ava II-Cfr101 fragment used as probes in studies of these regulatory regions. The topological arrangement of these sites in relationship to three of the c-*myc* promoters, their associated DH sites and restriction sites in the region is shown. References for these binding sites can be found in the text. Please refer to Figure 15 for additional *cis*-acting elements.

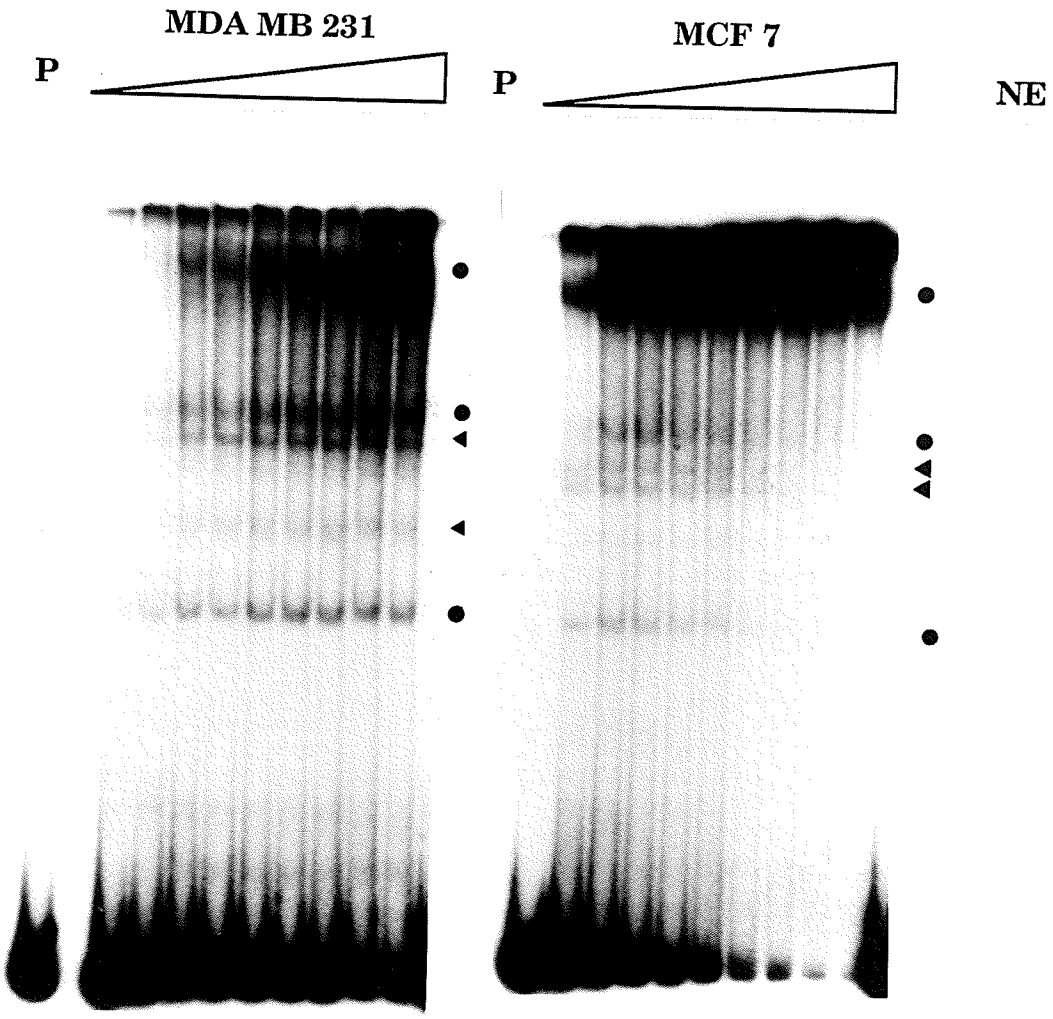


Figure 28. Interaction of DNA Binding Proteins with the c-*myc* DNase I Hypersensitive II_2 Region. Electrophoretic mobility shift assay of the DH II_2 c-*myc* DNA is shown. Increasing amounts of MDA MB 231 and MCF 7 breast cancer cell line nuclear extracts (2-10 μg) were incubated with 1 ng of the end labelled Hae III-Acc I c-*myc* fragment as described in Methods and Materials. The major protein-DNA complexes formed are indicated with circles and arrows. The arrows indicate differences in complex migration between the two extracts.

the mobility shift patterns generated. Similar patterns of complex formation were obtained irrespective of which end of the 203 bp fragment was labelled (Figure 29 B compare lanes 2, 3 and 14, 15).

The full length 203 bp fragment and the 140 bp Hae III-Nsi I fragment produced at least seven DNA-protein complexes, however only 3 to 4 complexes remained associated with the 90 bp Hae III-Asp I fragment (Figure 29 B, lanes 11, 12). Complexes 1, 2 and 4 disappeared as the Hae III-Nsi I fragment was shortened 17 bp by Alu I digestion to a 122 bp fragment (Figure 29 B, compare lanes 5, 6 with 8, 9). The binding sites for complexes 3, 5, 6 and 7, detected with probes labelled at the Hae III end, were therefore mapped between the Hae III site at -787 and the Asp I site at -718, while DNA-protein complexes 1, 2 and 4 were found to interact between the Alu I site at -687 and the Nsi I site at -670. The differences in the pattern of DNA-protein complexes were not enhanced between MCF 7 and MDA MB 231 nuclear extracts when the Hae III labelled c-*myc* fragment was shortened by restriction endonuclease digestion.

At least seven DNA-protein complexes were also detected when the 203 bp c-*myc* fragment and the 113 bp Asp I-Acc I shortened fragment were labelled at the Acc I end (Figure 29 B, lanes 14, 15 and 17, 18). Some of the complexes formed with the 113 bp Asp I-Acc I fragment were also generated with the 63 bp Nsi I-Acc I region (Figure 29 B, lanes 23, 24). Nuclear proteins interacting with this region of DH site II_2 have not been previously reported. Interestingly, as the c-*myc* fragment was shortened from the 81 bp Alu I-Acc I fragment to the 63 bp Nsi I-Acc I fragment, new DNA-protein complexes appeared (Figure 29 B, compare lanes 20, 21 with 23, 24). Of particular interest, however, was the observation that nuclear extracts from

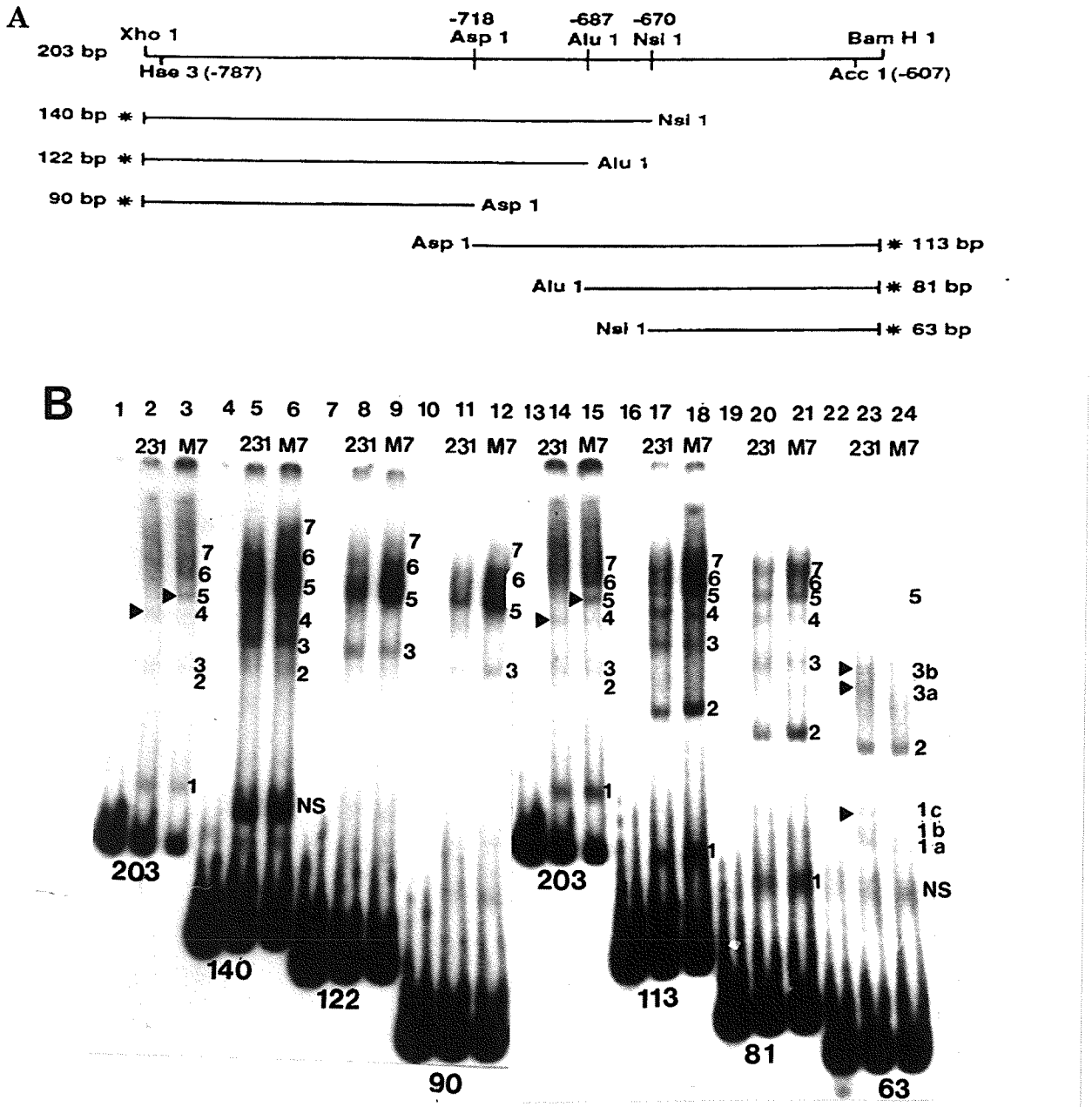


Figure 29. Stairway Assay of DNA Sequences Surrounding the c-*myc* DNase I Hypersensitive II_2 Region. The Hae III-Acc I DNA fragment was end labelled at either the Xho I or Bam HI end and then shortened using internal restriction sites as described in Methods and Materials. **(A)** The expected labelled fragments generated from the restriction endonuclease digestions are diagrammed. **(B)** The protein-DNA complexes generated with the labelled fragments and 2-3 μg of MDA MB 231 (lanes 2, 5, 8, 11, 14, 17, 20, 23) and MCF 7 (lanes 3, 6, 9, 12, 15, 18, 21, 24) nuclear extracts were resolved by electrophoresis. Complexes were labelled according to their migration and numbers do not reflect similarities between complexes. The lanes preceding each pair of reactions show the labelled probes without added proteins. The size of each probe is listed under each set of reactions and differences are noted with arrows.

MDA MB 231 and MCF-7 breast cancer cells formed several complexes with differing intensities (Figure 29 B, compare lanes 23 and 24).

Association of Sp1 and NF1 with DH Site II2

Nuclear extracts from HeLa cells have previously been used to demonstrate *in vitro* sequence specific DNA binding of Sp1, NF-1 and CBP to the DH II2 region (Lang *et al.*, 1991). These *trans*-acting factors were therefore candidates for the complexes formed with the 122 bp Hae III-Alu I DH II2 region, which contained the appropriate *cis*-acting elements for these proteins. The S1, S2, C1, and C4 oligonucleotides (Figure 27), which corresponded to the two previously described Sp1 sites plus previously described NF1-CBP and NF1 binding sites, respectively (Lang *et al.*, 1991) were synthesized for *in vitro* DNA binding studies. They were used in EMSA as probes and as competitors of complexes formed between the 140 bp Hae III-Nsi I fragment and nuclear extracts from breast cancer cells (Figure 30). A similar pattern of protein-DNA complexes was formed between the MCF 7 and MDA MB 231 nuclear extracts with these oligonucleotides. The pattern of complexes formed with the S2 oligonucleotide was similar to that seen with a high affinity Sp1 oligonucleotide (Figure 30 A, compare lanes 2, 3 with 5, 6). One difference in the EMSA patterns with S2 or Sp1 oligonucleotides with MCF 7 or MDA MB 231 human breast cancer nuclear extracts was in the abundance and/or activity of complex 1 which was typically greater with the MDA MB 231 nuclear extracts (Figure 30 A). The pattern obtained with the Sp1 oligonucleotide and either breast cancer cell nuclear extracts or chicken immature erythrocyte nuclear extracts was essentially identical (Figure 30 A, lane 7). Evidence has previously been provided for the interaction of Sp1 multimers from the chicken erythrocyte nuclear extract with the Sp1 oligonucleotide (Sun *et al.*, 1992). The S2 oligonucleotide-protein complexes

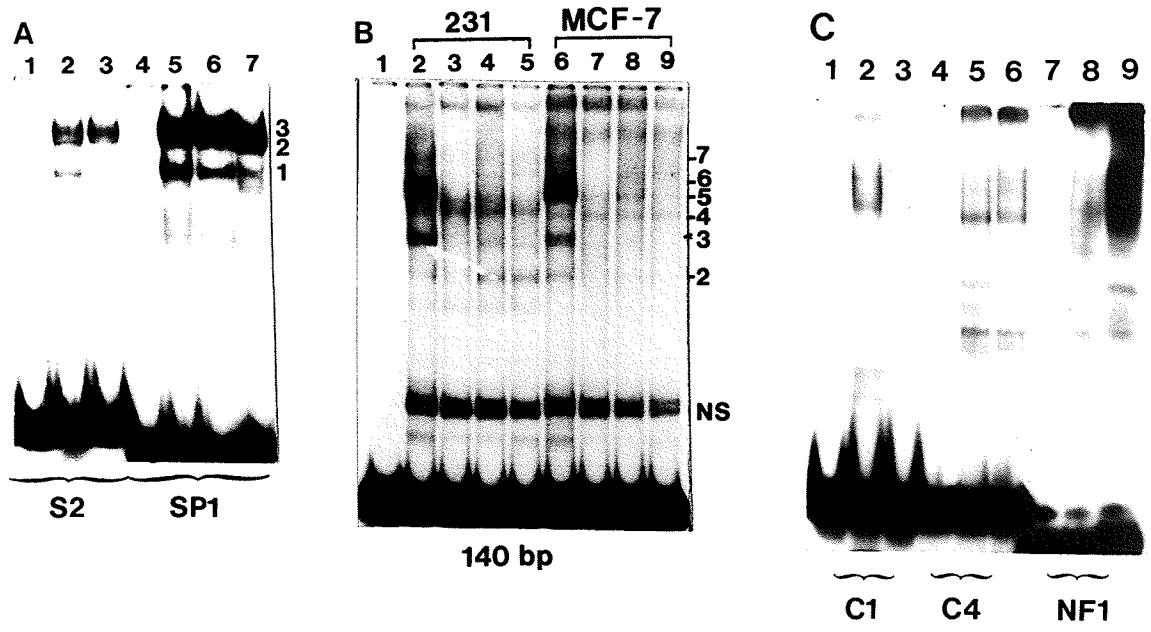


Figure 30. Sp1 and NF1 Interactions with Characterized Sites of the DNase I Hypersensitive II_2 Region. The oligonucleotides described by Lang *et al.*, (1991) were used in EMSA as described in Methods and Materials to determine if there were proteins available in nuclear extracts from breast cancer cell lines which could interact with these sequences. **(A)** The interaction of 2-3 μg of nuclear extracts from MDA MB 231 (lanes 2, 5), MCF 7 (lanes 3, 6) and chicken erythrocyte (lane 7) with 1 ng of the S2 (lanes 1-3) and Sp1 consensus (lanes 4-7) oligonucleotides. **(B)** The Hae III-Nsi I region of c-*myc* which contains the S2 consensus sequence for Sp1 was end labelled and used in EMSA as a probe for DNA-protein complexes from MDA MB 231 (lanes 1-5) and MCF 7 (lanes 6-9) nuclear extracts. A 100 fold molar excess of S1 (lanes 3 and 7), S2 (lanes 4 and 8), and Sp1 consensus (lanes 5 and 9) oligonucleotides were added as specific competitors of DNA-protein complexes. **(C)** The interaction of 2-3 μg of MDA MB 231 (lanes 2, 5, 8) and MCF 7 (lanes 3, 6, 9) nuclear extracts with 1 ng of the C1 (lanes 1-3), C4 (lanes 4-6) and NF1 consensus (lanes 7-9) oligonucleotides.

appeared to co-migrate with complexes 3, 5 and 6 formed with the 90 bp, 122 bp and 140 bp c-*myc* DNA fragments (compare Figures 29 and 30 A). Competition experiments (Figure 30 B) provided evidence consistent with the hypothesis that complexes 3, 5 and 6 were generated by Sp1-like multimers binding to the S2 Sp1 binding site. Complexes 3, 5, 6, and 7 formed with the 140 bp Hae III-Nsi I fragment (Figure 30 B) were eliminated by the inclusion of a 100 fold molar excess of unlabelled oligonucleotides containing Sp1 binding sites S1, S2 or Sp1.

Protein complexes were formed on the C1 and C4 oligonucleotides. Comparison with the consensus NF1 oligonucleotide identified some similarities and differences (Figure 30 C). The pattern of C4 complexes more closely resembled the pattern of complexes formed with the consensus NF1 oligonucleotide (Figure 30 C, compare lanes 5, 6 with 8, 9). Significantly, the C4 fragment represented two NF1 sites, while the C1 sequence has been described as having both NF1 and CBP binding sites (Lang *et al.*, 1991). There appeared to be quantitative differences in some of the DNA-protein complexes formed with the C1 and C4 oligonucleotides between the two nuclear extracts (Figure 30 C, compare lanes 2, 3 and 5, 6).

The NF1 and CBP binding sites associated with the C1 and C4 oligonucleotides are located between the Hae III site at -787 and the Alu I site at -687 (see Figure 27 or 29). The Asp I restriction site lies between the C1 and C4 sequences, and endonuclease digestion at Asp I would delete the NF1 binding sites associated with the C4 sequence from c-*myc* fragments labelled at the Hae III end (Figure 29). The NF1-CBP and NF1 sites previously described could therefore be involved in DNA-protein complexes formed within the Hae III-Alu I c-*myc* region. However, oligonucleotides for these sites were unable to compete for complexes formed on the larger Hae

III-Nsi I c-*myc* fragment (data not shown). Therefore, while binding sites for NF1 and CBP had previously been mapped in the Hae III-Alu I region (Lang *et al.*, 1991), complexes formed in this region using human breast cancer cell nuclear extracts did not appear to be due to NF1 or CBP interactions. Sp1-like proteins were already determined to be most likely responsible for complexes 3, 5, 6, and 7 formed within this region (Figure 30 B), however Sp1, NF1 or CBP were unlikely to account for complexes 2 and 4, which mapped within the Alu I-Nsi I region of the 140 bp fragment. Significantly, the *cis*-acting sites for Sp1, NF1 and CBP also did not correspond to protein binding sites for the additional protein interactions with the 63 bp Nsi I-Acc I region located between -670 and -607 relative to P_1.

Novel DNA-binding Sites at c-*myc* DH Site II_2

Consistent with previous DNase I footprinting studies (Lang *et al.*, 1991), protection of the Sp1, NF1-CBP and NF1 binding sites located in the Hae III-Acc I c-*myc* DNA was observed with breast cancer cell nuclear extracts (Figure 31). However, several other protected regions were located 3' to the NF1 binding sites in both cell lines (Figure 32). Protected region I was located between the Nsi I and Asp I sites and contained the Alu I site, while protected regions II to IV were found to interact with sequences between the Nsi I and Acc I sites. These protected regions are consistent with and most likely correspond to novel protein-DNA complexes identified in the EMSA described above (Figure 29). The protected regions generated with the MCF 7 and MDA MB 231 nuclear extracts were similar for the sense strand (Figure 32 A), but differences in the footprints over the antisense strand (Figure 32 B) were noted. With the antisense strand three *in vitro* DNase I hypersensitive sites were detected (Figure 32 B, see arrows). The first hypersensitive site mapped between footprints III and IV, while the other two

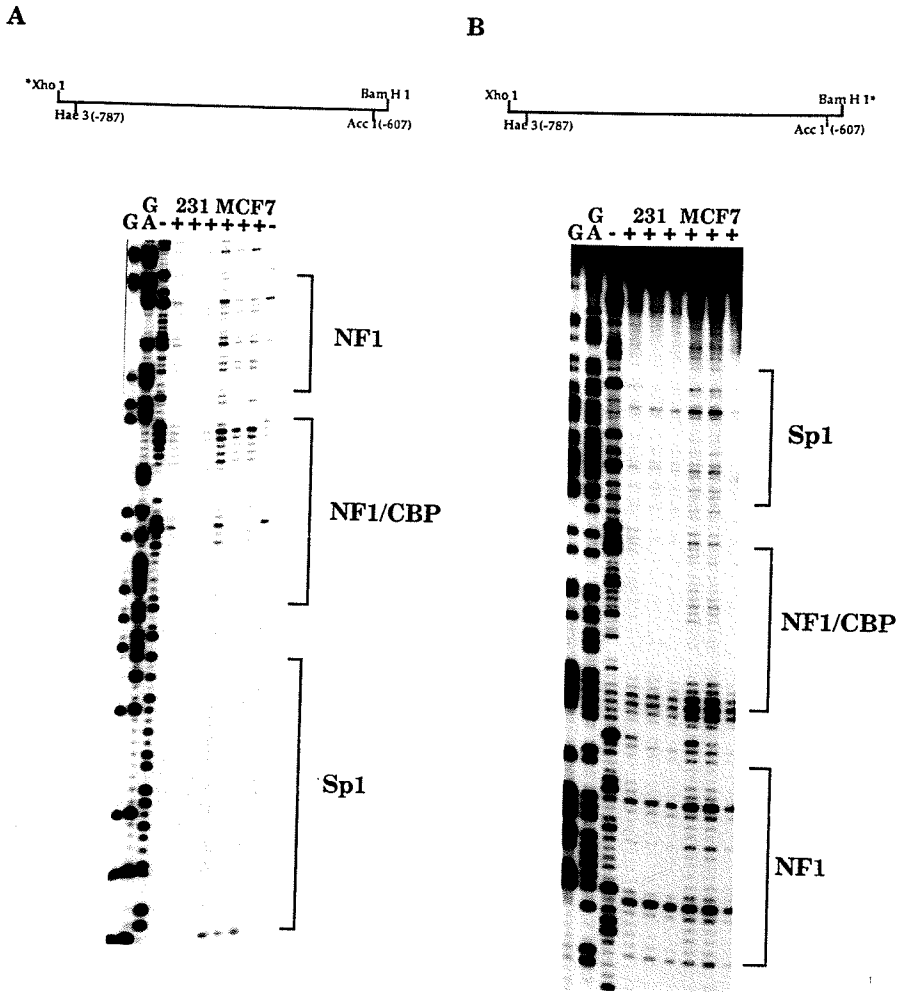


Figure 31. Footprinting Analysis of the Sp1 and NF1 c-*myc* Binding Sites. DNase 1 footprint analysis of the Hae III-Acc I c-*myc* region was performed as described in Methods and Materials. The G and G+A reactions were run in the first two lanes of each panel. The control DNA without any added nuclear extract was run on either side of the experimental footprint reactions (-). Three lanes each of the MDA MB 231 and MCF 7 reactions representing 2 ng of DNA, 50 µg of extract and DNase I concentrations of 0.2, 0.1 and 0.01 units per 40 µl reaction were run in each panel (+). The Hae III-Acc I fragment was labelled with Klenow at the Hae III end (antisense strand) (**A**) or the Acc I end (sense strand) (**B**) using convenient sites in the multicloning region of the plasmid.

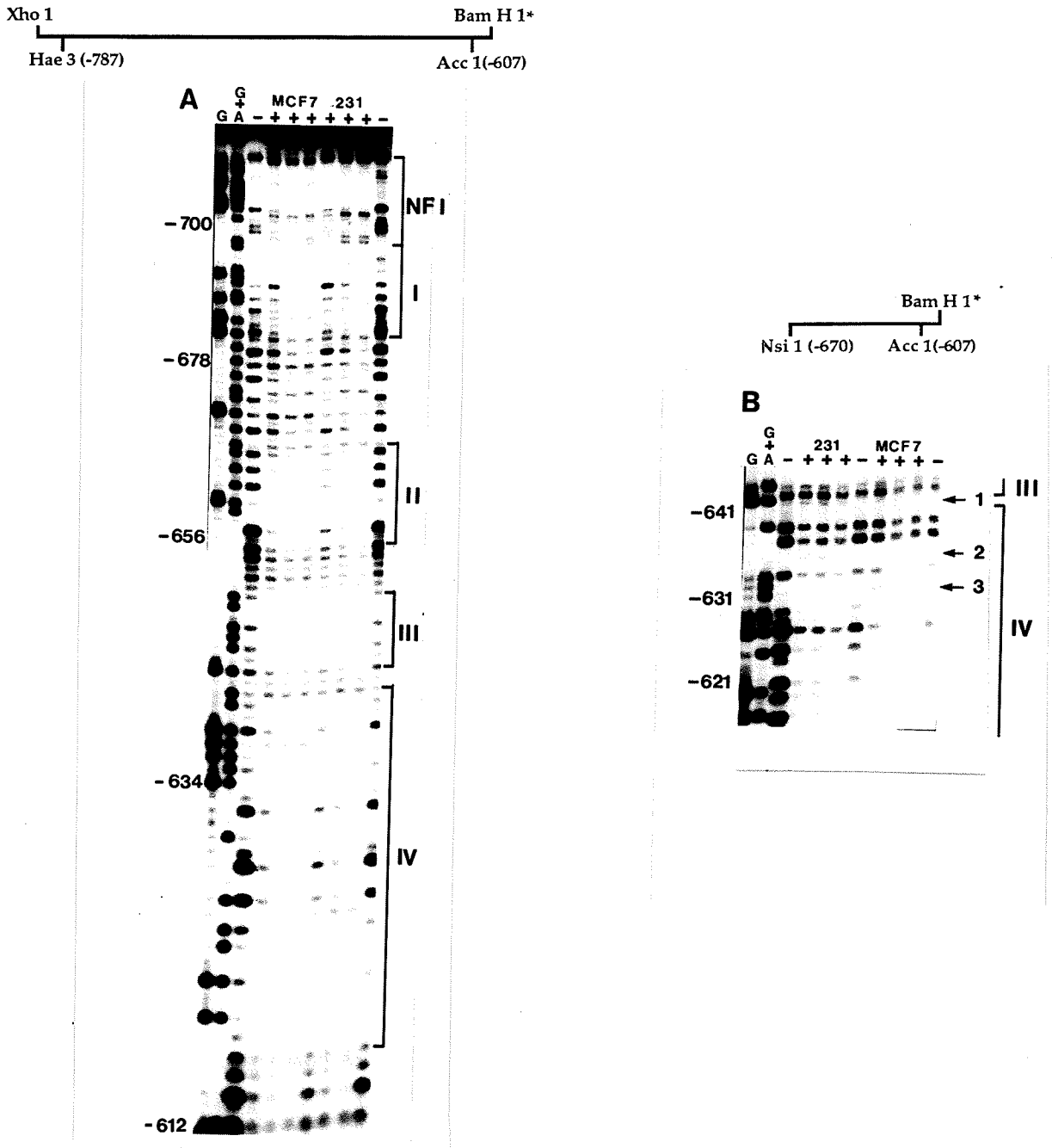


Figure 32. Novel Protein Interactions Identified in the Hae III-Acc I c-*myc* Region by DNase I Footprint Analysis. DNase I footprint analysis of the Hae III-Acc I and Nsi I-Acc I c-*myc* region was performed as described in Methods and Materials. The G and G+A reactions were run in the first two lanes. The reactions contained 3 ng of labelled DNA with (+) and without (-) 60 µg of nuclear extract from MDA MB 231 or MCF 7 breast cancer cells. (**A**) The Hae III-Acc I fragment was labelled with Klenow at the Hae III end (sense strand). The units of DNase I added to a 40 µl reaction were, from left to right, 0.01, 0.1 and 0.2 for both extracts and the control DNA without added protein. (**B**) The Nsi I-Acc I fragment was labelled with T4 polynucleotide kinase at the Acc I end using the Bam H I site of the plasmid (antisense strand). The units of DNase I added to a 40 µl reaction were 0.25, 0.2 and 0.15 for both extracts and 0.2, 0.15 and 0.1 for the control DNA without protein. The arrows indicate sites of *in vitro* DNase I hypersensitivity and the Roman numerals indicate the protected regions.

were localized within footprint IV. The MDA MB 231 nuclear extracts produced DNase I hypersensitive sites 1 and 2, while MCF 7 nuclear extracts generated DNase I hypersensitive site 3. Thus, the MCF 7 and MDA MB 231 nuclear extracts generated different EMSA and DNase I footprint patterns with the 63 bp Nsi I-Acc I region located between -670 and -607 relative to P_1.

In contrast to what was observed with the longer 203 bp Hae III-Acc I fragment (Figure 28), increasing the concentration of nuclear extract in DNA binding reactions with the shorter 63 bp Nsi I-Acc I fragment resulted in increased intensity of protein interactions with this region of c-*myc* (Figure 33 A). Additionally, self competition suggested that these interactions were specific (Figure 33 B). To tentatively assign the protein-DNA complexes formed with the 63 bp Nsi I-Acc I DNA fragment to specific footprints, oligonucleotides that encompassed footprints II to IV detected in the Nsi I-Acc I region were synthesized (Figure 34 A). Oligonucleotides O-II and O-II/III spanned protected regions II and II-III, respectively, with both overlapping the Nsi I site at the 5' end. The O-III/IV and O-IV oligonucleotides covered footprints III-IV and IV, respectively, with both overlapping a Bsp50I site at their 3' end. These two oligonucleotides were not synthesized to cover the Acc I site because Nsi I-Acc I and Nsi I-Bsp50I fragments gave the exact same pattern of DNA-protein complexes (data not shown). All four of the oligonucleotides were used in EMSA alongside the 63 bp Nsi I-Bam HI fragment which contained the Nsi I-Acc I c-*myc* region (Figure 34 B). As well, cross competitions between the oligonucleotides were used to confirm these preliminary protein assignments (Figures 35-37).

The DNA-protein complexes seen with the O-II/III oligonucleotide, which encompassed footprints II and III, correlated with complexes 3 (a or b)

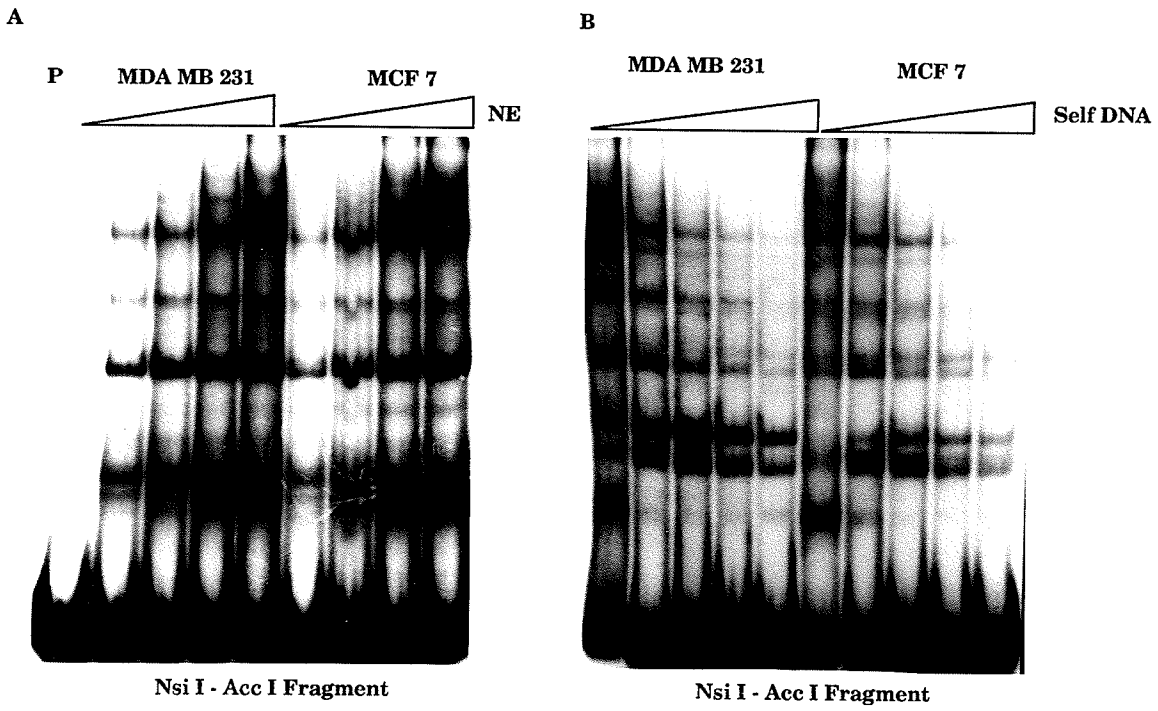


Figure 33. Protein Titration and Self Competition of the Nsi I-Acc I c-*myc* Region. The Nsi I-Acc I c-*myc* fragment was used in EMSA with increasing amounts of MCF 7 and MDA MB 231 breast cancer cell line nuclear extracts (NE) or with increasing amounts of competitor DNA (self DNA) as described in Methods and Materials. **(A)** End labelled DNA (1 ng) was incubated (from left to right) with no extract and 2.5, 5.0, 7.5 and 10.0 µg of MDA MB 231 or MCF 7 nuclear extracts. **(B)** Ten µg of MDA MB 231 or MCF 7 nuclear extracts were incubated with 1 ng of labelled Nsi I-Acc I c-*myc* DNA plus (from left to right) no competitor, 10 fold, 50 fold, 100 fold and 200 fold molar excess of unlabelled Nsi I-Acc I c-*myc* DNA.

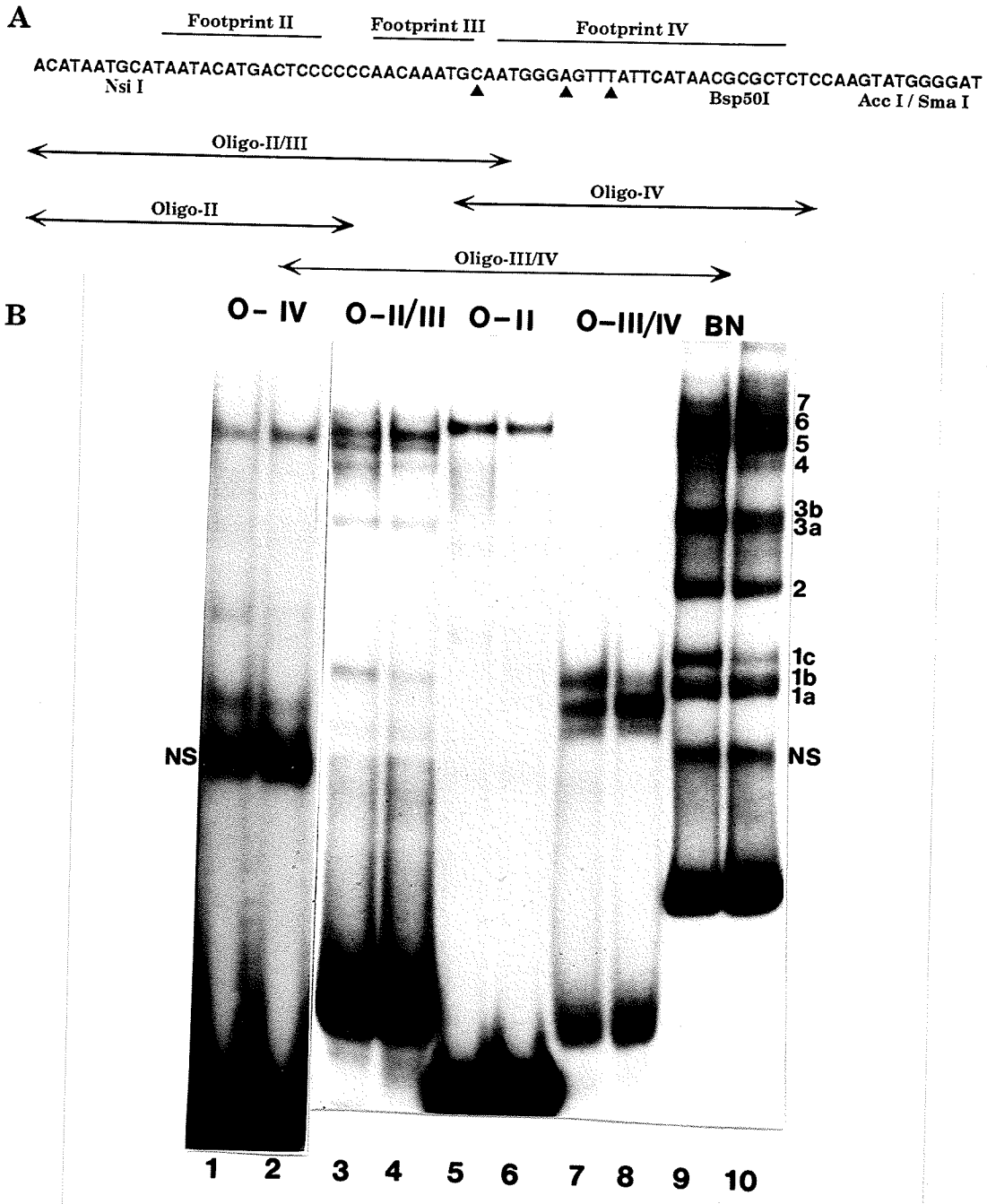


Figure 34. Protein Complexes Associated with the Novel DNase I Protected Sequences Identified in c-*myc*. A comparison of protein-DNA complexes between the oligonucleotides corresponding to the unique Nsi I-Acc I region of c-*myc*. **(A)** A diagram of the four oligonucleotides used in the EMSA and how they relate to the footprint regions within the Nsi I-Acc I c-*myc* DNA. The Nsi I and Bsp50I restriction sites and the Acc I/Sma I ligation site of c-*myc* and pSP73 are shown. The arrowheads indicate the three *in vitro* DH sites. **(B)** Nuclear extracts from MDA MB 231 (lanes 1, 3, 5, 7 and 9) and MCF 7 (lanes 2, 4, 6, 8 and 10) were incubated with the O-IV (lanes 1 and 2), O-II/III (lanes 3 and 4), O-II (lanes 5 and 6) and O-III/IV (lanes 7 and 8) oligonucleotides and compared to the protein-DNA complexes formed with the Nsi I-Acc I (BN) c-*myc* DNA (lanes 9 and 10).

and 4-6 seen with the 63 bp Nsi I-Bam HI fragment (Figure 34 B compare lanes 3 and 4 with 9 and 10). Conversely, the O-II oligonucleotide, which only encompassed footprint II, formed only one high molecular mass complex (Figure 34 B compare lanes 5 and 6 with 9 and 10). This suggested that protected region III was essential to the protein interactions (complexes 3a or 3b and 4-6) with O-II/III. Competition with the O-II/III (Figure 35 A), but not the O-II oligonucleotide (Figure 35 B), was able to competitively remove complexes 4-6 from O-II/III. Binding of the high molecular mass complex, possibly 7, formed with the O-II oligonucleotide was therefore solely due to the footprint II region (Figure 34 B, lanes 5 and 6). The O-III/IV oligonucleotide, which included footprints III and IV, bound proteins which correlated with complexes 1a-c (Figure 34 B compare lanes 7 and 8 with 9 and 10), as well as 2, 3 (a or b), 4 and 6 of the 63 bp Nsi I-Bam HI fragment, which were observed with other EMSAs (see panels B and D of Figures 36 and 37, or Figure 39). However, the O-IV oligonucleotide allowed the formation of only complexes 1c, 2 and 6 (Figure 34 B, compare lanes 1 and 2 with 9 and 10). The association of one or more of complexes 3 to 7 with all four oligonucleotides suggested that the generation of these high molecular mass complexes required footprints II, III and IV. Interestingly, the formation of complex 6 but not complex 4 on the O-III/IV oligonucleotide was competitively removed by both O-IV (Figure 36 D) and O-II/III (Figure 37B) oligonucleotides, but neither were removed with O-II (Figure 37 D). Similarly, the O-III/IV oligonucleotide and to a lesser extent the O-IV oligonucleotide, competitively removed complexes 4-6 from O-II/III (Figure 35 C and D). The formation of complex 3 (a or b) on both the O-II/III (Figure 35 A and C) and the O-III/IV (Figures 36 B and 37 B) oligonucleotides, was removed by self and cross competition with each other. Additionally, the

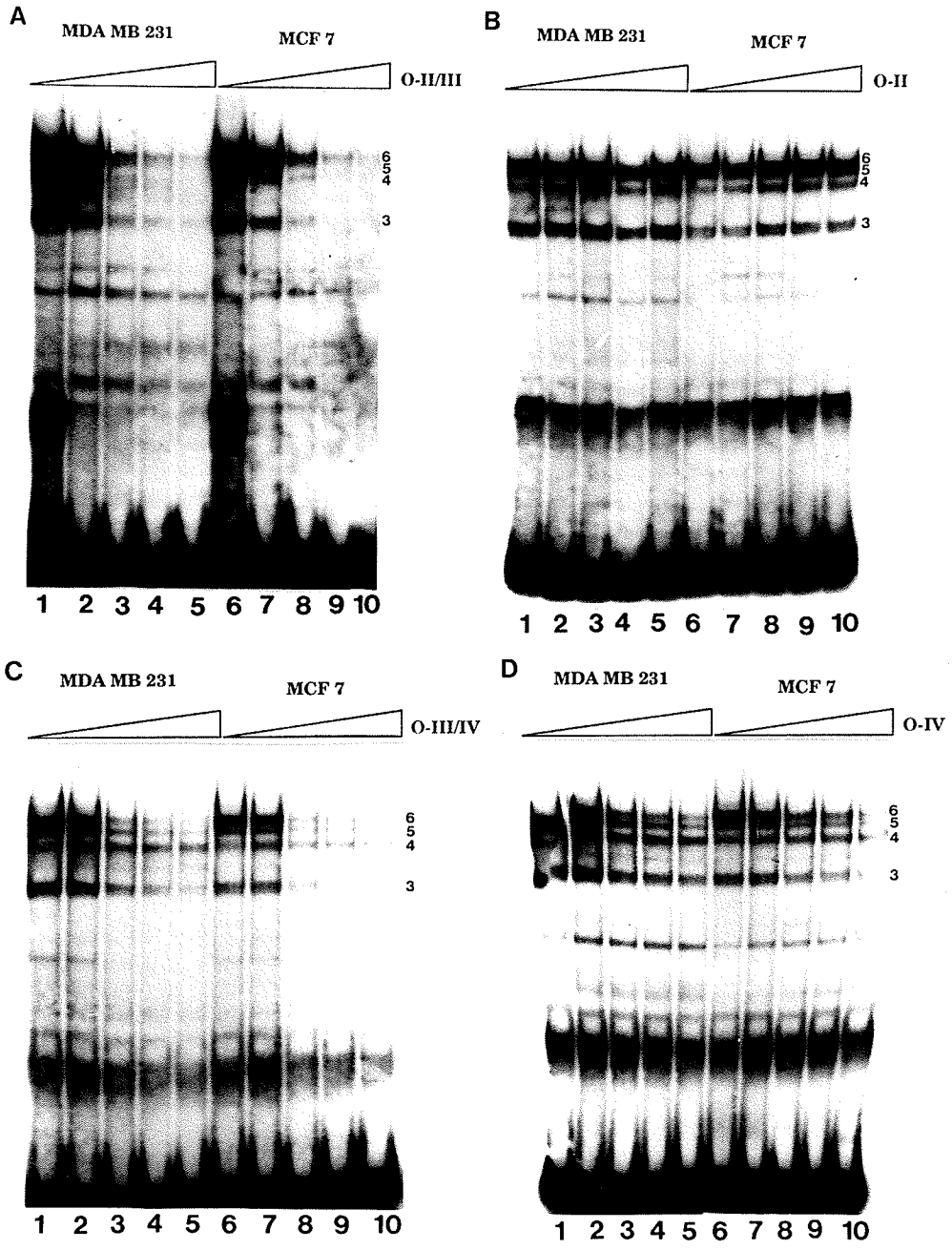


Figure 35. Protein Associations with c-*myc* Protected Region II/III. The four oligonucleotides which span the Nsi I-Acc I region of c-*myc* were used as competitors of protein-DNA complexes formed with O-II/III oligonucleotide as described in Methods and Materials. Complexes formed between 1 ng of this oligonucleotide and 10 μg of nuclear extracts from MDA MB 231 (lanes 1 to 5) and MCF 7 (lanes 6 to 10) breast cancer cell lines were competed for with 10 fold (lanes 2 and 7), 50 fold (lanes 3 and 8), 100 fold (lanes 4 and 9) and 200 fold (lanes 5 and 10) molar excess of the oligonucleotides O-II/III **(A)**, O-II **(B)**, O-III/IV **(C)** and O-IV **(D)**. The high molecular mass complexes 3 to 6 are indicated.

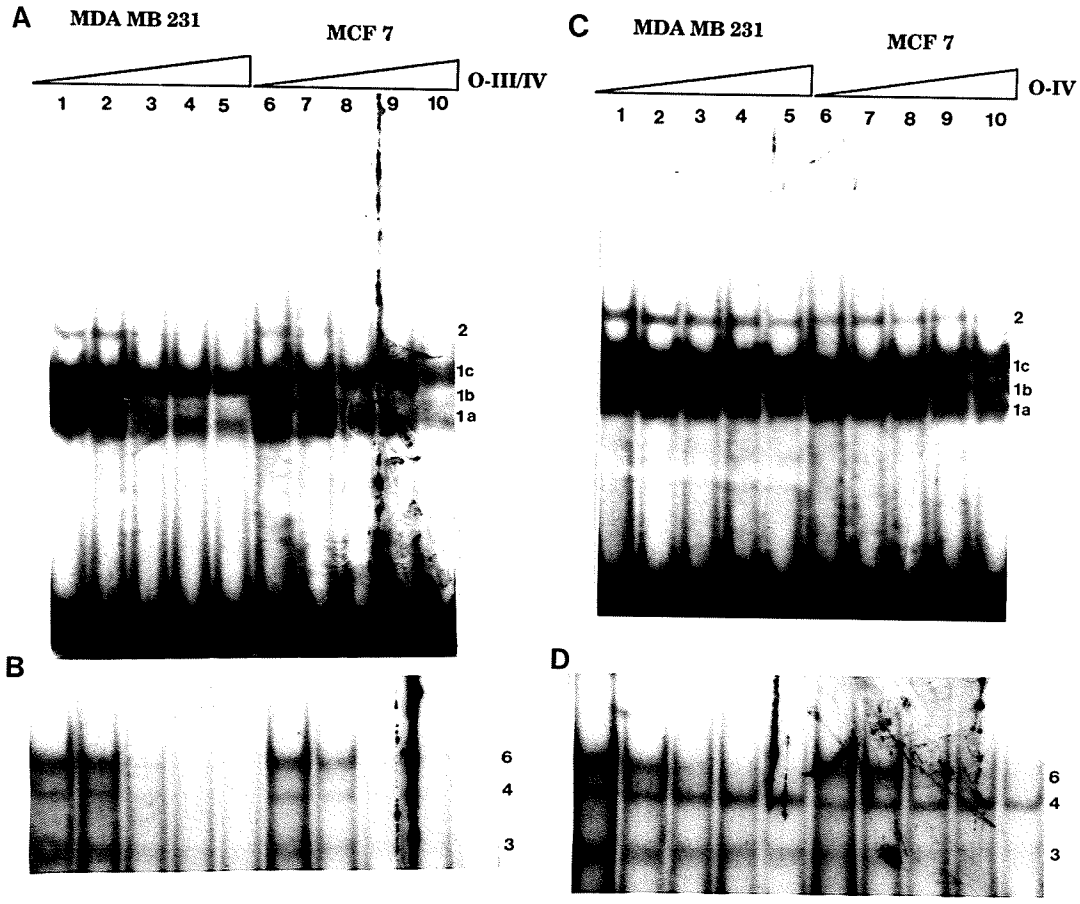


Figure 36. Protein Associations and Cross Competitions with c-*myc* Protected Regions III and IV. Oligonucleotide O-III/IV, which encompasses protected regions III and IV of the Nsi I-Acc I c-*myc* region, was end labelled and used in EMSA as described in Methods and Materials. Ten µg of nuclear extracts from MDA MB 231 (lanes 1 to 5) and MCF 7 (lanes 6 to 10) were incubated with 1 ng of the O-III/IV oligonucleotide and increasing amounts of competitor DNA. No competitor (lanes 1 and 6), 10 fold (lanes 2 and 7), 50 fold (lanes 3 and 8), 100 fold (lanes 4 and 9) and 200 fold (5 and 10) molar excess of the O-III/IV oligonucleotide (**A and C**) or the O-IV oligonucleotide (**B and D**) were used to compete for protein-DNA complexes. Specific protein-DNA complexes are labelled 1a, 1b, 1c, 2, 3, 4 and 6. Panels **B** and **D** represent exposures of the upper portion of the gel, where complexes 3, 4 and 6 migrate, which were at least twice as long as panels **A** and **C** and **Figure 34 B**.

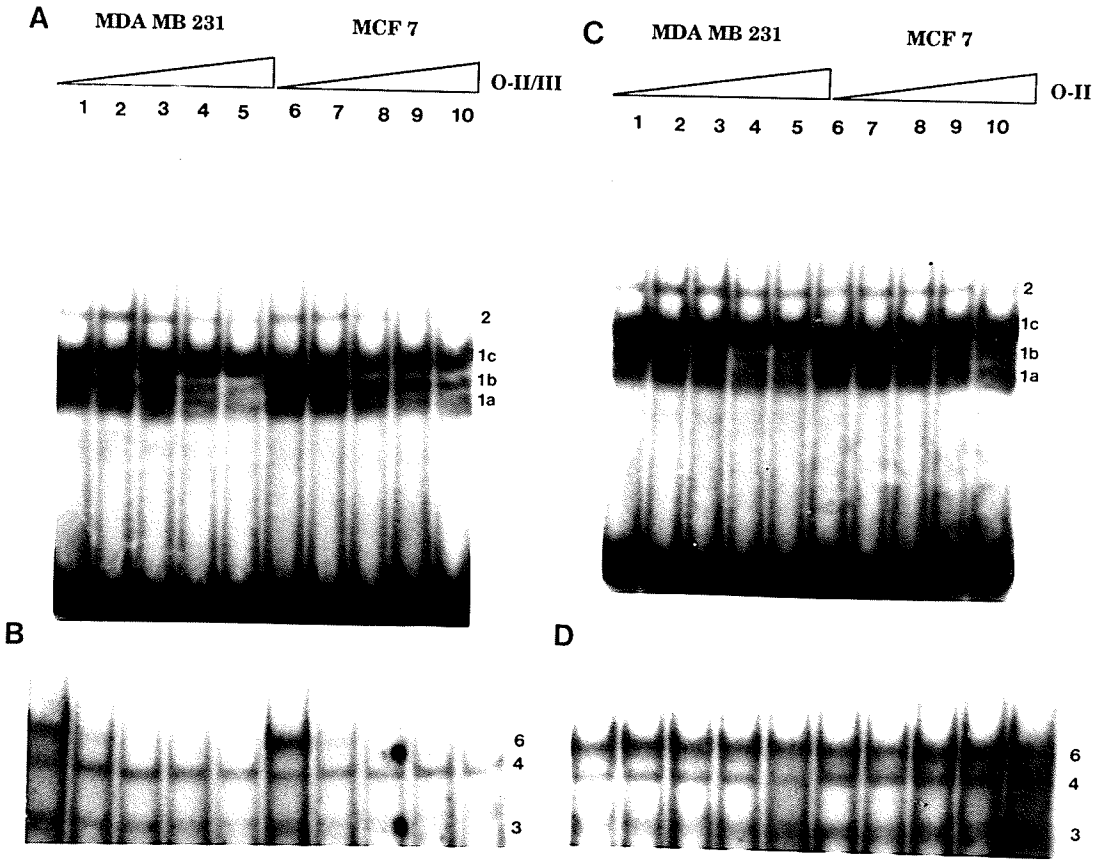


Figure 37. Competition of Proteins Associated with c-*myc* Protected Region III/IV by Protected Regions II and III. Oligonucleotide O-III/IV, which encompasses protected regions III and IV of the Nsi I-Acc I c-*myc* region, was end labelled and used in EMSA as described in Methods and Materials. Ten µg of nuclear extracts from MDA MB 231 (lanes 1 to 5) and MCF 7 (lanes 6 to 10) were incubated with 1 ng of the O-III/IV oligonucleotide and increasing amounts of competitor DNA. No competitor (lanes 1 and 6), 10 fold (lanes 2 and 7), 50 fold (lanes 3 and 8), 100 fold (lanes 4 and 9) and 200 fold (5 and 10) molar excess of the O-II/III oligonucleotide **(A and C)** or the O-II oligonucleotide **(B and D)** were used to compete for protein-DNA complexes. Specific protein-DNA complexes are labelled 1a, 1b, 1c, 2, 3, 4 and 6. Panels **B** and **D** represent exposures of the upper portion of the gel, where complexes 3, 4 and 6 migrate, which were at least twice as long as panels **A** and **C** and **Figure 34 B**.

O-IV oligonucleotide (Figures 35 D and 36 D) but not the O-II oligonucleotide (Figures 35 B and 37 D) was able to compete for this complex on either oligonucleotide. Significantly, the relative level and/or activity of complex 3 (a or b) was reproducibly greater in MDA MB 231 compared to MCF 7 nuclear extracts.

Differences in factor level and/or activity between the two breast cancer nuclear extracts were also observed with the O-III/IV and the O-IV oligonucleotide (Figure 34 B, see lanes 7,8 and 1,2). Complex 1b, seen only on O-III/IV, was consistently more abundant with MCF 7 nuclear extracts while complex 1c, observed with both oligonucleotides O-III/IV and O-IV, was reproducibly greater with MDA MB 231 nuclear extracts. The formation of complexes 1b and 1c appeared to require both footprints III and IV since these complexes were best observed with the O-III/IV oligonucleotide (Figure 34), which included sequences for both of these regions. Significantly, these observations were consistent with the *in vitro* DNase I hypersensitivity differences seen between the two extracts, which were also associated with footprints III and IV (Figure 32 B). The ability of oligonucleotides O-III/IV and O-II/III to competitively remove complexes 1a-c from the O-III/IV oligonucleotide supported these conclusions (Figures 36 A and 37 A, respectively). Additionally, complex 2, which migrated above complexes 1a-c, was also competitively removed with the O-III/IV oligonucleotide and weakly removed with O-II/III oligonucleotide (Figures 36 A and 37 A), suggesting this complex may also require the footprint III region of c-*myc* for binding.

None of the complexes formed on the O-II/III or the O-III/IV oligonucleotides could be removed with the addition of consensus AP1 oligonucleotides (Figure 38). This observation provided further support for the specific nature of the DNA-protein interactions with the O-II/III and

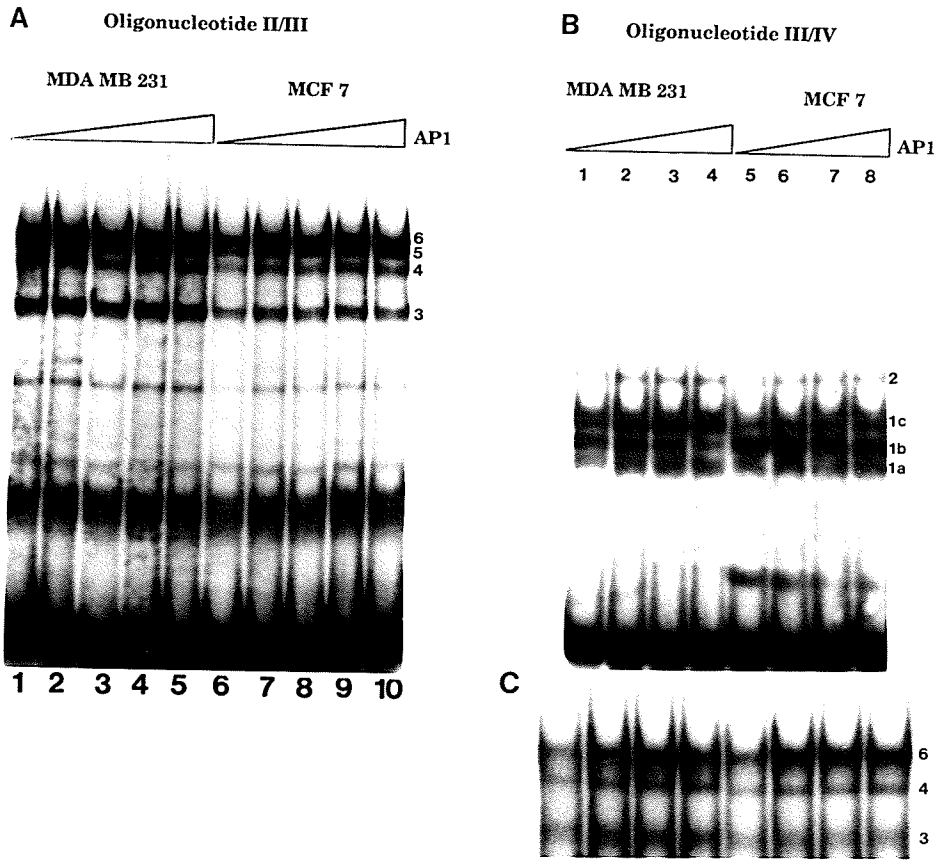


Figure 38. AP1 is Unable to Disrupt Protein Interactions with c-*myc* Protected Regions II/III and III/IV. Oligonucleotides O-II/III and O-III/IV which encompass the Nsi I-Acc I c-*myc* protected regions II/III and III/IV, respectively were end labelled and used in EMSA as described in Methods and Materials. **(A)** Ten μg of nuclear extracts from MDA MB 231 (lanes 1 to 5) and MCF 7 (lanes 6 to 10) were incubated with 1 ng of the O-II/III oligonucleotide in the presence of no competitor (lanes 1 and 6), 10 fold (lanes 2 and 7), 50 fold (lanes 3 and 8), 100 fold (lanes 4 and 9) and 200 fold (lanes 5 and 10) molar excess of consensus Ap1 oligonucleotide. **(B)** Ten μg of nuclear extracts from MDA MB 231 (lanes 1 to 4) and MCF 7 (lanes 5 to 8) were incubated with 1 ng of the O-III/IV oligonucleotide in the presence of no competitor (lanes 1 and 5), 6.75 fold (lanes 2 and 6), 40 fold (lanes 3 and 7) and 65 fold (lanes 4 and 8) molar excess of consensus Ap1 oligonucleotide. **(C)** A longer exposure of the top portion of panel **B**. Specific protein-DNA complexes are indicated beside each panel.

O-III/IV oligonucleotides. Interestingly, methylation of the Nsi I-Acc I fragment, the O-III/IV and the O-II/III oligonucleotides disrupted the ability of MCF 7 nuclear extracts to form DNA-protein interactions (Figure 39). Specifically, the high molecular mass complexes formed with both the unmethylated Nsi I-Acc I fragment and O-II/III oligonucleotide were decreased with a concomitant increase in the faster migrating complexes in the presence of methylated DNA (Figure 39 A compare lanes 2-4 with 6-8; Figure 39 B compare lanes 3 and 4 with 7 and 8). As well, complex 1a was unable to form with the methylated O-III/IV oligonucleotide (Figure 39 B compare lanes 1 and 2 with 5 and 6).

The *in vitro* DNA binding studies therefore suggested that the protein interactions at footprints II, III and IV were responsible for the multiple complexes formed between MCF 7 and MDA MB 231 nuclear extracts and the 63 bp Nsi I-Acc I c-*myc* fragment. Proteins associated with the footprint I region, which maps 5' of the Nsi I site, do not appear to be essential for the generation of protein interactions with the footprint II, III and IV sequences. Preliminary assignments suggest that most of the protein interactions with the Nsi I-Acc I c-*myc* region required footprint III or IV, since oligonucleotide O-II/III formed complexes 3 (a or b) plus 4-6, oligonucleotide O-III/IV formed complexes 1a-c, 2, 3 (a or b), 4 and 6, while oligonucleotide O-IV formed complexes 1c, 2 and 6. Preliminary evidence suggests that some of the protein-DNA interactions which formed the higher molecular mass complexes and complex 1a may be disrupted by methylation at guanine residues. The footprint II region appeared to be independently necessary only for complex 7, which was formed on oligonucleotide O-II. Interestingly, the differential *in vitro* DNase I hypersensitive sites detected between ER^+ MCF 7 and ER^- MDA MB 231 nuclear extracts were observed between footprints III and IV

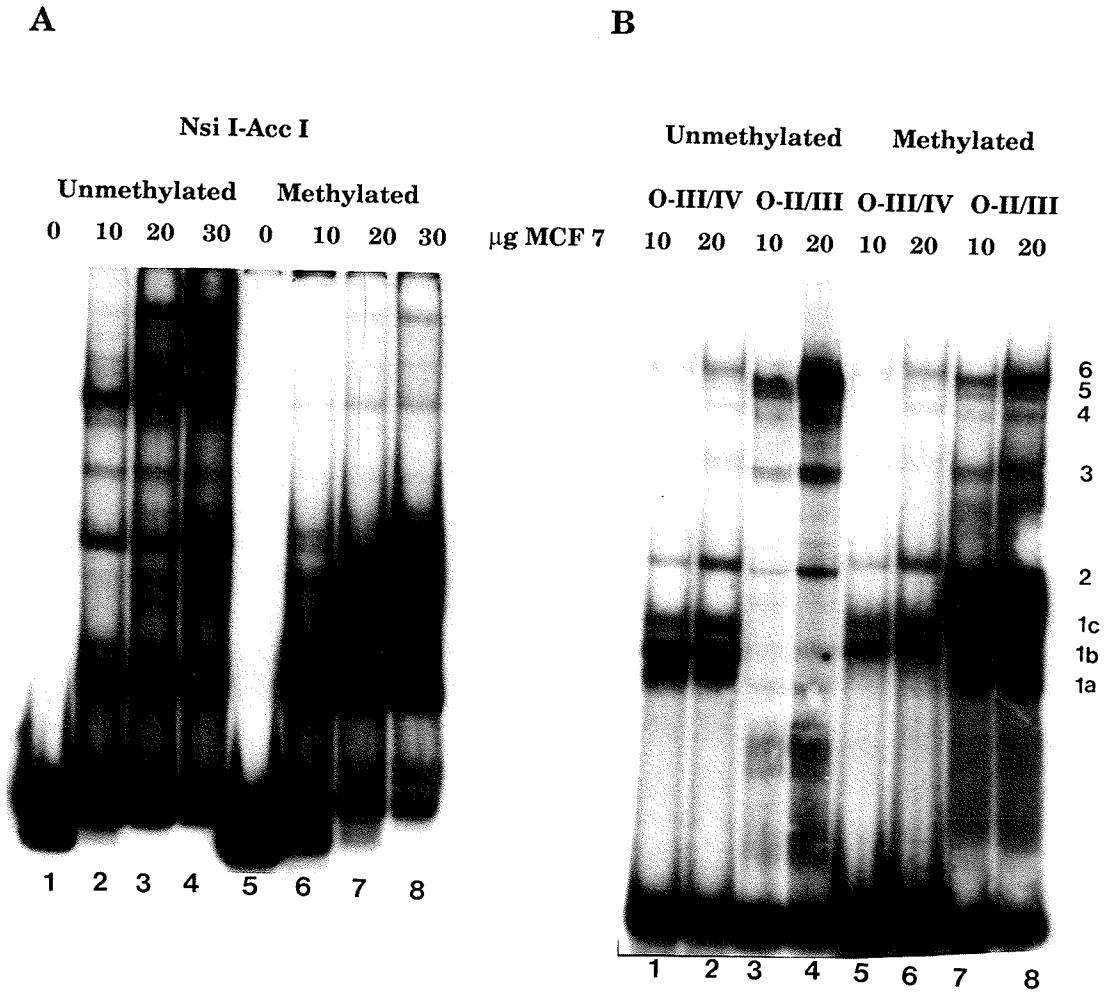


Figure 39. Methylation at Guanine Residues Disrupts Protein Interactions. Nsi I-Acc I c-*myc* fragment and oligonucleotides O-II/III and O-III/IV were used in EMSA with and without methylation at guanine residues as described in Methods and Materials. **(A)** No extract (lanes 1 and 5) and 10 µg (lanes 2 and 6), 20 µg (lanes 3 and 7) and 30 µg (lanes 4 and 8) of MCF 7 nuclear extracts were incubated with unmethylated (lanes 1 to 4) and methylated (lanes 5 to 8) end labelled Nsi I-Acc I fragment. **(B)** Ten µg (lanes 1, 3, 5 and 7) and 20 µg (lanes 2, 4, 6, and 8) of MCF 7 nuclear extract were incubated with unmethylated O-III/IV (lanes 1 and 2), unmethylated O-II/III (lanes 3 and 4), methylated O-III/IV (lanes 5 and 6) and methylated O-II/III (lanes 7 and 8). Specific protein-DNA complexes are indicated on the side of each panel.

and within footprint IV. This correlated with some of the differences in protein complex formation. Although nuclear extracts isolated from MDA MB 231 and MCF 7 breast cancer cell lines contained a similar spectrum of sequence specific DNA binding proteins that interacted with the DH II_2 5' regulatory regions of the c-*myc* gene, differences observed in the formation of complexes 3 (a or b), 1b and 1c suggested that the nuclear protein profile (abundance and/or activity) of these cells was not identical.

Protein Associations with DH Site III_2

The major c-*myc* promoter used in both hormonally dependent and independent human breast cancer cells was P_2 (Miller *et al.*, 1993). DNA sequences immediately 5' of this promoter were determined to be necessary for estrogen regulated transcription and a *cis*-acting element responding to estrogen has been localized in the +25 to +141 region of the c-*myc* gene (Dubik and Shiu, 1992). Furthermore, this 116 bp region of the c-*myc* gene was associated with the DH III_2 site (see Figures 16 and 27). It was therefore important to investigate if proteins present in human breast cancer cell nuclear extracts could specifically interact with this region. Initially, c-*myc* fragments encompassing the Ava II-Cfr 101 sites at -1 to +140 were used for EMSA (Figure 40), and no differences in complex pattern were observed between the MCF 7 and MDA MB 231 nuclear extracts (Figure 40 lanes 1-5).

The Ava II-Cfr101 region of the c-*myc* gene has several binding sites for previously described *trans*-acting proteins between +67 and +140 relative to P_1 (see Figure 27), including ME1a1, ME1a2, E2F and MBP-1 (Asselin *et al.*, 1989; Hall, 1990; Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Moberg *et al.*, 1991,1992; Plet *et al.*, 1992; Ray and Miller, 1991; Thalmeier *et al.*, 1989). To determine which of these binding sites was responsible for the protein

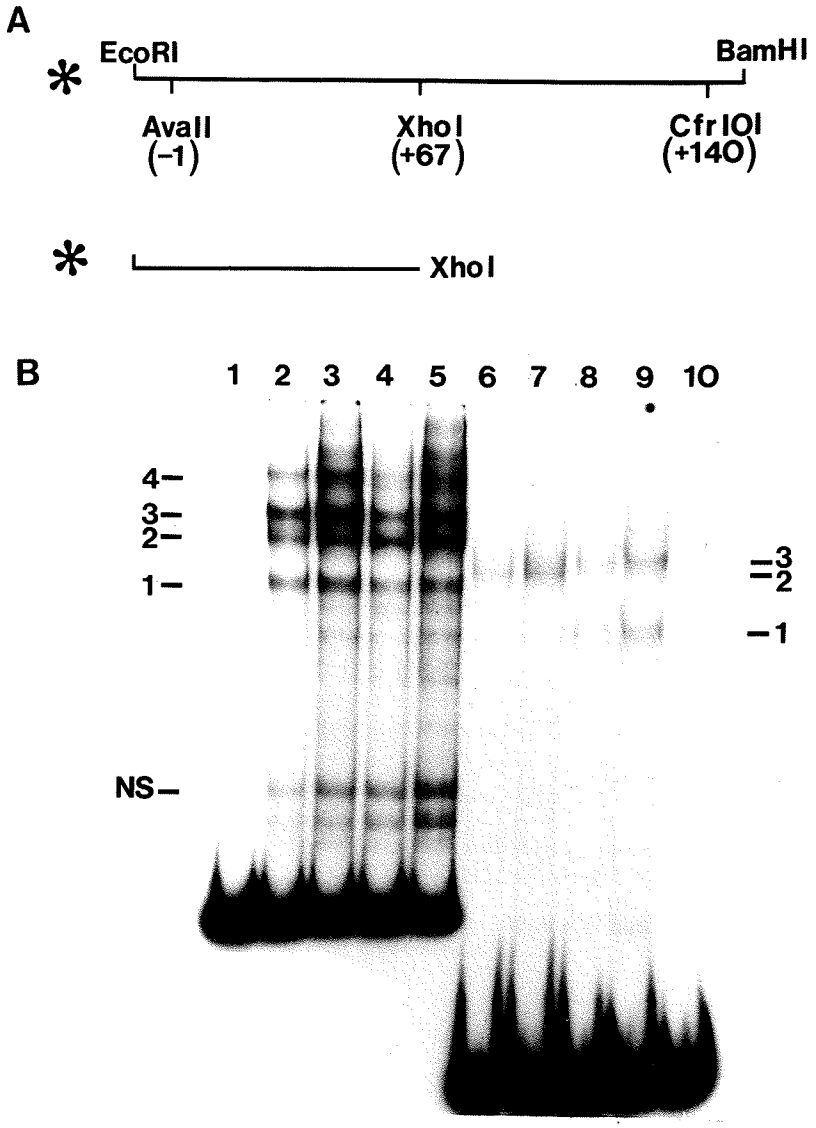


Figure 40. Protein Interactions at c-*myc* DNase I Hypersensitive Site III2. The Ava II-Cfr101 and Ava II-Xho I sequences surrounding the c-*myc* DH III2 were end labelled and used in EMSA as described in Methods and Materials. **(A)** The Ava II-Cfr 101 and Ava II-Xho I fragments were labelled as shown by the asterisks. **(B)** The Ava II-Cfr 101 (lanes 1 to 5) and the Ava II-Xho I (lanes 6 to 10) fragments were incubated with no extract (lanes 1 and 10), 2 μg (lanes 2, 4, 6, 8) or 4 μg (lanes 3, 5, 7, 9) of nuclear extracts from MCF 7 (lanes 4, 5, 8, 9) and MDA MB 231 (lanes 2, 3, 6, 7) breast cancer cell lines. The specific complexes are labelled and NS indicates non-specific binding. This study was done in collaboration with Dr. Yan Jin.

interactions identified with MCF 7 and MDA MB 231 nuclear extracts, oligonucleotides representing these factors were used as competitors for complexes with this c-*myc* sequence (Figure 41). The E2F and MBP-1 oligonucleotides did not competitively remove any of the complexes (Figure 41 lanes 6,7 and 8,9, respectively), however, ME1a1 and ME1a2 oligonucleotides competed equally for complexes 1 to 4 (Figure 41 lanes 2,3 and 4,5, respectively). When the oligonucleotides were end labelled and used as probes in EMSA, specific complexes were formed with the ME1a1, ME1a2 and MBP-1 oligonucleotides (Figure 42 A). However, protein-DNA complexes did not form with the E2F oligonucleotide (data not shown). The complexes formed with the ME1a1, ME1a2, and S2 oligonucleotides had similar electrophoretic mobilities which were quite different from that of the complex produced with MBP-1 (Figure 42 A). The S2 oligonucleotide represented the previously characterized Sp1 binding site of the P_0 promoter (Lang *et al.*, 1991). These results and a comparison of the putative Sp1 binding sequences with the consensus Sp1 site (Figure 42 B) provided evidence that Sp1-like proteins were most likely interacting with ME1a1 and ME1a2 binding sites.

Binding sites for sequence specific DNA-proteins have not been previously identified 5' of the Xho I site in the Ava II-Cfr 101 fragment (Figures 40 A, 27 and 15). Therefore, a fragment encompassing the Ava II-Xho I sites at -1 to +67, relative to P_1 was isolated (Figure 40 A) and EMSA was used to determine whether proteins from breast cancer cell nuclear extracts could recognize DNA elements in this region. Three protein-DNA complexes were formed with this Ava II-Xho I fragment (Figure 40 B, lanes 6 to 10). The results suggested that uncharacterized binding proteins interacted with this region and differences between nuclear extracts from MDA MB 231 and MCF 7 cells were apparent (Figure 40 B, compare

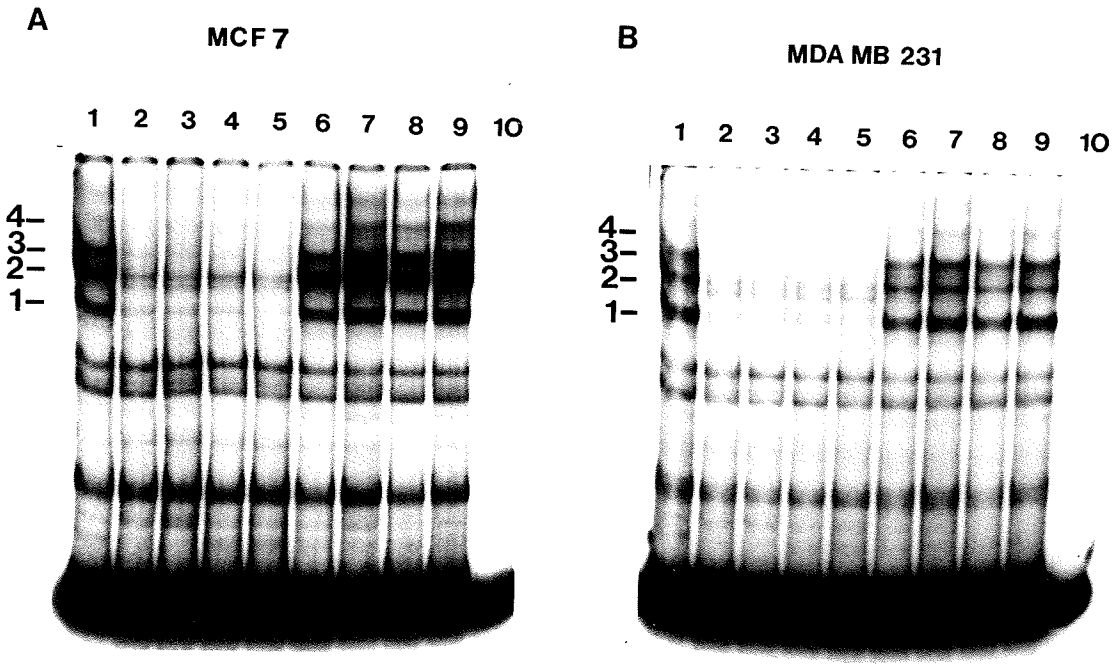
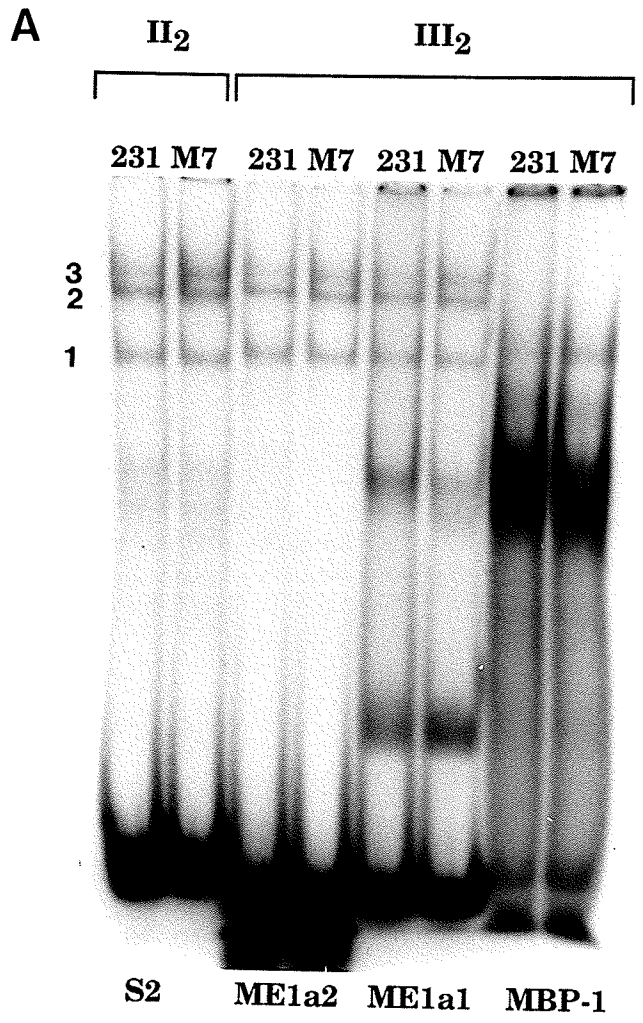


Figure 41. Competition of the Ava II-Cfr101 c-*myc* Fragment with Binding Sites for ME1a1, ME1a2, E2F and MBP-1. The Ava II-Cfr101 DH III$_2$ region was end labelled and used in EMSA in the presence of oligonucleotides representing previously identified protein binding sites as described in Methods and Materials. End labelled AvaI II-Cfr 101 fragment was incubated without (lane 10) or with (lanes 1 to 9) 4 μg of MCF 7 (**A**) or MDA MB 231 (**B**) nuclear extracts plus 50 fold (lanes 2, 4, 6 and 8) or 100 fold (lanes 3, 5, 7 and 9) molar excess of synthetic oligonucleotides ME1a1 (lanes 2 and 3), ME1a2 (lanes 4 and 5), E2F (lanes 6 and 7) and MBP-1 (lanes 8 and 9). This study was done in collaboration with Dr. Yan Jin.



B

Consensus Sp1	G/T	G/A	G	G	T A C	G/T	G/A	G/A	G/T
Sp1 Oligo	G	G	G	G	C	G	G	G	G
S2	C	A	G	G	A	G	G	G	G
ME1a1/CT-I2	A	G	G	G	A	G	G	G	A
ME1a2	A	G	G	G	C	A	G	G	G
CT 1-5	G/T	G/A	G	G/A	A	G	G	G	G

Figure 42. Sp1-like Proteins Interact with Binding Sites S2, ME1a2 and ME1a1 of the c-*myc* Gene. The S2, ME1a2, ME1a1 and MBP-1 oligonucleotides were end labelled and used in EMSA as described in Methods and Materials. **(A)** The oligonucleotides S2, ME1a1, ME1a2 and MBP-1 were incubated with 4 μg of nuclear extract from MDA MB 231 (231) and MCF 7 (M7) breast cancer cell lines. **(B)** A comparison of the consensus Sp1 binding site with potential Sp1 binding sites within the human c-*myc* gene. Bold letters represent non-consensus sequences. This study was done in collaboration with Dr. Yan Jin.

lanes 6 and 7 with 8 and 9). In particular, nuclear extracts from MDA MB 231 cells formed complexes 1 to 3 with complex 2 being the most prominent (Figure 40 B lanes 6 and 7). MCF 7 nuclear extracts generated similar complexes, but complex 3 was the most abundant (Figure 40 B lanes 8 and 9).

Novel DNA-binding Sites at c-*myc* DH Site III_2

Using the DNase I footprinting technique, proteins from both breast cancer cell nuclear extracts were found to interact with sequences located in the Ava II-Cfr 101 c-*myc* DNA (Figure 43). Protected region II contained the ME1a2 sequence and protected region III contained the binding sites for E2F, ME1a1 and MBP-1. Two *in vitro* DNase I hypersensitive sites were noted between footprints II and III (see arrows in Figure 43). No major differences between cell lines were observed in these footprints. Footprint I spanned the region +45 to +65, which was detected 5' to the Xho I site at +67 (Figure 27). *In vitro* DNase I hypersensitive sites were observed 5' to footprint I, however the intensities of these hypersensitive sites differed between nuclear extracts from MCF 7 and MDA MB 231 human breast cancer cell lines (see arrows Figure 43). DNase I footprinting of the same fragment using nuclear extracts from HeLa cells demonstrated a similar pattern to that obtained with MCF 7 nuclear extracts (data not shown). The results provided evidence for potentially novel factor binding sites in the DH site III_2 region spanning -1 to +66 relative to P_1. Additionally, a different level and/or activity of sequence specific DNA binding proteins from MDA MB 231 and MCF 7 cells were found to interact with the c-*myc* DH III_2 region.

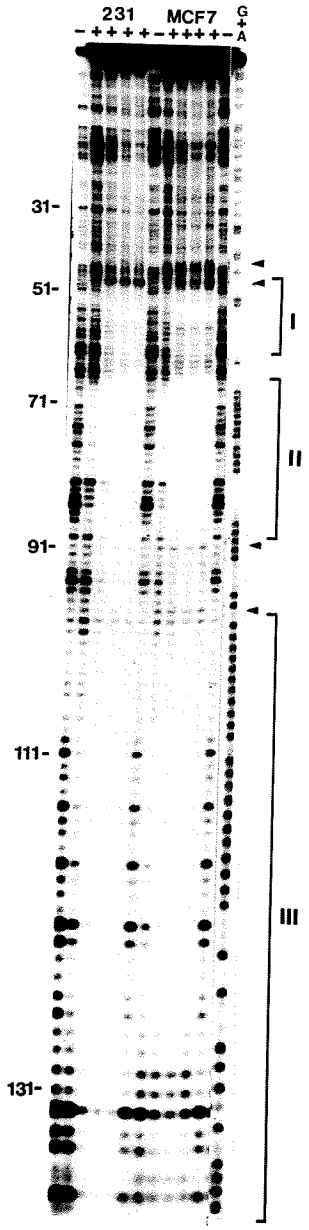


Figure 43. Novel Protein Associations with the c-*myc* Ava II-Xho I Region Identified by DNase I Footprinting. The Ava II-Cfr 101 c-*myc* fragment was end labelled at the Cfr 101 end using the Bam HI site of the plasmid (sense strand) and used in DNase 1 footprinting as described in Methods and Materials. The AvaII-Cfr 101 fragment was incubated without any added nuclear extract (-) or with 50 μg of nuclear extract (+) from MDA MB 231 (231) or MCF 7 breast cancer cell lines. The amount of DNase I added to the 50 μl reactions from left to right was 0.2, 0.1, 0.2, 0.4, 0.1, 0.2, 0.4 and 0.01 units. Protected regions are labelled I, II (ME1a2) and III (E2F, ME1a1 and MBP-1). The arrows identify *in vitro* DNase I hypersensitive sites which differ between MDA MB 231 and MCF 7 cells. This study was done in collaboration with Dr. Yan Jin and Jian-Min Sun.

Discussion

Breast cancer is a major health problem facing women from western society. A major component of the clinical management of this disease has focused on hormonal manipulation, but quite often the initial successes are followed by treatment failure (McGuire *et al.*, 1987; Bonadonna and Valagussa, 1988). Therefore, an important problem in the biology of breast cancer concerns the molecular mechanisms responsible for the progression from hormonal dependence to hormonal independence. Although loss of the estrogen receptor could explain some of this progression, a significant proportion of breast tumors classified as ER^+ do not respond to the antiestrogen tamoxifen or other forms of endocrine therapy (McGuire, 1980; Arafah and Pearson, 1986). The activity of the ER is generally mediated through an interaction with DNA elements on estrogen responsive genes (Kumar and Chambon, 1988; Klein-Hitpass *et al.*, 1989; Gorski, 1993). Loss of the ER DNA binding or *trans*-activating function, which would not be detected by hormone binding assays, has therefore been suggested as another reason for therapy failure (Foster *et al.*, 1991). Alternatively, a loss in the ability of estrogen regulated genes to respond to the presence or absence of estrogens may be involved in the molecular mechanisms responsible for the hormonally independent phenotype (Darbre and King, 1987).

The c-*myc* gene is differentially regulated in ER^+ and ER^- human breast cancer cells (Dubik *et al.*, 1987; Dubik and Shiu,1988). This suggested that the c-*myc* gene may be an excellent model for analyzing the progression of gene expression from hormonal dependence to independence. Significantly, the overexpression or amplification of this gene is correlated with a poor patient prognosis (Berns *et al.*, 1992). In addition to regulating proliferation

(Kaczmarek *et al.*, 1985), c-Myc has been reported to be a regulator of both differentiation and apoptosis (Miner and Wold, 1991; Evan *et al.*, 1992). The demonstration that c-Myc exerts these functions by acting as a transcription factor (Blackwell *et al.*, 1990; Prendergast and Ziff, 1991) suggests that alterations in its regulation will have profound cellular effects. This is supported by the observation that the relative levels of Myc, Max and other related proteins are important determinants of movement into and out of the cell cycle (Amati *et al.*, 1993; Amin *et al.*, 1993).

Estrogen induces the proliferation of responsive neoplastic breast cells both in culture and when grown in the nude mouse (Engel and Young, 1978; Soule and McGrath, 1980). Inhibition of c-*myc* proto-oncogene expression with antisense RNA prevents the estrogen induced proliferation of ER^+ breast cancer cells and also arrests the growth of estrogen independent cells (Watson *et al.*, 1991). Therefore, the c-*myc* gene plays an important role in the proliferation of both hormone responsive and nonresponsive human breast cancer cell lines. Although estrogen stimulates c-*myc* expression in ER^+ breast cancer cells, there is no correlation between the level of c-*myc* expression and estrogen receptor status of breast cancer cell lines (Figure 18). This is consistent with what has been found in human breast cancer biopsies (Escot *et al.*, 1986). Additionally, the proliferation rate of breast cancer cell lines is not associated with ER status, but does correlate somewhat with the level of c-*myc* expression. Generally those cell lines with higher steady state levels of c-*myc* mRNA, MCF 7, T47 D5, MDA MB 231, and ZR 75-1 (Figure 18), also have the faster doubling times (Murphy and Dotzlaw, 1989c). An exception is the T47 D5 cell line, which has a higher steady state level of c-*myc* expression than the MCF 7 cell line even though the average doubling time of T47 D5 cells (1.33 days) is somewhat longer than MCF 7 cells (1.27

days) (Murphy and Dotzlaw, 1989c). However, evidence suggests that there is a higher level of c-*myc* amplification in T47 D5 cells compared to MCF 7 cells which may contribute to this difference (Wong and Murphy, 1991).

In hormonally dependent, estrogen receptor positive human breast cancer cells, c-*myc* expression is transcriptionally regulated by estrogen (Dubik and Shiu, 1988). However, in hormonally independent, estrogen receptor negative human breast cancer cells, constitutive c-*myc* expression is observed (Dubik *et al.*, 1987; Dubik and Shiu, 1988). The molecular mechanisms associated with this differential regulation of c-*myc* in breast cancer cells are largely unknown, but have been proposed to be partially due to the interaction of the ER with a half site ERE located between the P_1 and P_2 promoters (Dubik and Shiu, 1992). The presence of the ER may not be solely responsible for regulatory differences in c-*myc* expression between ER^+ and ER^- breast cancer cells, since reexpressing the ER does not reconstitute an estrogen increased proliferative response in breast cancer cells (Jiang and Jordan, 1992) or in other cell types (Kushner *et al.*, 1990; Touitou, *et al.*, 1990; Maminta *et al.*, 1991). Indeed, estrogen, if anything, decreased proliferation, although an appropriate estrogen response was observed with an exogenously added estrogen responsive reporter gene, the endogenous PR gene (Jiang and Jordan, 1992) and the endogenous cathepsin D gene (Touitou, *et al.*, 1990; Cavaillès *et al.*, 1993). Furthermore, an estrogen stimulated increase in the expression of the endogenous c-*myc* gene has not been reconstituted (Maminta *et al.*, 1991), and preliminary studies suggest that c-*myc* expression in the S30 cell line, which has had the ER reintroduced, may be decreased with the addition of estrogen (V.C. Jordan, personal communication). Together, these studies suggest that additional

factors are necessary for the reconstitution of estrogen induced cell growth and c-*myc* gene expression.

At least two different levels of regulation are involved in controlling gene expression (reviewed in Weintraub, 1985; Mitchell and Tjian, 1989; Felsenfeld, 1992). A general mechanism of regulation, which must occur prior to gene expression, involves the unfolding of the chromatin domain into a conformation that is more accessible to the transcription apparatus. These transcriptionally competent chromatin domains are then specifically regulated by mechanisms which determine the rate of gene transcription. These processes require the coordinated action of factors and elements which determine nuclear matrix attachments (Getzenberg *et al.*, 1991; Bonifer *et al.*, 1991), histone H1 interactions (Grunstein, 1992; Workman and Buchman, 1993), nucleosome positioning (Lee and Garrard, 1991; Hayes and Wolffe, 1992) and transcription factor associations (Mitchell and Tjian, 1989; Martin, 1991; Herschlag and Johnson, 1993) with chromatin. The implementation of these regulatory mechanisms results in differences in chromatin structure between transcribed and nontranscribed genes, which are detected as changes in nuclease sensitivity and hypersensitivity. These measurable alterations in chromatin structure can therefore be used to analyze regulatory differences in gene expression.

The chromatin structure of the human c-*myc* gene has been previously analyzed in cell types other than breast cancer cells. At least eight DNase I hypersensitive sites (Figure 11) have been associated with the 5' untranslated regions of c-*myc* (Siebenlist *et al.*, 1984; Dyson *et al.*, 1985; Bentley and Groudine, 1986a; 1986b; Eick *et al.*, 1990) and four of these DH sites are associated with promoter elements (Siebenlist *et al.*, 1984; Bentley and Groudine, 1986a; 1986b; Eick *et al.*, 1990). The nuclease sensitivity of c-*myc*

chromatin has also suggested that positioned nucleosomes may be present 5' of the P_0 promoter, while the region between the P_0 and P_1 promoters may exclude nucleosomes (Kumar and Leffak, 1989; 1991). A detailed analysis of the chromatin structure associated with the estrogen regulated c-*myc* proto-oncogene compared to that of the constitutively activated gene in hormonally independent breast cancer cells is therefore crucial to an understanding of molecular changes involved in the progression of this gene from hormonal dependence to independence.

The results of the DNase I hypersensitivity studies of the 5' flanking region of c-*myc* gene chromatin in both ER^+ and ER^- human breast cancer cells (Figures 19 and 20) indicate that DH sites I, II_2, III_2, IVb and V are located in the same positions as those characterized in HL60 cells (Bentley and Groudine, 1986a; Siebenlist *et al.*, 1988). In the c-*myc* chromatin of breast cancer cells DH site III_1 is occasionally resolved from DH site III_2 and is also similarly located in HL60 cells (Siebenlist *et al.*, 1988). DNase I hypersensitivity at sites III_1 and III_2 has been associated with the transcriptional activity of the P_1 and P_2 promoters, respectively (Siebenlist *et al.*, 1984; 1988). For all of the breast cancer cell lines studied, DH site III_2 is more intense than DH site III_1 suggesting that in breast cancer cells, the P_2 promoter is the more active of the two promoters regardless of ER status. Studies using the RNase protection assay support this suggestion (Figure 24), and provide no evidence for differential promoter usage between ER^+ and ER^- cells.

The DNase I hypersensitivity analysis also provides evidence that the human c-*myc* chromatin of ER^- breast cancer cells is different from that of ER^+ cells since there is differential accessibility of the DH sites to nuclease attack. Increased sensitivity to DNase I digestion at DH site II_2 is observed

in ER$^-$ cells relative to ER$^+$ cells (Figures 19 and 20). The relative intensities of DH sites I, II$_2$ and III$_2$ in ER$^-$ cells are similar to those of undifferentiated HL60 cells, while the intensities of these DH sites in ER$^+$ cells are similar to those of differentiated HL60 cells (Bentley and Groudine, 1986a; Siebenlist *et al.*, 1988). This correlation is interesting since ER$^+$ breast cancer cells are considered to be phenotypically more differentiated than ER$^-$ cells (Williams *et al.*, 1987). Consistent with this hypothesis is the observation that expression of the ER in breast tumors is correlated with a better prognosis and longer overall survival of patients (McGuire *et al.*, 1975; Williams *et al.*, 1987; Blanco *et al.*, 1984; Barbi *et al.*, 1987).

The DH II$_2$ site is located about 100 bases upstream of the human c-*myc* P$_0$ promoter and increased nuclease sensitivity at this site has been suggested to reflect the function of this promoter (Bentley and Groudine, 1986a; 1986b). Additionally, these authors have proposed that P$_0$ may affect the function of the P$_1$ and P$_2$ promoters by acting as an entry site for RNA polymerases. Differential DNase I hypersensitivity of DH site II$_2$ between ER$^+$ and ER$^-$ human breast cancer cells may reflect differential c-*myc* promoter usage between these cells. However, under the conditions of these experiments no significant use of the P$_0$ promoter for the initiation of c-*myc* transcription was observed and the relative use of the P$_1$ and P$_2$ promoters was similar for both cell types (Figure 24). This suggests that the differential sensitivity of DH site II$_2$ in ER$^+$ and ER$^-$ breast cancer cells is not associated with differential promoter usage.

The c-*myc* proto-oncogene is genetically altered in up to 56% of primary human breast carcinomas (Escot *et al.*, 1986; Garcia *et al.*, 1989; Callahan and Campbell, 1989). Significantly, current evidence demonstrates that genetic alterations in c-*myc* may be an important prognostic indicator in

human breast cancer (Mariani-Constantini *et al.*, 1988; Münzel *et al.*, 1991; Berns *et al.*, 1992; Pavelic *et al.*, 1992). In particular, amplification of c-*myc* has been strongly correlated with an early disease recurrence and poor patient survival. An increased frequency of alterations in the c-*myc* gene has also been associated with age, suggesting that it may play a role in the development and/or progression of breast cancer (Escot *et al.*, 1986). Furthermore, the overexpression of c-*myc* in the mammary gland of transgenic mice results in an increased incidence of breast neoplasia (Stewart *et al.*, 1984; Schöenenberger *et al.*, 1988) and a loss of the hormonal dependence of the mammary specific *Wap* and *ß casein* milk protein genes (Schöenenberger *et al.*, 1988).

Evidence for amplification of the c-*myc* gene has been previously observed and confirmed (Figure 25) for the BT 20, MCF 7 and T47 D5 cell lines used in this study (Dubik *et al.*, 1987; Wong and Murphy, 1991). As previously indicated, the nuclease hypersensitivity of c-*myc* chromatin at DH site II_2 in ER^+ breast cancer cells is less than ER^- cells. A lower level of hypersensitivity at DH site II_2 has also been observed in the malignant fibrous histiocytoma cell line, PC3, and was suggested to reflect a compensatory response to c-*myc* amplification (Gibson and Croker, 1992). It is possible that this association is related to the decreased hypersensitivity observed for the ER^+ MCF 7 and T 47 D5 cell lines, both of which have been characterized to carry an amplified c-*myc* gene. However, the observation that the unamplified c-*myc* chromatin of ER^+ ZR 75-1 cells has lower hypersensitivity at DH site II_2 relative to the c-*myc* chromatin of ER^- MDA MB 231 cells (Figure 19) suggests that the previous association can not explain the decrease in nuclease sensitivity at DH site II_2 in all ER^+ breast cancer cell lines.

The regulation of c-*myc* expression is complex and can occur at the levels of transcription initiation (Bentley and Groudine, 1986a), transcript elongation (Bentley and Groudine, 1986b; 1988; Eick and Bornkamm, 1986; Nepveu and Marcu, 1986) and messenger RNA stability (Piechaczyk *et al.*, 1985; 1987). Moreover, the human c-*myc* gene has at least three promoters, and several negative and positive *cis*-acting regulatory sequences, located upstream and downstream of these promoters, are involved in the control of c-*myc* transcription (Figure 14). There are, however, conflicting reports as to whether a particular region exerts a negative or positive effect. Some of these contradictions may be due to differences between the cellular backgrounds used in the studies, while other discrepancies may be attributed to the size of the region analyzed.

Both positive and negative *cis*-acting regulatory DNA sequences have been localized at or near human c-*myc* DH site II_2 (Figure 14). The region -1300 to -300 relative to P_1 was shown to have a positive effect on transcription (Hay *et al.*, 1987), however, multiple negative elements have also been described in this region (Chung *et al.*, 1986; Lang *et al.*, 1988; Whitelaw *et al.*, 1988; Lipp *et al.*, 1987; Remmers *et al.*, 1986; Weisinger *et al.*, 1988; Hay *et al.*, 1987). Although the larger negative control regions include the DH II_2 region, the negative effects could be due to smaller negative elements which have been described. In particular, the mouse sequences -1060 to -923 and -615 to -424, which are equivalent to human sequences -1141 to -1004 and -702 to -511 (Remmers *et al.*, 1986; Weisinger *et al.*, 1988), and the human sequences -608 to -407, -400 to -293 and -353 to -293 have all been described as negative elements (Chung *et al.*, 1986; Hay *et al.*, 1987). The negative element described in the mouse c-*myc* gene, located at -702 to -511 in human c-*myc*, overlaps the -787 to -608 DH II_2 region.

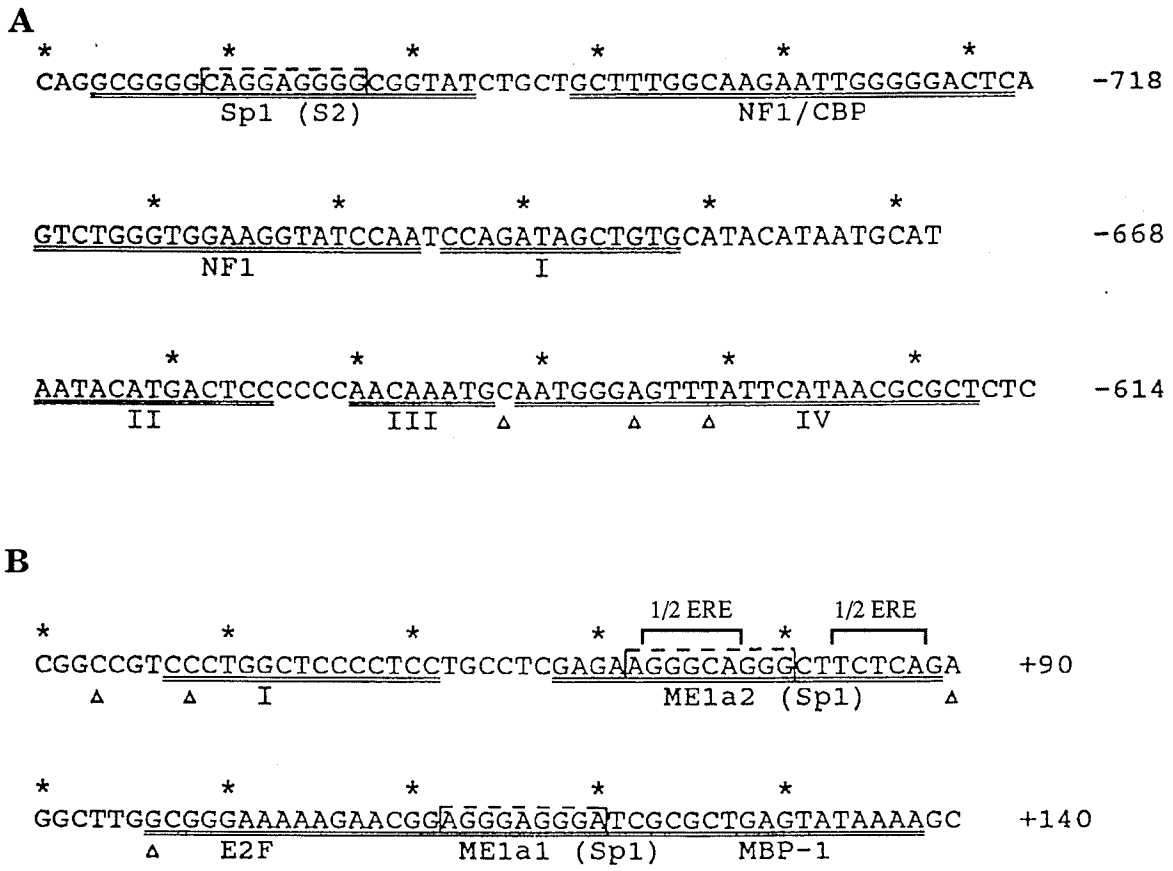
However, the negative function of this region may not be associated with these sequences of human c-*myc*, since sequence identity was observed with only one factor binding site at -595 to -581 (Weisinger *et al.*, 1988). The other negative elements are located 5' and 3' to the -787 to -608 DH II_2 region analyzed in this study (Figures 14 and 27).

Previous *in vitro* DNA binding studies of the c-*myc* gene using nuclear extracts from murine fibroblasts, human epithelial (Lang *et al.*, 1988, 1991; Whitelaw *et al.*, 1988) and human lymphoblastoid cells (Siebenlist *et al* 1984) have identified putative positive and negative *cis*-acting sequences and *trans*-acting factors which bind at the DH site II_2 region (Figure 14). Binding sites for transcription factors Sp1, NF1 and CBP (Figure 27) are located in the 5' portion of the -787 to -608 DH II_2 region (Lang *et al.*, 1988; 1991; Whitelaw *et al.*, 1988). Since different cell lines may have a different collection of *trans*-acting factors to bind and determine the function of *cis*-acting DNA elements, it is important to analyze the function of the various c-*myc* DNA elements in the cell type of interest. The differential sensitivity of DH site II_2 strongly suggests that ER^- breast cancer cells have a complement or activity of *trans*-acting factors, which bind to this region, that is different from those of ER^+ cells. These *trans*-acting factors may be important in the differential regulation of c-*myc* gene expression in human breast cancer cells.

The results of *in vitro* DNA binding studies of the Hae III-Acc I c-*myc* fragment demonstrate that many proteins bind to sequences in the DH site II_2-P_0 region of the c-*myc* gene (Figures 28 and 29). The Stairway and DNase I footprinting assays provide evidence that sequence-specific DNA-binding proteins are located along the entire length of this region *in vitro* (Figures 29, 31 and 32). The MCF 7 and MDA MB 231 nuclear extracts

show similar results to those described for HeLa cell nuclear extracts (Lang *et al.*, 1991) in that the previously described Sp1, NF1 and CBP binding sites were protected from nuclease digestion (Figure 31). Although NF1 and CBP factors did not appear to be responsible for the protein-DNA complexes detected by EMSA with the -787 to -670 c-*myc* region, Sp1-like interactions were apparent in this assay (Figure 30). The interaction of other proteins, such as MAZ/ZF87 (Bossone *et al.*, 1992; Pyrc *et al.*, 1992), ETF-1 (Kageyama *et al.*, 1988) or GCF (Kageyama and Pastan, 1989), which are reported to bind G+C-rich elements, can not be excluded and it is possible that competing interactions between these factors are involved in the regulation of c-*myc* transcription in human breast cancer cells.

In addition to the previously identified Sp1, NF1/CBP and NF1 factors, the interaction of other proteins with the c-*myc* region -694 to -607 relative to P_1 is apparent (Figure 32). Sequence-specific DNA-binding proteins interacting with the -694 to -607 5' flanking region of the human c-*myc* gene have not been previously reported. Using DNase I footprinting, four protected regions were detected, suggesting at least four *trans*-acting factors can bind to this region (Figure 32). Footprint I (-694 to -682) is located between the Asp I and Nsi I sites and contains the Alu I site (Figures 27 and 44). This protected region contains the sequence 5'-AGATAG-3', which is similar to the binding site for the GATA family of transcription factors (Faist and Meyer, 1992). The other three protected regions are located between -670 to -617, relative to P_1. Footprint IV, which contains the sequence 5'-TTATTCAT-3', shares sequence similarity with the consensus site for Pit-1/GHF-1/GHF-5/GHF-7. These factors are involved in activation of the growth hormone and prolactin genes whose regulatory regions contain multiple copies of this consensus sequence (Nelson *et al.*, 1986; Nelson, *et al.*,



C

Oligonucleotide	Complexes	Competitor
O-II	7	ND
O-II/III	3(a or b), 4, 5, 6	O-II/III, O-III/IV, O-IV
O-III/IV	1a, 1b, 1c, 3(a or b), 4, 6	O-II/III, O-III/IV, O-IV
O-IV	1c, 2, 6	ND

Figure 44. Summary of New Protein Binding Sites Identified in c-*myc* Regulatory Regions. The c-*myc* DH II_2 region from -771 to -614 (**A**) and the c-*myc* DH III_2 region from +40 to +140 (**B**), relative to P_1, is shown. Double underlines indicate the DNase I protected sequences and arrowheads indicate the *in vitro* DNase I hypersensitive sites. Roman numerals in **A** and **B** correspond to those shown in Figures 32 and 43, respectively. The putative 1/2 ERE sites are as described in Figure 16. (**C**) A summary of the protein interactions with the various footprints of theDH II_2 region based on EMSA data and cross competitions. ND indicates competitions were not done.

1988; Bodner *et al.*, 1988; Schaufele *et al.*, 1990). Although these factors have been described to require two sites of this sequence (Schaufele *et al.*, 1990), only one site is present in this region of the c-*myc* gene.

This same site, 5'-TTATTCA-3' and a second site 5'-TGACTCC-3', which maps to Footprint II, were identified using the IBI MacVector program as possible AP-1 sites. However, competition studies with consensus AP-1 oligonucleotides suggest that AP-1-like factors do not bind this region (Figure 38). The sequence 5'-TGACTC-3', in Footprint II, is also identical to the consensus sequence for GCRE, a GCN4-like factor (Hope and Strühl, 1985). However, this factor requires nine deoxy-thymidine residues downstream from the binding site for optimal binding and these residues are not present in this region of c-*myc*. There are two repeats of the sequence CCCTC, one of which lies between Footprints II and III and a second which is located within Footprint IV. These sites correspond to part of the binding site for CTCF (Lobanenkov *et al.*, 1990), which binds to three repeats of the motif CCCTC spaced 5-6 bp apart in the chicken c-*myc* gene. In the DH II_2 region of human c-*myc*, these two repeats lie 15 bases apart and are on opposite strands, therefore it is unlikely that this factor is responsible for generating these protected regions. No other sequence similarities were identified in this c-*myc* region, therefore it appears that some of the proteins interacting with this region of c-*myc* are novel and uncharacterized.

Oligonucleotides that contained one or more of the protected regions II, III and IV were used in EMSA to assign the protein-DNA complexes to the various footprints (for summary see Figure 44). Similar interactions with the -670 to -607 c-*myc* region were observed with nuclear extracts form MCF 7 and MDA MB 231 human breast cancer cells (Figures 34-37). However, differences in the abundance and/or activity of three protein-DNA complexes

were apparent between the two nuclear extracts. Proteins generating complex 3 (a or b) and 1c were relatively higher with nuclear extracts from the ER^- MDA MB 231 cells, while the level of factor responsible for complex 1b was relatively higher in ER^+ MCF 7 cells. The generation of complexes 1b and 1c and the associated differences in abundance and/or activity were apparent with the 63 bp Nsi I-Acc I fragment and the O-III/IV oligonucleotide (Figure 34). Footprint regions III and IV were present in both of these DNA sequences, suggesting a requirement for one or more of these elements. Similarly, complex 3 (a or b) was formed on the 63 bp Nsi I-Acc I fragment, the O-II/III and the O-III/IV oligonucleotides, all of which contained the Footprint III sequences (Figures 34-37). A higher abundance and/or activity of both complex 3a and 3b was apparent in MDA MB 231 nuclear extracts using the Stairway assay (Figure 29). However, complexes 3a and 3b were unresolved in subsequent studies (Figures 34-37). It is therefore possible that these two complexes could represent separate interactions with the O-II/III and O-III/IV oligonucleotides.

The integrity of the Footprint III sequences of c-*myc* was also required for the generation of most of the other protein-DNA complexes corresponding to DH II_2. The presence of this protected region was required for the formation of complexes 1a, 2, 4, 5 and 6, in addition to its integral role in generating complexes 3 (a or b) and 1b (Figures 34-37). Conversely, complexes 7 and 1c were generated in the absence of this element (Figure 34). These results suggest there is a high level of interaction between proteins associated with these sequences of the c-*myc* gene. These interactions may be due to different factors, which bind to overlapping sites, or may be the result of sequence-specific DNA-binding proteins that interact with different non-DNA binding proteins (e.g. co-activators) to yield different mobilities by

EMSA but similar protected regions with DNase I footprinting. Significantly, the faster migrating complexes (1a, 1b, 1c) formed on the O-III/IV oligonucleotide and the 63 bp Nsi I-Acc I fragment show similar intensities of binding, but the slower migrating complexes (2, 3a, 3b, 4, 5, 6, 7) are less intense on the oligonucleotides (Figure 34). This observation and the ability of Footprint III and IV sequences to decrease the binding of the slower migrating complexes on both the O-II/III and O-III/IV oligonucleotides (Figures 35-37) supports the suggestion that the slower migrating complexes may be generated by interactions with non-DNA-bound proteins.

Coincident with the observed differences in factor abundance and/or activity at Footprints III and IV was the detection of three differential *in vitro* DNase I hypersensitive sites which mapped to these protected regions (Figure 32). Nuclear extracts from ER^- MDA MB 231 cells generated *in vitro* hypersensitive sites 1 and 2, while ER^+ MCF 7 nuclear extracts were responsible for the formation of *in vitro* hypersensitive site 3. The first of these sites was located between the Footprint III and IV regions, while the other two sites were located within Footprint IV (Figure 44). These differences in *in vitro* DNase I hypersensitivity reflect differences in the protein contacts within this region of the c-*myc* gene between the ER^- and ER^+ breast cancer nuclear extracts. These data support the hypothesis that the differential *in vivo* DNase I hypersensitivity detected at DH site II_2 between ER^- MDA MB 231 and ER^+ MCF 7 breast cancer cells may be the result of cellular differences in the complement or activity of factors interacting with the c-*myc* chromatin spanning this regulatory region. Notably, none of the Sp1 and NF1 sites mapped previously to the c-*myc* DH II_2 region can account for the differential interactions of factors associated with the -670 to -607 c-*myc* sequences (Figure 44). Since Sp1

binding is unaffected by methylation (Harrington *et al.*, 1988), the methylation sensitivity of protein interactions in this region suggests that the generation of these protected regions are not the result of Sp1 binding to these sites (Figure 39).

The most common methylated nucleotide *in vivo* is 5-methylcytosine, which occurs largely in the dinucleotide sequence mCpG. These modified nucleotides are found in the 5' regions of genes (Bird, 1986) and enhance the probability that a stretch of DNA will assume a Z configuration (Behe and Felsenfeld, 1981). The observation that the c-*myc* P_0 region forms Z DNA during active transcription, which is lost during differentiation (Wittig *et al.*, 1992), suggests that differential methylation in the DH II_2 region may be involved in the formation of differential hypersensitivity at this site in breast cancer cells. The formation of Z DNA has not been detected near the P_1 and P_2 promoters of the normal human c-*myc* gene (Wittig *et al.*, 1992) however, a methylated GC-rich island covering exons 1 and 2 has been reported for the truncated but still active c-*myc* gene (Bianchi *et al.*, 1989). Conversely, hypomethylation at CpG sequences is observed during the cellular progression to neoplasia (Jones and Buckley, 1990), and the description of an age associated change in the methylation status of c-*myc* may contribute to tissue-specific changes in expression (Ono *et al.*, 1989).

An analysis of the 5' flanking region of the c-*myc* gene in MCF 7 and MDA MB 231 cell lines using methylation sensitive restriction endonucleases, found no detectable difference in DNA methylation patterns (Figure 26). Isoschizomeric restriction endonucleases however, detect only about 4% of methylcytosines. Therefore, a site specific method, such as genomic sequencing, is a far superior method for DNA methylation analysis (Saluz and Jost, 1989). Furthermore, many of these enzymes are incapable of

distinguishing between hemimethylated and fully methylated sites (Gross and Garrard, 1988). It is therefore possible that the method used in these studies was unable to detect a critical methylation site which may be involved in determining the differential accessibility of the chromatin to nucleases. Significantly, DNA methylation suppresses DH site formation (Gross and Garrard, 1988), while hypomethylation creates a permissive environment for establishing hypersensitive sites (Groudine *et al.*, 1981; Saluz *et al.*, 1988; Adams, 1990).

Decreased CpG methylation at many loci *in vivo* also correlates with increased gene expression, perhaps due to the increased access of regulatory *trans*-activators to DNA elements (Cedar, 1988). DNA methylation has previously been implicated in the disruption of sequence specific protein interactions (Prendergast and Ziff, 1991), therefore differences in the methylation pattern of c-*myc* DNA could potentially explain the differential hypersensitivity and regulation of this gene in breast cancer cell lines if the modification altered a critical binding site in the 5' flanking region of the gene. Significantly, in the region -787 to -607, relative to P_1, which contains the DH II_2 site and the P_0 promoter, there are five CpG sites which could potentially be methylated. One of these CpG sites was found to be unmethylated (Figure 26; Hha I at -618) but a second CpG which would not be detected by this analysis is immediately 5' to this restriction site in the Footprint IV region (Figure 44). Although this uncharacterized CpG site is not near the *in vitro* hypersensitive sites identified in Footprint IV, methylation here could still interfere with protein interactions. Two other potential methylation sites are present in the Sp1 binding sequence. Although Sp1 binding is unaffected by methylation (Harrington *et al.*, 1988), the presence of 5-methylcytosine in this site may disrupt the binding of other

factors which recognize G+C rich sites. A fifth CpG is located upstream from the Sp1 site at -781, relative to P_1. Protein interactions near this site have not been reported and were not characterized in this study. Therefore, at this stage it would be premature to rule out the possibility that methylation is involved in generating the differential hypersensitivity at DH site II_2.

The P_2 promoter is the preferred promoter of human breast cancer cells (Figure 24). Analysis of nuclear proteins from MCF 7 and MDA MB 231 human breast cancer cells by EMSA and DNase I footprinting detects several proteins which interact with DNA sequences localized between +67 to +140 (Figures 40 and 43). This region corresponds to the DH III_2 region, where binding sites for ME1a2, E2F, ME1a1 and MBP-1 have been previously identified (Figures 15 and 27). The *cis*-elements for ME1a2 and ME1a1 are thought to have an important role in regulating transcription from the P_2 promoter (Asselin *et al.*, 1989; DesJardins and Hay, 1993; Hall 1990; Hiebert *et al.*, 1989; Lipp *et al.*, 1989; Moberg *et al.*, 1991, 1992; Plet *et al.*, 1992; Thalmeier *et al.*, 1989). Synthetic oligonucleotides to ME1a1 and ME1a2 were able to compete for these complexes (Figure 41), however interactions due to the E2F and MBP-1 binding sites were not apparent. Significantly, the complexes formed with ME1a2 and ME1a1 were identical with those formed with the Sp1-like binding site (S2) located in the DH site II_2 region (Figure 42A). Moreover, the pattern of complexes formed was typical of that formed by the interaction of Sp1 with other Sp1 *cis*-elements (Sun *et al.*, 1992). A comparison of S2, ME1a1 and ME1a2 sequences revealed sequence similarities with other Sp1-like binding sites (Figure 42B). In addition, it has recently been reported that Sp1 can bind to the ME1a1 site (DesJardins and Hay, 1993). These data suggest that Sp1-like factors can

interact with sequences thought to be of major importance in regulating transcription of the c-*myc* gene from the P_2 promoter.

Different intensities of the protein-DNA complexes interacting with the -1 to +67 c-*myc* region occurred between MCF 7 and MDA MB 231 nuclear extracts (Figure 40). DNase I footprint analysis identified a protected region at +48 to +62 (Footprint I) which is 5' of both the ME1a2 region at +71 and the Xho I site at +67 (Figure 27 and 43). While similar footprints were found using MCF 7 and MDA MB 231 nuclear extracts, distinct *in vitro* hypersensitive sites were generated near Footprint I between the two cell lines (Figure 43). DNA binding sites in this region of the c-*myc* gene have not previously been reported (Figures 15 and 27) and therefore, represent potentially novel *trans*-acting factor binding sites for c-*myc*. The protected DNA in this region has sequence similarity to CTCF (Lobanenkov *et al.*, 1990), which binds to three repeats of the motif CCCTC spaced 5-6 bp apart in the chicken c-*myc* gene. Within Footprint I of the human c-*myc* DH III_2 region (Figure 44), two similar repeats, 5'-T<u>CCCTG</u>GCTC<u>CCCTC</u>C-3', are found. However, only the second of the potential CTCCC motifs is conserved between human, mouse and rat (Morberg *et al.* 1991), suggesting this factor is unlikely to associate with this region. This motif also resembles the CT elements described to be Sp1 sites (DesJardins and Hay, 1993), suggesting Sp1-like factors may be interacting here. These data, together with the DNase I footprinting data, imply that previously unidentified transcription factors and their corresponding *cis*-acting elements may be involved in regulating transcription of the c-*myc* gene in human breast cancer cells.

An obvious question concerns the role of the ER in the differential hypersensitivity of DH site II_2 in these breast cancer cells. It was shown

previously (Dubik *et al.*, 1987) that c-*myc* expression is estrogen responsive in MCF 7 and T 47D cells. In these experiments estrogen was used to rescue mono-hydroxytamoxifen treated cells. However, the ER liganded with mono-hydroxytamoxifen can still interact *in vitro* with estrogen responsive elements (Fawell *et al.*, 1990). Tamoxifen is also not a pure estrogen antagonist, having partial estrogen agonist activity (Jordan and Murphy, 1990). This makes it likely that the basal level of c-*myc* transcription in the absence of estrogen may be significantly lower than that previously reported. Additionally, the antiestrogen nafoxidine is capable of interacting with an ER-VP16 chimeric protein and inducing an altered chromatin structure of a yeast reporter gene containing two EREs (Pham *et al.*, 1991a). The effect of estrogen on the DH pattern in MCF 7 cells was therefore investigated by treatment of cells which had been maintained under essentially estrogen free conditions (phenol red free, twice charcoal/dextran stripped fetal bovine serum).

As expected, RNA analysis demonstrates that 10 nM 17ß-estradiol markedly increases the steady state level of c-*myc* mRNA in cells cultured under these conditions (Figure 17). A 7 fold increase in c-*myc* expression above estrogen deprived cells is observed at 30 and 60 minutes after treatment. This observation agrees with those previously published (Dubik *et al.*, 1987). An alteration in chromatin structure associated with gene activation is expected to occur prior to the observation of increased mRNA levels, therefore, estrogen induced changes in nuclease sensitivity were analyzed up to 30 minutes after treatment. The relative hypersensitivity of DH site II_2 in the c-*myc* chromatin of ER^+ MCF 7 cells grown in estrogen depleted conditions remained low compared to ER^- MDA MB 231 cells grown under the same conditions (Figure 21). Additionally, the DNase I

hypersensitivity of DH site II_2 in the c-*myc* chromatin of MCF 7 cells was unaffected by estrogen addition (Figure 22). The relative promoter usage remains unaltered under these conditions suggesting that increased use of the P_0 promoter is not responsible for these changes (data not shown). Furthermore, the reintroduction of the ER does not reconstitute the ER^+ c-*myc* chromatin structure, although selective progression to the ER^- phenotype appears to coincide with an increase in DNase I hypersensitivity at DH site II_2 (Figure 23).

Although nucleosome positions for c-*myc* chromatin have not been determined in this study, the region between the P_0 and P_1 promoters of c-*myc* is proposed to resist packaging into nucleosomes, while an ordered chromatin structure, possibly due to positioned nucleosomes is evident 5' of the P_0 promoter (Kumar and Leffak, 1989). These regions remained hypersensitive when transduced into a new site in the genome, suggesting the underlying sequence may be important for organizing the c-*myc* chromatin structure (Kumar and Leffak, 1991). Some of the DH sites in active genes are maintained even if the gene is infrequently transcribed (Svaren and Chalkly, 1990). Significantly, all of the c-*myc* DH sites that are observed in actively transcribed cells are maintained in quiescent cells which have little transcription but are still capable of being induced (Siebenlist *et al.*, 1988; Spencer and Groudine, 1991). These same sites are lost when the c-*myc* gene is repressed, such as in the nontranslocated allele of BL cells (Siebenlist *et al.*, 1984, 1988; Kakkis *et al.*, 1986; Dyson and Rabbitts, 1985).

The treatment of HL60 cells with DMSO (Dyson *et al.*, 1985; Siebenlist *et al.*, 1988), dibutyryl cyclic AMP (Siebenlist *et al.*, 1988) or retinoic acid (Bentley and Groudine, 1986a) induces their differentiation and results in the disappearance of DH site II_2. Expression of c-*myc* is markedly reduced

under these conditions and is associated with increased nuclease sensitivity at DH site IVb (Bentley and Groudine, 1986a; Chung *et al.*, 1986). This DH site is located in intron 1 and maps to a block in c-*myc* transcriptional elongation (Bentley and Groudine, 1986a). In ER^+ and ER^- breast cancer cells, DH site II_2 is differentially sensitive to nuclease treatment, however, DH site IVb is weakly sensitive in both cell types (Figures 19, 20 and 23). Increased nuclease sensitivity at DH site IVb is not readily apparent in breast cancer cells grown under estrogen depleted conditions (Figure 21). As well, estrogen treatment of MCF 7 cells grown under these conditions does not contribute to differences in nuclease attack at this site (Figure 22). These observations suggest that decreased c-*myc* gene transcription in estrogen depleted cells is not primarily due to attenuation of transcript elongation. Similarly, although some reduction may occur, nuclease sensitivity at DH site II_2 is not lost upon estrogen withdrawal (Figures 21 and 22), suggesting that c-*myc* transcriptional competence is maintained in breast cancer cells grown in estrogen depleted conditions. This is consistent with the observation that MCF 7 cells appear to proliferate, although much more slowly, under these conditions.

Estrogen withdrawal causes the regression of ER^+ MCF 7 tumors grown in nude mice. This was characterized as a cytotoxic effect which results in programmed cell death (Kyprianou *et al.*, 1991). Significantly, the overexpression of c-*myc* in MCF 7 cells results in slower *in vitro* growth rates (Liu *et al.*, 1989), which may also reflect an increase in cell death. The determination that the c-*myc* proto-oncogene is involved in regulating apoptosis (Ffrench-Constant, 1992; Evan *et al.*, 1992), makes it a prime target for alterations that would suppress this function but enhance its proliferative function. The outgrowth of estrogen independent but responsive breast

cancer sublines of MCF 7 cells grown as tumors in nude mice for extended periods in estrogen depleted conditions (Brünner *et al.*, 1993) and the observation that most breast tumors eventually progress to a hormonally independent phenotype (McGuire *et al.*, 1975; McGuire, 1980) supports this possibility. The *bcl-2* gene, which codes for a putative G protein (Haldar *et al.*, 1989), can suppress c-Myc induced apoptosis but does not effect the proliferative effects of c-Myc (Bissonnette *et al.*, 1992; Fanidi *et al.*, 1992). Although no data are available on the effects of extended growth in the presence of antiestrogens, the growth of MCF 7 cells in the presence of adriamycin results in drug resistance, loss of ER expression (Vickers *et al.*, 1988) and enhanced hypersensitivity at DH site II_2 of c-*myc* (Figure 23). Possibly chemotherapeutic agents induce the *bcl-2* signaling pathway. This may indirectly alter the regulatory proteins involved in the induction of c-*myc* expression, which are reflected as differences in hypersensitivity at DH site II_2, and promote the development of hormone resistant breast cancer cells.

Although estrogen does not appear to alter c-*myc* chromatin structure, estrogen inducible hypersensitive sites are observed in the presence of palindromic EREs (reviewed in Gross and Garrard, 1988; Elgin, 1988) or multiple ERE half sites (Kaye *et al.*, 1986; Kato *et al.*, 1992b). As well, estrogen induced changes in the chromatin structure of chromosomally integrated gene constructs (Seyfred and Gorski, 1990; Pham *et al.*, 1991; Gilbert *et al.*, 1992) or reconstituted templates (Schild *et al.*, 1993) have also been reported. In constructs containing more than one consensus ERE estrogen induces dramatic alterations in chromatin structure, however, in the presence of only one ERE changes in chromatin structure are barely detectable (Pham *et al.*, 1991; Gilbert *et al.*, 1992). These observations have been suggested to reflect the binding of one receptor dimer to a nucleosomal

ERE without disrupting nucleosome positioning (Pham *et al.*, 1991). This is consistent with the observation that the interaction of the GR with a GRE positioned within a nucleosome induces only transient (Reik *et al.*, 1991) or minor changes in nucleosome structure (Perlmann and Wrange, 1988). Alternatively, limited chromatin alterations in the presence of only one ERE could also be consistent with the presence of a positioned nucleosome between the ERE and the promoter, as has been observed for the *Xenopus* vitellogenin B1 gene (Schild *et al.*, 1993).

Interestingly, even in the presence of multiple EREs the interaction of receptor with DNA is suggested to be a relatively weak or transient event, since protection of ERE guanine residues is not observed with *in vivo* DMS footprinting (Gilbert *et al.*, 1992). Hormone dependent *in vitro* DMS protection (Becker *et al.*, 1986) and DNase I footprinting (Perlmann and Wrange, 1988) has been observed for the GRE, however another group has shown this same region to be protected by a protein interacting with an overlapping binding site (Rigaud *et al.*, 1991). Similarly, the promoter proximal half-site ERE found in the ovalbumin gene can also bind AP1 proteins (Gaub *et al.*, 1990) and can be replaced by the binding sites of other *trans*-activators, including Sp1 (Kato *et al.*, 1992b). These combined observations support the hypothesis that receptor interactions with chromatin elements may recruit other *trans*-acting factors to these sites in an attempt to establish active transcription complexes.

The co-transfection of a c-*myc*-*CAT* reporter gene and an ER expression vector into HeLa cells has been used to identify regulatory regions responsive to estrogen stimulation (Dubik and Shiu, 1992; Dubik, 1991). The results suggest that the highest level of estrogen response (9.2 fold) requires the integrity of sequences from -667 to +141 of the c-*myc* gene (Figure 16). The

addition of sequences 5' of -667 or 3' of +141 reduces this effect such that sequences from -607 to +202, -1255 to +202 and -2327 to +202 result in 5.1, 4.6 and 5.7 fold, respectively. From these studies it is difficult to determine the specific effect of the DH II_2 region, but it appears to enhance the estrogen activation of c-*myc*. The critical region for estrogen activity was determined to be located within +25 to +202, relative to P_1, which supported a 6.2 fold response to estrogen. The DNA-binding domain of the ER was essential for this estrogen regulated c-*myc-CAT* expression.

The removal of sequences from +67 to +141 negated any estrogen effect (Figure 16). This 116 bp c-*myc* region contains the c-*myc* P_2 promoter region and binding sites for ME1a2, E2F, ME1a1, MBP-1, and E2F (Figure 27). Although binding of an *in vitro* transcribed and translated ER has not been detected at this site (Dubik and Shiu, personal communication), crude nuclear extracts isolated from breast cancer cells produce protein-DNA complexes with sequences encompassing this region (Figures 40, 41 and 43). In particular, the sites for ME1a1 and ME1a2 appear to be interacting with Sp1-like proteins (Figure 42). Others have also suggested that Sp1 interacts with these sequences (Asselin *et al.*, 1989; Hall, 1990; Moberg *et al.*, 1991; 1992; DesJardins and Hay, 1993). The ME1a1 element has been described as inhibitory to P_1 transcription, but stimulatory for P_2 transcription, while the ME1a2 element has been characterized as necessary for optimal transcription (Asselin *et al.*, 1989; Moberg *et al.*, 1991; DesJardins and Hay, 1993). Additionally, the ME1a1 element has been identified as a regulator of transcript elongation from both P_1 and P_2 (Miller *et al.*, 1989b, Wright *et al.*, 1991; Bossone *et al.*, 1992; Dufort *et al.*, 1993). Thus the c-*myc-CAT* construct lacking sequences +67 to +141 would delete this critically important regulatory region.

The construct which contains the internal deletion of -101 to +25 also results in a significant decrease in estrogen stimulated c-*myc*-*CAT* activity, to only 3.1 fold (Figure 16). This region overlaps with one of the CT-elements located between -157 to -97, which have also been shown to bind Sp1. The integrity of these five DNA elements is necessary for initiation from P_1 and required for maximal P_2 transcription (DesJardines and Hay, 1993). These authors have also demonstrated that the ratio of transcripts initiating from P_1 and P_2 can be modulated by the disruption of these elements. The presence and integrity of Sp1-like sites near the strong P_2 promoter (Me1a2 and Me1a1/CT-I_2) and proximal to the weaker but still active P_1 promoter (CT-elements 1 to 5) is therefore essential to estrogen induced expression of c-*myc*.

A classical estrogen responsive element (ERE) is not located in the c-*myc* DH II_2 region or in any of the c-*myc* sequences from -2327 to +5743 relative to P_1 (Dubik and Shiu, 1992; Dubik and Shiu, personal communication). However, sequences similar to half site EREs have been identified by computer analysis 3' of P_1 (data not shown). One of these sites is located within the DH site III_2 sequences which have been determined to be necessary for estrogen activation of c-*myc* (Dubik and Shiu, 1992). Neither an ERE nor a 5'-GGTCA-3' half-palindromic motif is present in this region, but this region does contain the sequence 5'-GGGCA-3' which is similar to the classical ERE half-site (Figures 3 and 16). However, this putative site overlaps with the ME1a2 binding site and therefore may have to compete with Sp1 for binding. Another ERE-like half site is located immediately 3' to the ME1a2 site and contains the sequence 5'-TCTCA-3' (Figure 16). Together these elements make up a nonpalindromic ERE with the sequence, 5'-GGGCA-n_5-TCTCA-3', which conforms to the lower strand of the consensus

sequence previously described (Figure 3C). Nuclease protection of these putative ERE-like sequences and the ME1a2 binding site is seen with nuclear extracts from both ER^+ MCF 7 and ER^- MDA MB 231 breast cancer cell lines (Figure 43 and 44).

Binding of ER monomers, has been observed both *in vivo* (Gorski *et al.*, 1993) and *in vitro* (Medici *et al.*, 1991). This interaction is not dependent on a palindromic ERE and has been observed when the dyad symmetry of the ERE is disrupted (Medici *et al.*, 1991). DNA binding of ER monomers is not dependent on ligand interaction (Gorski *et al.*, 1993; Brown and Sharp, 1990) however, the *trans*-activation function of the hormone binding domain (TAF-2) does require estrogen binding (Kumar *et al.*, 1987; Webster *et al.*, 1988). The function of hormone binding has therefore been suggested to increase the interaction of the ER with transcription factors or other chromatin bound proteins (Gorski *et al.*, 1993). The final result of these associations would be an increase in the transcriptional activity of estrogen responsive genes.

The interaction of ER monomers with multiple half site EREs in the chicken ovalbumin gene has been suggested as one mechanism for estrogen action in the absence of a palindromic ERE (Kato *et al.*, 1992b). Conversely, the interaction of Sp1 with an element 3' of an ERE half site in the creatine kinase B gene has been proposed as an alternate mechanism for stabilizing monomeric ER binding and inducing the estrogen response (Wu *et al.*, 1991). Similarly, Sp1 or Sp1-like interactions at the ME1a2 element (Asselin *et al.*, 1989; Hall 1990; Moberg *et al.*, 1991; 1992; DesJardins and Hay, 1993; see also Figure 42) or at an additional Sp1 site slightly downstream (Nishikura, 1986) could be involved in stabilizing monomeric interactions of the ER with the putative ERE half sites in the human c-*myc* gene. Alternatively, the

novel protein binding sites identified at DH sites II_2 and III_2 in this study may be involved in stabilizing the ER at these half sites in the c-*myc* chromatin of hormone responsive breast cancer cells. The observation that a maximal estrogen response (9.2 fold) requires c-*myc* sequences from -670 to +140 (Figure 16), which contains the novel protein binding sites identified at DH sites II_2 and III_2 in addition to the Sp1 sites previously characterized for the DH III_1 and III_2 regions (Figure 27 and 44), supports a role for these factors in facilitating ER action.

The establishment of a preinitiation complex by the basic transcription factors is thought to be stabilized by the interaction of proteins associated with distal enhancers and promoter proximal elements. These interactions facilitate the formation of an active transcription complex and increase transcript initiation. Interactions between *trans*-acting factors, effector proteins and the general transcription complex are therefore important for the regulation of transcriptional activity (Martin, 1991; Ham *et al.*, 1992; Herschlag and Johnson, 1993). A role for intermediary factors in the *trans*-activation function of steroid receptors, including the ER has previously been suggested (Bocquel *et al.*, 1989). Similarly, the interaction of nuclear receptors with intermediary factors found in the TFIID fraction has been reported for the PR (Shemshedini *et al.*, 1992). The overexpression of c-Jun and c-Fos in MCF 7 cells inhibits ER activity (Doucas *et al.*, 1991). Similarly, these *trans*-activators have been characterized as regulators of the activation and suppression functions of the GR by an interaction of these proteins at overlapping or adjacent sites (Diamond *et al.*, 1990). AP1 elements are not present near the putative ERE half-sites of the human c-*myc* gene, although an overlapping AP1/ERE site is present upstream of the c-*fos* gene (Weisz and Rosales, 1990). There is however, an AP1 site 5' of c-*myc* DH site

III_1(Figure 27), which is associated with a negative control element (Hay *et al.*, 1989; Takimoto *et al.*, 1989). The role of this region in the estrogen regulation of c-*myc* has not been determined.

Transcriptional synergy between steroid receptors and Sp1 has been demonstrated with constructed elements (Schüle *et al.*, 1988a; 1988b; Strähle *et al.*, 1988) and the degree of synergy between the GR and other *trans*-activators is suggested to be inversely related to the strength of the GRE (Schüle *et al.*, 1988a). One method of determining if Sp1 and ER were acting synergistically would be to use antibodies to these factors to disrupt *in vitro* protein interactions with this region of c-*myc*. Alternatively, the combined interactions of purified Sp1 and ER with oligonucleotides to this region and the disruption of binding with site specific mutations would provide further useful information. However, it may be difficult to detect if the ER is interacting directly with either of the c-*myc* ERE-like half elements *in vivo*, since *in vivo* genomic footprinting of the mouse c-*myc* gene suggests that protein interactions at the ME1a2 site are lost only when the gene is silent or truncated (Plet *et al.*, 1992). This observation, as well as the analysis of c-*myc* chromatin (Figures 21 and 22), which failed to detect an estrogen responsive DH site, suggest that the ER may interact with other *trans*-acting factors to potentiate its function.

Sequences that generate DH sites may induce or suppress the formation of other DH sites even over long distances (Gross and Garrard, 1988), suggesting that the DNase I hypersensitive sites of c-*myc* may influence the nuclease sensitivity of each other. This current study provides evidence that Sp1-like proteins interact with ME1a1 and ME1a2 in human breast cancer cells and confirms the presence of Sp1-like binding sites in the DH site II_2 region of the human c-*myc* gene. Others have described the

interaction of Sp1 with the ME1a1 site (DesJardins and Hay, 1993) and have described five additional Sp1 sites at -157 to -97 within the DH III_1 region. The interaction of Sp1-like proteins located at DH site II_2 with Sp1 bound at multiple *cis*-acting elements in the DH III_1 and III_2 regions, provides a mechanism for the association of other *trans*-acting proteins located within these three elements (Figure 45).

Specific DNA-protein complexes at these DH sites may be maintained on the c-*myc* gene regardless of the transcriptional state and these established structures may be required for the rapid response of this gene to hormonal stimulation. These interactions may stabilize ER monomers at their putative binding sites in c-*myc* and facilitate an association with the P_2 promoter region in hormone-dependent breast cancer cells. Significantly, interactions between distal and promoter proximal regulatory regions are required for estrogen activation of the prolactin gene (Seyfred and Gorski, 1990). The absence of this receptor in hormone-independent breast cancer cells suggests that other factors are involved in the constitutive expression of c-*myc*. The differential interactions of nuclear proteins from MCF 7 and MDA MB 231 breast cancer cells with previously uncharacterized regulatory elements in the DH II_2 and III_2 regions may be involved in this altered regulation and may influence the intensity of nuclease sensitivity at DH site II_2.

Cooperative interactions of Sp1 multimers bound at distal and promoter proximal sites have been previously characterized (Courey *et al.*, 1989). These associations require the integrity of the Sp1 *trans*-activation domains (Pascal and Tjian, 1991) and the presence of at least one promoter proximal site for transcriptional activation (Seipel *et al.*, 1992). The interaction of Sp1 multimers at separated sites induces chromatin loop

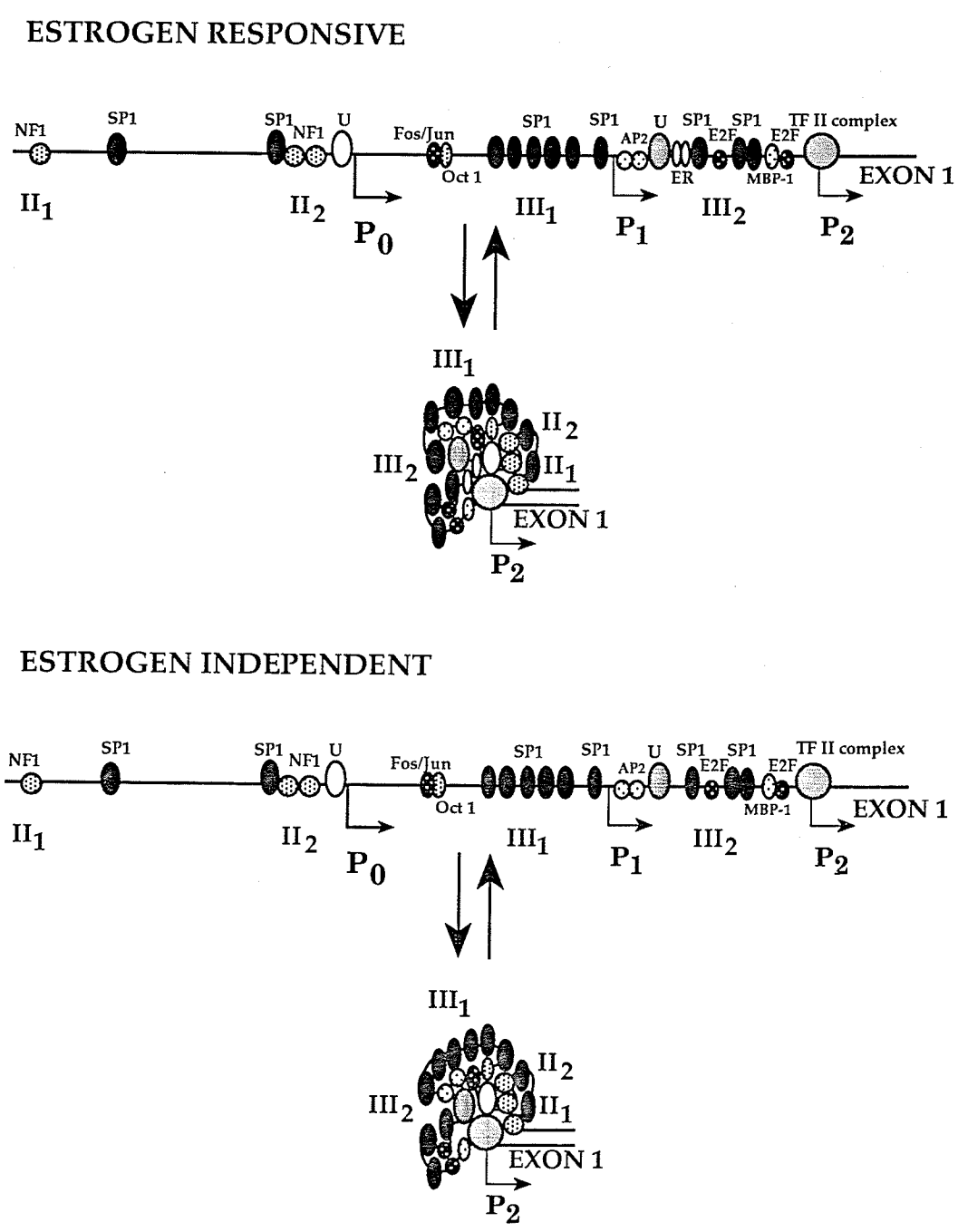


Figure 45. Model for Transcriptionally Active c-*myc* Chromatin. Multiple Sp1 binding sites in the 5' flanking sequences of the c-*myc* gene are located at DH sites II_1, II_2, III_1 and III_2. The interaction of Sp1 multimers allows for the association of these regulatory regions with each other and with the transcriptional machinery. The association of the ER with the DH III_2 region of responsive c-*myc* chromatin plays an integral role in estrogen regulation of this gene, while novel proteins (U) associated with the DH II_2 and III_2 regulatory regions may allow for the constitutive expression of c-*myc* in estrogen independent cells.

formation, prevents H1 mediated repression and enhances the formation of the preinitiation complex (Laybourn and Kadonaga, 1991; Li *et al.*, 1991; Su *et al.*, 1991; Hoey *et al.*, 1993). Additionally, Sp1 proteins have been identified in both nuclear matrix and nonmatrix protein fractions (van Wijnen *et al.*, 1993), suggesting that this *trans*-acting factor may be responsible for maintaining the transcriptionally competent c-*myc* chromatin conformation.

Support for this suggestion comes from the observation that the transcriptionally active c-*myc* allele in BL cells associates with the nuclear matrix, while the repressed allele does not (Andreeva *et al.*, 1992). Furthermore, this association is through the interaction of nuclear proteins at multiple attachment sites, arranged throughout the entire gene locus. The estrogen receptor has also been localized within the nucleus to the nuclear matrix component (Barrack and Coffey, 1980; Alexander *et al.*, 1987). This affiliation is both steroid and tissue specific and can be modulated by the hormonal state of the animal. Significantly, ligand interaction with the ER results in a tighter association with the nucleus and conformational changes in the receptor, which may lead to an association with other chromatin associated proteins (Gorski *et al*, 1993). The PR and the transcription factor ATF both facilitate the formation of a stable preinitiation complex (Klein-Hitpass *et al.*, 1990; Horikoshi *et al.*, 1988) and significantly ATF-like proteins also associate with the nuclear matrix (Dwortzky *et al.*, 1992). Possibly the interaction of the ER with the c-*myc* gene chromatin of hormone responsive breast cancer cells also stabilizes the preinitiation complex and alters the interactions of this domain with the nuclear matrix. These interactions may in turn influence protein associations with upstream

elements of c-*myc* which result in decreased DNase I hypersensitivity at the DH II_2 region in these cells.

Summary

The expression of the c-*myc* proto-oncogene is estrogen regulated in ER positive, hormone-dependent human breast cancer cells, but it is constitutively active in ER negative, hormone-independent human breast cancer cells. Furthermore, it appears that mere expression of the ER in an originally ER^- cell line does not reconstitute an estrogen induced growth response. A detailed comparison of the chromatin structure in the 5' flanking region of the human c-*myc* gene between hormone-dependent and hormone-independent human breast cancer cells was undertaken in order to obtain clues to the possible molecular mechanisms responsible for this differential regulation.

An analysis of the DNase I hypersensitive sites in the 5' flanking region of the c-*myc* gene in a range of human breast cancer cell lines demonstrated that DH sites I, II_2, III_1, III_2, IV, and V were present in all the cell lines studied. Although the position of the DH sites were identical, their relative intensities differed. Specifically, the intensity of DH site II_2 in the c-*myc* gene chromatin of ER negative cell lines was relatively strong in comparison to ER positive cell lines. The progression of an ER^+ cell line to an ER^- phenotype resulted in an increase in hypersensitivity of this site that was similar to other ER^- cell lines, however the reexpression of the ER was unable to reconstitute the ER^+ c-*myc* chromatin structure at DH site II_2. This observation supports the use of the c-*myc* gene as a model for studying molecular changes which may lead to progression of breast cancer cells to the hormonally independent phenotype. Furthermore, the DNase I

hypersensitivity results support the conclusion that the chromatin structure of the 5' flanking region of the human c-*myc* gene is different in ER positive and ER negative human breast cancer cells.

Differential use of the P_0 promoter was not observed between ER^+ and ER^- cell lines and therefore could not account for the differential hypersensitivity observed at DH site II_2. For all human breast cancer cell lines studied, DH site III_2 was more intense than DH site III_1, which is consistent with the observation that the P_2 promoter is more active than P_1 in these cells. Differences in c-*myc* amplification did not appear to be responsible for differential hypersensitivity at DH site II_2. Additionally, CpG sites in this region appeared to be undermethylated in both cell types, suggesting that differential methylation may not be involved in generating structural differences in the c-*myc* chromatin of breast cancer cell lines. Alternatively, the results of *in vitro* DNA binding assays revealed potential differences in the interaction of nuclear proteins from ER^+ and ER^- breast cancer cells with the DH II_2 and DH III_2 regions of c-*myc* .

The *in vitro* DNA binding assays demonstrated that Sp1-like proteins isolated from breast cancer cell nuclear extracts interact with previously characterized Sp1 binding sites located in the DH II_2 and DH III_2 regions of c-*myc*. More importantly, novel factor binding sites in the c-*myc* DH site II_2 region (-687 to -607) and in the DH site III_2 region (-1 to +66) were identified. Nuclear protein interactions with these c-*myc* sequences revealed differential *in vitro* DNase I hypersensitive sites and differences in complex intensities between ER^+ MCF 7 and ER^- MDA MB 231 breast cancer cells. These differences could be due to factor abundance, differential levels of protein modifications or differences in the association of sequence-specific DNA binding factors with non-DNA binding chromatin associated proteins.

Furthermore, these *in vitro* differences in protein interactions with c-*myc* regulatory regions may be responsible for the *in vivo* DNase I hypersensitivity observed between ER^- and ER^+ breast cancer cells.

Although inducible, tissue specific and developmental DH sites may require displacement of nucleosomes prior to gene activation, other DH sites may be nucleosome free as a result of specifically positioned nucleosomes or their exclusion during the replicative process. Previous studies suggest that the c-*myc* gene may be maintained in an open chromatin conformation with nucleosomes excluded from promoter proximal regions and precisely positioned in upstream flanking regions. Estrogen induced DH sites have been observed in other genes, but there is no direct evidence to suggest that the ER is able to displace nucleosomes. Therefore, it is not clear whether the ER induces an altered chromatin structure that is revealed by DNase I hypersensitivity or if the ER binds to an already altered chromatin structure, which is generated by other proteins. The absence of an estrogen induced change in c-*myc* chromatin structure in ER^+ breast cancer cells supports the hypothesis that proteins are already assembled at the promoter regions. The maintenance of this open conformation may be through the association of Sp1-like proteins with sites dispersed throughout the c-*myc* gene. The role of the ER may then be to increase the efficiency of transcription by enhancing the stability of the initiation complex through an association with the nuclear matrix.

A model for the association of factors interacting with the DH II_2, III_1 and III_2 regulatory regions was proposed based on the multiple Sp1 interactions found in the c-*myc* gene and the available literature supporting Sp1 multimers in the juxtapositioning of laterally separated elements. In hormone responsive breast cancer cells the interaction of the ER with the

c-*myc* chromatin may be involved in facilitating the interaction of these elements with the transcription initiation complex and the nuclear matrix, while other factors possibly associated with the novel DNA binding sites identified in this study may be involved in this process in hormonally independent cells. The results indicate that nuclear extracts of ER^+ MCF 7 and ER^- MDA MB 231 cells contain a similar, but not identical, spectrum of sequence specific DNA-binding proteins that bind to the DH II_2 and the DH III_2 regions of the human c-*myc* gene. These differences may be involved in the differential regulation of c-*myc* transcription in breast cancer cells and provide a basis for continuing the analysis of progression from hormonal dependence to independence.

Conclusion

This research has demonstrated that the chromatin structure in the 5' flanking region of the c-*myc* gene is different between ER^+ and ER^- human breast cancer cells. Although the positions of the DH sites were identical, their relative intensities differed. Specifically, the intensity of DH site II_2 of the c-*myc* gene chromatin was relatively strong in all ER^- cell lines studied, while this site was relatively weak in the ER^+ cell lines. The decreased hypersensitivity of this site in ER^+ cells could not be reconstituted by stable expression of the ER into originally ER^- cells. Conversely, the progression of ER^+ cells to an ER^- phenotype resulted in increased hypersensitivity at DH site II_2 which was similar to other ER^- cell lines. The molecular mechanisms responsible for this alteration in c-*myc* chromatin structure did not appear to be due to differential promoter usage, gene amplification or methylation at CpG sites. Differential protein interactions detected at previously uncharacterized DNA binding sites located in the DH II_2 and III_2 regions of c-*myc* are proposed to be involved in the generation of the different chromatin

conformations. Sp1-like proteins, interacting with elements in the DH II_2 and III_2 regions, are proposed to induce looping out of the intervening sequences and facilitate the interaction of regulatory proteins associated with these and other DH sites with the preinitiation complex. The association of the ER with hormone responsive c-*myc* chromatin is suggested to be involved in these interactions in ER^+ breast cancer cells, but other factors, possibly those associated with DNA binding sites identified in this study, are proposed to compensate for the loss of ER activity in hormonally independent breast cancer cells.

References

Abrams, H.D., Rohrschneider, L.R., Eisenman, R.N. 1982. Nuclear location of the putative transforming protein of avian myelocytomatosis virus. Cell 29: 427-439.

Abulafia, R., Ben-Ze'ev, A., Hay, N., Aloni, Y. 1984. Control of late simian virus-40 transcription by the attenuation mechanism and transcriptionally active ternary complexes are associated with the nuclear matrix. J. Mol. Biol. 172: 467-487.

Adams, C.C. and Workman, J.L. 1993. Nucleosome displacement in transcription. Cell 72: 305-308.

Adams, J.M., Haris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., Brinster, R.L. 1985. The c-*myc* oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature 318: 533-538.

Adams, R.L.P. 1990. DNA methylation - the effect of minor bases on DNA-protein interactions. Biochem. J. 265: 309-320.

Adler, S., Waterman, M.L., He, X., Rosenfeld, M.G. 1988. Steroid receptor mediated inhibition of rat prolactin gene expression does not require the receptor DNA-binding domain. Cell 52: 685-695.

Alexander, R.B., Greene, G.L., Barrack, E.R. 1987. Estrogen receptors in the nuclear matrix: direct demonstration using monoclonal antireceptor antibody. Endocrinol. 120: 1851-1857.

Ali, I.U., Merlo, G., Callahan, R., Liderau, R. 1989. The amplification unit of chromosome 11q13 in aggressive primary breast tumors entail the *bcl-1*, *int-2* and *host* loci. Oncogene 4: 89-92.

Alitalo, K., Ramsay, G., Bishop, J.M., Pfeifer, S.O., Colby, W.W., Levinson, A.D. 1983. Identification of nuclear proteins encoded by viral and cellular *myc* oncogenes. Nature 306: 274-277.

Alonso, M.C. and Cabrera, C.V. 1988. The *achaete-scute* gene complex of *Drosophila melanogaster* comprises four homologous genes. EMBO J. 7: 2585-2591.

Alt, F.W., DePinho, R.A., Zimmerman, K., Legouy, E., Hatton, K., Ferrier, P., Tesfaye, A., Yancopoulos, G.D., Nisen, P. 1986. The human *myc* gene family. Cold Spring Harbor Symposia on Quantitative Biology 51: 931-941.

Alvarez, E., Northwood, I.C., Gonzalez, F.A., Latour, D.A., Seth, A., Abate, C., Curran, T., Davis, R.J. 1991. Pro-Leu-Ser/Thr-Pro is a consensus primary sequence for substrate protein phosphorylation - characterization of the phosphorylation of c-Myc and c-Jun proteins by an epidermal growth factor receptor threonine-669 protein-kinase. J. Biol. Chem. 266: 5277-5285.

Amati, B., Brooks, M.W., Levy, N., Littlewood, T.D., Evan, G.I., Land, H. 1993. Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell 72: 233-245.

Amati, B., Dalton, S., Brooks, M.W., Littlewood, T.D., Evan, G.I., Land, H. 1992. Transcriptional activation by the c-Myc oncoprotein in yeast requires interaction with Max. Nature 359: 423-426.

Amin, C., Wagner, A.J., Hay, N. 1993. Sequence-specific transcriptional activation by Myc and repression by Max. Mol. Cell. Biol. 13(1): 383-390.

Anderson, D.E. 1974. Genetic study of breast cancer - identification of a high risk group. Cancer 34: 1090-1097.

Anderson, G.M. and Freytag, S.O. 1991. Synergistic activation of a human promoter *in vivo* by transcription factor Sp1. Mol. Cell. Biol. 11: 1935-1943.

Andreeva, M., Markova, D., Loidl, P., Djondjurov, L. 1992. Intranuclear compartmentalization of transcribed and nontranscribed c-*myc* sequences in Namalva-S cells. Eur. J. Biochem. 207: 887-894.

Angel, P. and Karin, M. 1991. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochem. Biophys. Acta 1072: 129-157.

Arafah B.M . and Pearson O.H. 1986. In: Estrogen/Antiestrogen Action and Human Breast Cancer Therapy. pp 417-429. Jordan V.C., ed. University of Wisconsin Press.

Archer, T.K., Lefebvre, P., Wolford, R.G., Hager, G.L. 1992. Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. Science 255: 1573-1576.

Armstrong, B.K. 1988. Estrogen therapy after the menopause - boon or bane. Med. J. Austral. 148: 213-214.

ar-Rushdi, A., Nishikura, K., Erikson, J., Watt, R., Rovera, G., Croce, C.M. 1983. Differential expression of the translocated and the untranslocated c-*myc* oncogene in Burkitt lymphoma. Science 222: 390-393.

Arents, G., Burlingame, R.W., Wang, B.-C., Love, W.E., Moudrianakis, E.N. 1991. The nucleosomal core histone octamer at 3.1 Å resolution: A tripartite protein assembly and left-handed superhelix. Proc. Natl. Acad. Sci. 88: 10148-10152.

Ariga, H., Imamura, Y., Iguchi-Ariga, S.M.M. 1989. DNA replication origin and transcriptional enhancer in c-*myc* gene share the c-*myc* protein binding sequences. EMBO J. 8(13): 4273-4279.

Armelin, H.A., Armelin, M.C.S., Kelly, K., Stewart, T., Leder, P., Cochran, B.H., Stiles, C.D. 1984. Functional role for c-*myc* in mitogenic response to platelet-derived growth factor. Nature 310: 655-660.

Asselin, C., Nepveu, A., Marcu, K.B. 1989. Molecular requirements for transcriptional initiation of the murine c-*myc* gene. Oncogene 4: 549-558.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds.) 1989. Current Protocols in Molecular Biology. 3rd edition. Wiley Interscience.

Avigan, M.I., Strober, B., Levens, D. 1989. A far upstream element stimulates c-*myc* expression in undifferentiated leukemia cells. J. Biol. Chem. 265: 18538-18545.

Aviv, H. and Leder, P. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. Proc. Natl. Acad. Sci. 69: 1408-1412.

Ayer, D.E., Kretzner, L., Eisenman, R.N. 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell 72: 211-222.

Baas, F., Bikker, H., Geurts van Kessel, A., Melsert, R., Pearson, P.L., de Vijlder, J.J., van Ommen, G.J. 1985. The human thyroglobulin gene: a polymorphic marker localized distal to c-*myc* on chromosome 8 band q 24. Hum. Gen. 69: 138-143.

Bagchi, S., Weinmann, R., Raychaudhuri, P. 1991. The retinoblastoma protein copurifies with E2F-1, an E1A-regulated inhibitor of the transcritpion factor E2F. Cell 65: 1063-1072.

Bandara, L.R. and La Thangue, N.B. 1991. Adenovirus-E1A prevents the retinoblastoma gene-product from complexing with a cellular transcription factor. Nature 351: 494-497.

Barbi, G.P., Marroni, P., Bruzzi, P., Nicolò, G., Paganuzzi, M., Ferrara, G.B. 1987. Correlation between steroid hormone receptors and prognostic factors in human breast cancer. Oncology 44: 265-269.

Barr, L.F., Mabry, M., Nelkin, B.D., Tyler, G., May, W.S., Baylin, S.B. 1991. c-*myc* gene-induced alterations in protein kinase C expression: a possible mechanism facilitating *myc-ras* gene complementation. Cancer Res. 51: 5514-5519.

Barrack, E.R. and Coffey, D.S. 1980. The specific binding of estrogens and androgens to the nuclear matrix of sex hormone responsive tissues. J. Biol. Chem. 255: 7265-7275.

Baserga, R. 1978. Resting cells and the G_1 phase of the cell cycle. J. Cell Physiol. 95: 377-386.

Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G., Leder, P. 1983. The human c-*myc* oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell 34: 779-787.

Beato, M. 1989. Gene regulation by steroid hormones. Cell 56: 335-344.

Beatson, G.T. 1896. On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment, with illustrative cases. Lancet 2: 104-107.

Becker, P.B., Gloss, B., Schmid, W., Strahle, U., Schutz, G. 1986. *In vivo* protein DNA interactions in a glucocorticoid response element require the presence of the hormone. Nature 324: 686-688.

Beckman, H., Su, L.-K., Kadesch, T. 1990. TFE3: a helix-loop-helix protein that activates transcription through the immunoglobin enhancer μE3 motif. Genes Dev. 4: 167-179.

Behe, M. and Felsenfeld, G. 1981. Effects of methylation on a synthetic polynucleotide - the B-Z transition in poly(dG-m5dC) · poly (dG-m5dC). Proc. Natl. Acad. Sci. 78: 1619-1623.

Bello-Fernandez, C., Packham, G., Cleveland, J.L. 1993. The ornithine decarboxylase gene is a transcriptional target of c-myc. Proc. Natl. Acad. Sci. 90: 7804-7808.

Bentley, D.L. and Groudine, M. 1986a. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature 321: 702-706.

Bentley, D.L. and Groudine, M. 1986b. Novel promoter upstream of the human c-*myc* gene and regulation of c-*myc* expression in B-cell lymphomas. Mol. Cell. Biol. 6: 3481-3489.

Bentley, D.L. and Groudine, M. 1988. Sequence requirements for premature termination of transcription in the human c-*myc* gene. Cell 53: 245-256.

Berberich, S., Hyde-DeRuyscher, N., Espenshade, P., Cole, M. 1992. *max* encodes a sequence-specific DNA-binding protein and is not regulated by serum growth factors. Oncogene 7: 775-779.

Berberich, S.J. and Cole, M.D. 1992. Casein kinase II inhibits the DNA-binding activity of Max homodimers but not Myc/Max heterodimers. Genes Dev. 6: 166-176.

Berezney, R. 1991. The nuclear matrix: a heuristic model for investigating genomic organization and function in the cell nucleus. J. Cell. Biochem. 47: 109-123.

Berezney, R. and Coffey,D. 1974. Identification of a nuclear protein matrix. Biochem. Biophys. Res. Commun. 60: 1410-1419.

Berg, J.M. 1988. Proposed structure for the zinc-binding domains from transcription factor-IIIA and related proteins. Proc. Natl. Acad. Sci. 85: 99-102.

Bergemann, A.D. and Johnson, E.M. 1992. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. Mol. Cell. Biol. 12(3): 1257-1265.

Bernard, O., Cory, S., Gerondakis, S., Webb, E., Adams, J.M. 1983. Sequence of the murine and human cellular *myc* oncogenes and two modes of *myc* transcription resulting from chromosome translocation in B lymphoid tumours. EMBO J. 2: 2375-2383.

Bernards, R., Dessain, S., Weinberg, R. 1986. N-*myc* amplification causes down-modulation of MHC class I antigen expression in neuroblastoma. Cell 47: 667-674.

Berns, E.M.J.J., Klijn, J.G.M., van Putten, W.L.J., van Staveren, I.L., Portengen, H., Foekens, J.A. 1992. c-*myc* amplification is a better prognostic factor than *HER2/neu* amplification in primary breast cancer. Cancer Res. 52: 1107-1113.

Bernstein, P.L., Herrick, D.J., Prokipcak, R.D., Ross, J. 1992. Control of c-*myc* mRNA half-life *in vitro* by a protein capable of binding to a coding region stability determinant. Genes Dev. 6: 642-654.

Berry, M., Nunez, A.-M., Chambon, P. 1989. Estrogen-responsive element of the human pS2 gene is an imperfect palindromic sequence. Proc. Natl. Acad. Sci. 86: 1218-1222.

Berthois, Y., Katzenellenbogen, J.A., Katzenellenbogan, B.S. 1986. Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. Proc. Natl. Acad. Sci. 83: 2496-2500.

Bianchi, N.O., Bianchi, M.S., Alitalo, K., de la Chapelle, A. 1989. The methylation pattern of normal and truncated amplified human c-*myc* oncogenes. Biochem. Biophys. Acta 1007: 350-355.

Bidwell, J.P., van Wijnen, A.J., Fey, E.G., Dworetzky, S., Penman, S., Stein, J.L., Lian, J.B., Stein, G.S. 1993. Osteocalcin gene promoter-binding factors are tissue-specific nuclear matrix components. Proc. Natl. Acad. Sci. 90: 3162-3166.

Bird, A. 1992. The essentials of DNA methylation. Cell 70: 5-8.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209-213.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513-1523.

Birrer, M.J., Segal, S., DeGreve, J.S., Kaye, F., Sausville, E.A., Minna, J.D. 1988. L-*myc* cooperates with *ras* to transform rat embryo fibroblasts. Mol. Cell. Biol. 8: 2668-2673.

Bissonnette, R.P., Echeverri, F., Mahboubi, A., Green, D.R. 1992. Apoptotic cell death induced by c-*myc* is inhibited by *bcl-2*. Nature 359: 552-554.

Bister, K., Trachmann,C., Jansen, H.W., Schroeer, B., Patchinsky, T. 1987. Structure of mutant and wild-type MC29 v-*myc* alleles and biochemical properties of their protein products. Oncogene 1: 97-109.

Black, D.M. and Solomon, E. 1993. The search for the familial breast-ovarian cancer gene. Trends Genet. 9(1): 22-26.

Blackwell, T.K. and Weintraub, H. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding-site selection. Science 250: 1104-1110.

Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N., Weintraub, H. 1990. Sequence-specific DNA binding by the c-Myc protein. Science 250: 1149-1151.

Blackwood, E.M. and Eisenman, R.N. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science 251: 1211-1217.

Blackwood, E.M., Lüscher, B., Eisenman, R.N. 1992. Myc and Max associate *in vivo*. Gene Dev. 6: 71-80.

Blanco, G., Alavaikko, M., Ojala, A., Collan, Y., Heikinnen, M., Hietanen, T., Aine, R., Taskinen, P.J. 1984. Estrogen and progesterone receptors in breast cancer: relationships to tumour histopathology and survival of patients. Anticancer Res. 4: 383-390.

Bocquel, M.T., Kumar, V., Stricker, C., Chambon, P., Gronemeyer, H. 1989. The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. Nucleic Acids Res. 17: 2581-2595.

Bode, J., Kohwi, Y., Dickinson, L., Jon, T., Klehr, D., Mielke, C., Kowhi-Shigematsu, T. 1992. Biological significance of unwinding capability of nuclear matrix-associating DNAs. Science 255: 195-197.

Bodner, M., Castrillo, J-L., Theill, L.E., Deerinck, T., Ellisman, M., Karin, M. 1988. The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell 55: 505-518.

Bonadonna, G. and Valagussa, P. 1988. The contribution of Medicine to the primary treatment of breast cancer. Cancer Res. 48: 2314-2324.

Bonifer, C., Hecht, A., Saueressig, H., Winter, D.M., Sippel, A.E. 1991. Dynamic chromatin: the regulatory domain organization of eukaryotic gene loci. J. Cell. Biochem. 47: 99-108.

Bonnieu, A., Roux, P., Marty, L., Jeanteur, P., Piechaczyk, M. 1990. AUUUA motifs are dispensible for rapid degradation of the mouse c-*myc* mRNA. Oncogene 5: 1670-1674.

Bossone, S.A., Asselin, C., Patel, A.J., Marcu, K.B. 1992. MAZ, a zinc finger protein, binds to c-*myc* and *C 2* gene sequences regulating transcriptional initiation and termination. Proc. Natl. Acad. Sci. 89: 7452-7456.

Boulikas, T. 1992. Homeotic protein binding sites, origins of replication and nuclear matrix anchorage sites share the ATTA and ATTTA motifs. J. Cell. Biochem. 50: 111-123.

Boulikas, T., Wiseman, J.M., Garrard, W.T. 1980. Points of contact between histone H1 and the histone octamer. Proc. Natl. Acad. Sci. 77: 127-131.

Bradford, M.M. 1976. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principles of protein-dye binding. Anal. Biochem. 72: 248-254.

Bresciani, F. 1965. Effect of ovarian hormones on duration of DNA synthesis in cells of the C3H mouse mammary gland. Exp. Cell Res. 38: 13-32.

Bresciani, F. 1968. Topography of DNA synthesis in the mammary gland of the C3H mouse and its control by ovarian hormones: an autoradiographic study. Cell Tissue Kinet. 1: 51-63.

Brewer, G. 1991. An A+U-rich element RNA-binding factor regulates c-*myc* mRNA stability *in vitro*. Mol. Cell. Biol. 11: 2460-2466.

Brinkley, D. and Haybittle, J.L. 1975. The curability of breast cancer. Lancet 2: 95-97.

Brinton, L.A., Hoover, R., Fraumeni, J.F. Jr. 1986. Menopausal estrogens and breast cancer risk - an expanded case-control study. Br. J. Cancer 54: 825-832.

Broome, H.E., Reed, J.C., Godillot, E.P., Hoover, R.G. 1987. Differential promoter utilization by the c-*myc* gene in mitogen and interleukin-2 stimulated human lymphocytes. Mol. Cell. Biol. 7: 2988-2993.

Brown, M. and Sharp, P.A. 1990. Human estrogen receptor forms multiple protein-DNA complexes. J. Biol. Chem. 265: 11238-11243.

Brünner, N., Boulay, V., Fojo, A., Freter, C.E., Lippman, M.E., Clarke, R. 1993. Acquisition of hormone-independent growth in MCF-7 cells is accompanied by increased expression of estrogen-regulated genes without detectable DNA amplifications. Cancer Res. 53: 283-290.

Butnick, N.Z., Miyamoto, C., Chizzonite, R., Cullen, B.R., Ju, G., Skalka, A.M. 1985. Regulation of the human c-*myc* gene: 5' noncoding sequences do not affect translation. Mol. Cell. Biol. 5: 3009-3016.

Cairns, J. 1975. Mutation selection and natural history of cancer. Nature 255: 197-200.

Cailleau, R., Young, R., Olive, M., Reeves, W.J., Jr. 1974. Breast tumor cell lines from pleural effusions. J. Natl. Cancer Inst. 53: 661-674.

Callahan, R. 1989. Genetic alterations in primary breast cancer. Breast Cancer Res. Treat. 13: 191-203.

Callahan, R. and Campbell, G. 1989. Mutations in human breast cancer: an overview. J. Natl. Cancer Inst. 81: 1780-1786.

Campisi, J., Gray, H.E., Pardee, A.B., Dean, M., Sonenshein, G.E. 1984. Cell-cycle control of c-*myc* but not c-*ras* expression is lost following chemical transformation. Cell 36: 241-247.

Carr, C.S. and Sharp, P.A. 1990. A helix-loop-helix protein related to the immunoglobulin E box-binding proteins. Mol. Cell. Biol. 10: 4384-4388.

Carter, K.C. and Lawrence, J.B. 1991. DNA and RNA within the nucleus: how much sequence-specific spatial organization? J. Cell. Biochem. 47: 124-129.

Caudy, M., Vässin, H., Brand, M., Tuma, R., Jan, L.Y., Jan, Y.N. 1988. *Daughterless*, a *Drosophila* gene essential for both neurogenesis and sex determination, has sequence similarities to *myc* and the *achaete-scute* complex. Cell 55: 1061-1067.

Cavaillès, V., Augereau, P., Rochefort, H. 1993. Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells. Proc. Natl. Acad. Sci. 90: 203-207.

Cedar, H. 1988. DNA methylation and gene activity. Cell 53: 3-4.

Ceriani, R.L. 1970. Fetal mammary gland differentiation *in vitro* in response to hormones. II. Biochemical findings. Dev. Biol. 21: 530-546.

Cesarman, E., Dalla-Favera, R., Bentley, D., Groudine, M. 1987. Mutations in the first exon are associated with altered transcription of c-*myc* in Burkitt Lymphoma. Science 238: 1272-1275.

Cesarone, C.F., Bolognesi, C., Santi, L. 1979. Improved microfluorometric DNA determination in biological material using 33258 Hoechst. Anal. Biochem. 100: 188-197.

Chalbos, D., Vignon, F., Keydar, I., Rochefort, H. 1982. Estrogens stimulate cell proliferation and induce secretory proteins in a human breast cancer cell line (T 47D). J. Clin. Endocrinol. Met. 55: 276-283.

Charron, J., Malynn, B.A., Fisher, P., Stewart, V., Jeannotte, L., Goff, S.P., Robertson, E.J., Alt, F.W. 1992. Embryonic lethality in mice homozygous for a targeted disruption of the N-*myc* gene. Genes Dev. 6: 2248-2257.

Chellappan, S.P., Hiebert, S., Mudryj, M., Horowitz, J.M., Nevins, J.R. 1991. The E2F transcription factor is a cellular target for the RB protein. Cell 65: 1053-1061.

Chirgwin, J., Pryzbyla, A., MacDonald, R., Rutter, W. 1979. Isolation of biologically active ribonucleic acid from sources enriched for ribonuclease. Biochemistry 18: 5294-5298.

Chou, R.H., Churchill, J.R., Flubacher, M.M., Mapstone, D.E., Jones, J. 1990. Identification of a nuclear matrix-associated region of the c-*myc* proto-oncogene and its recognition by a nuclear protein in the human leukemia HL-60 cell line. Cancer Res. 50: 3199-3206.

Chung, J., Sinn, E., Reed, R.R., Leder, P. 1986. *Trans*-acting elements modulate expression of the human c-*myc* gene in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. 83: 7918-7922.

Chung, J., Sussman, D.J., Zeller, R., Leder, P. 1987. The c-*myc* gene encodes superimposed RNA polymerase II and III promoters. Cell 51: 1001-1008.

Churchill, M.E.A. and Travers, A.A. 1991. Protein motifs that recognize structural features of DNA. Trends Biochem. Sci. 16: 92-97.

Churchill, M.E.A. and Suzuki, M. 1989. Spkk motifs prefer to bind to DNA at A/T-rich sites. EMBO J. 8: 4189-4195.

Clarke, R., Paik, S., Norman, M.J., Katzenellenbogen, B.S., Koppi, C., Thompson, E.W., Dickson, R.B., Brünner, N., Lippman, M.E. 1989. Progression of human-breast cancer-cells from hormone-dependent to hormone-independent growth both *in vitro* and *in vivo*. Proc. Natl. Acad. Sci. 86: 3649-3653.

Clewell, D.B. 1972. Nature of Col E1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. 110: 667-676.

Cockerill, P.N. and Garrard, W.T. 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell 44: 273-282.

Cogswell, J.P., Cogswell, P.C., Kuehl, W.M., Cuddihy, A.M., Bender, T.M., Engelke, U., Marcu, K.B., Ting, J. P.-Y. 1993. Mechanism of c-*myc* regulation by c-Myb in different cell lineages. Mol. Cell. Biol. 13 (5) : 2858-2869.

Cole. M.D. 1986. The *myc* oncogene: its role in transformation and differentiation. Annu. Rev. Genet. 20: 361-384.

Cole. M.D. 1991. Myc meets its Max. Cell 65: 715-716.

Compton, J.L., Bellard, M., Chambon, P. 1976. Biochemical evidence of variability in the DNA repeat length in the chromatin of higher eukaryotes. Proc. Natl. Acad. Sci. 73(12): 4382-4386.

Cook, P.R. 1989, The nucleoskeleton and the topology of transcription. Eur. J. Biochem. 185: 487-501.

Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J., Hogan, M.E. 1988. Site-specific oligonucleotide binding represses transcription of the human c-*myc* gene *in vitro*. Science 241: 456-459.

Cory, S. 1986. Activation of cellular oncogenes in hematopoietic cells by chromosome translocation. Adv. Cancer Res. 47: 189-234.

Courey, A.J., Holtzman, D.A., Jackson, S.P. Tjian, R. 1989. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell 59: 827-836.

Crépin, M., Salle, V., Raux, H., Berger, R., Hamelin, R., Brouty-Boyé, D., Israel, L. 1990. Steroid hormone receptors and tumorigenicity of sublines from breast tumor metastatic MDA MB 231 cell line. Anticancer Res. 10: 1661-1666.

Croston, G.E., Kerrigan, L.A., Lira, L.M., Marshak, D.R., Kadonaga, J.T. 1991. Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. Science 251: 643-649.

Dalla-Favera. R., Gelmann, E.P., Martinotti, S., Franchini, G., Papas, T.S., Gallo, R.C., and Wong-Staal, F. 1982. Cloning and characterization of different human sequences related to the onc gene (v-*myc*) of avian myelocytomatosis virus (MC29). Proc. Natl. Acad. Sci. 79: 6497-6501.

Dang, C.V. 1991. c-Myc oncoprotein function. Biochem. Biophys. Acta, 1072: 103-113.

Dang, C.V. and Lee, V.M.F. 1988. Identification of the human c-*myc* protein nuclear translocation signal. Mol. Cell. Biol. 8: 4048-4054.

Dang, C.V., McGuire, M., Buckmire, M., Lee, W.M.F. 1989. Involvement of the 'leucine zipper' region in the oligomerization and transforming activity of human c-*myc* protein. Nature 337: 664-666.

Darbre, P.D. and King, R.J.B. 1987. Progression to steroid insensitivity can occur irrespective of the presence of functional steroid receptors. Cell 51: 521-528.

Darbre, P., Yates, J., Curtis, S., King, R.J.B. 1983. Effect of estradiol on human breast cancer cells in culture. Cancer Res. 43: 349-354.

Darnell, J.E. Jr. 1982. Variety in the level of gene control in eukaryotic cells. Nature 297: 365-371.

Darveau, A., Pelletier, J., Sonenberg, N. 1985. Differential efficiencies of *in vitro* translation of mouse c-*myc* transcripts differing in the 5' untranslated region. Proc. Natl. Acad. Sci. 82: 2315-2319.

Darzynkiewicz, Z. 1983. Molecular interactions and cellular changes during the cell cycle. Pharm. Ther. 21: 143-188.

Davidson, N.E., Prestigiacomo, L.J., Hahm, H.A. 1993. Induction of Jun gene family members by transforming growth factor-alpha but not 17-beta-estradiol in human breast-cancer cells. Cancer Res. 53(2): 291-297.

Davie, J.R. and Murphy, L.C. 1990. Level of ubiquitinated histone H2B in chromatin is coupled to ongoing transcription. Biochemistry 29: 4752-4757.

Davis, L.G., Dibner, M.D., Battey, J.F. (eds). 1986. Basic Methods in Molecular Biology. Elsevier Publishers, New York.

Davis, R.L., Weintraub, H., Lasser, A.B. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987-1000.

Davis, T.L., Firulli, A.B., Kinniburgh, A.J. 1989. Ribonucleoprotein and protein factors bind to an H-DNA-forming c-*myc* DNA element: possible regulators of the c-*myc* gene. Proc. Natl. Acad. Sci. 86: 9682-9686.

Dean, M., Kent, R.B., Sonenshein, G.E. 1983. Transcriptional activation of immunoglobulin alpha heavy-chain genes by translocation of the c-*myc* oncogene. Nature 305: 443-446.

Dean, M., Levine, R.A., Ran, W., Kindy, M.S., Sonenshein, G.E. 1986. Regulation of c-*myc* transcription and mRNA abundance by serum growth factors and cell contact. J. Bio. Chem. 261: 9161-9166.

DeGreve, J., Battey, J., Fedorko, J., Birrer, M., Evan, G., Kaye, F., Sausville, E., Minna, J. 1988. The human L-*myc* gene encodes multiple phosphoproteins from alternatively processed mRNAs. Mol. Cell. Biol. 8: 4381-4388.

Demczuk, S., Donovan, M., Franklin, G., Ohlosson, R. 1991. Order of probe and nuclear protein extract addition can determine specificity of protein-DNA complexes in tested mobility shift assays. Nucleic Acids Res. 19: 677-678.

Denhardt, D.T. 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. 23: 641-646.

DePinho, R.A., Schreiber-Agus, N., and Alt, F.W. 1991. *myc* family oncogenes in the development of normal and neoplastic cells. Adv. Cancer Res. 57: 1-46.

DePinho, R.A., Mitsock, L., Hatton, K., Ferrier, P., Zimmerman, K., Legouy, E., Tesfaye, A., Collum, R., Yancopoulos, G., Nisen, P., Kriz, R., Alt, F. 1987a. Myc family of cellular oncogenes. J. Cell. Biochem. 33: 257-266.

DePinho, R.A., Hatton, K.S., Tesfaye, A., Yancopoulos, G.D., Alt, F.W. 1987b. The human *myc* gene family: structure and activity of L-*myc* and an L-*myc* pseudogene. Genes Dev. 1: 1311-1326.

DesJardins, E. and Hay, N. 1993. Repeated CT-elements bound by zinc finger proteins control the absolute and relative activities of the two principle human c-*myc* promoters. Mol. Cell. Biol. 13: 5710-5724.

Desrosiers, R. and Tanguay, R.M. 1985. The modifications in the methylation patterns of H2B and H3 after heat shock can be correlated with the inactivation of normal gene expression. Biochem. Biophys. Res. Commun. 133: 823-829.

DeVos, A.M., Tong, L., Milburn, M.V., Matias, P.M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., Kim, S-H. 1988. Three dimensional structure of an oncogene protein: catalytic domain of human c-H-*ras* p21. Science 239: 888-893.

Diamond, M.I., Miner, J.N., Yoshinaga, S.K., Yamamoto, K.R. 1990. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science 249: 1266-1272.

Didier, P., Kumar. V., Chambon, P., Yamamoto, K.R. 1990. Signal transduction by steroid hormones: nuclear localization is differentially regulated in estrogen and glucocorticoid receptors. Cell Regulation 1: 291-299.

Dignam, J.D., Lebovitz, R.M., Roeder, R.G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11: 1475-1489.

Dillon, N. and Grosveld, F. 1993. Transcriptional regulation of multigene loci: multilevel control. Trends Genet. 9: 1000-1003.

Donner, P., Greiser-Wilke, I, Moelling, K. 1982. Nuclear localization and DNA binding of the transforming gene product of avian myelocytomatosis virus. Nature 296: 262-266.

Dotzlaw, H., Miller, T., Karvelas, J., Murphy, L.C. 1990. Epidermal growth factor gene expression in human breast cancer biopsy samples: relationship to estrogen and progesterone receptor gene espression. Cancer Res. 50: 4204-4208.

Doucas, V., Spyrou, G., Yaniv, M. 1991. Unregulated expression of c-Jun or c-Fos proteins but not Jun D inhibits oestrogen receptor activity in human breast cancer derived cells. EMBO J. 10: 2237-2245.

Downs, K.M., Martin, G.R., Bishop, J.M. 1989. Contrasting patterns of c-*myc* and N-*myc* expression during gastrulation of the mouse embryo. Genes Dev. 3: 860-869.

Draetta, G., Piwnica-Worms, H., Morrison, D., Druker, B., Roberts, T., Beach, D. 1988. Human cdc2 protein kinase is a major cell-cycle regulated tyrosine kinase substrate. Nature 336: 738-744.

Drews, U. and Drews, U. 1977. Regression of mouse mammary anlage in recombinants of Tfm and wild-type tissues: testosterone acts via the mesenchyme. Cell 10: 401-404.

Dubik, D. 1991. Molecular mechanism of estrogen activation of c-*myc* oncogene expression in human breast cancer cells. PhD thesis, University of Manitoba.

Dubik, D. and Shiu, R.P.C. 1988. Transcriptional regulation of c-*myc* oncogene expression by estrogen in hormone-responsive human breast cancer cells. J. Biol. Chem. 263: 12705-12708.

Dubik, D. and Shiu, R.P.C. 1992. Mechanism of estrogen activation of c-*myc* oncogene expression. Oncogene 7: 1587-1594.

Dubik, D., Dembinski, T.C. and Shiu, R.P.C. 1987. Stimulation of c-*myc* oncogene expression associated with estrogen-induced proliferation of human breast cancer cells. Cancer Res. 47: 6517-6521.

Duesberg, P.H., Bister, K. and Vogt, P.K. 1977. The RNA of avian acute leukemia virus MC29. Proc. Natl. Acad. Sci. 74: 4320-4324.

Dufort, D., Drolet, M., Nepveu, A. 1993. A protein binding site from the murine c-*myc* promoter contributes to transcriptional block. Oncogene 8: 165-171.

Durrin, L.K., Mann, R.K., Grunstein, M. 1992. Nucleosome loss activates *cup1* and *his3* promoters to fully induced levels in the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. 12: 1621-1629.

Duyao, M.P., Buckler, A.J., Sonenshein, G.E. 1990. Interaction of an NF-κB-like factor with a site upstream of the c-*myc* promoter. Proc. Natl. Acad. Sci. 87: 4727-4731.

Dworetzky, S.I., Wright, K.L., Fey, E.G., Penman, S., Lian, J.B., Stein, J.L., Stein, G.S. 1992. Sequence-specific DNA-binding proteins are components of a nuclear matrix-attachment site. Proc. Natl. Acad. Sci. 89: 4178-4182.

Dyson, P.J. and Rabbitts, T.H. 1985. Chromatin structure around the c-*myc* gene in Burkitt lymphomas with upstream and downstream translocation points. Proc. Natl. Acad. Sci. 82: 1984-1988.

Dyson, P.J., Littlewood, T.D., Forster, A., Rabbitts, T.H. 1985. Chromatin structure of transcriptionally active and inactive human c-*myc* alleles. EMBO J. 4: 2885-2891.

Easton, D.F., Bishop, D.T., Ford, D., Cockford, G.P. 1993. Genetic-linkage analysis in familial breast and ovarian cancer - results from 214 families. Am. J. Hum. Genet. 52: 678-701.

Edmondson, D.G. and Olson, E.N. 1989. A gene with homology to the *myc* similarity region of *myod1* is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev. 3: 628-640.

Egli, G.E., Egli, N.S., Newton, M. 1961. The influence of the number of breast-feedings on milk production. Pediatric 27: 314-317.

Eick, D. and Bornkamm, G.W. 1986. Transcriptional arrest within the first exon is a fast control mechanism in the c-*myc* gene expression. Nucleic Acids. Res. 14: 8331-8345.

Eick, D., Polack, A., Kofler, E., Lenoir, G.M., Rickinson, A.B., Bornkamm, G.W. 1990. Expression of P_0- and P_3-RNA from the normal and translocated c-*myc* allele in Burkitt's lymphoma cells. Oncogene 5: 1397-1402.

Eilers, M., Picard, D., Yamamoto, K.R., Bishop, J.M. 1989. Chimaeras of Myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. Nature 340: 66-68.

Eilers, M., Schirm, S., Bishop, J.M. 1991. The Myc protein activates transcription of the *α-prothymosin* gene. EMBO J. 10: 133-141.

Eisenman, R.N., Tachibani, C.Y., Abrams, H.D., Hann, S.R. 1985. v-*myc* and the c-*myc* encoded proteins are associated with the nuclear matrix. Mol. Cell. Biol. 5: 114-126.

Elgin, S.C.R. 1988. The formation and function of DNase I hypersensitive sites in the process of gene activation. J. Biol. Chem. 263: 19259-19262.

Ellis, R.E., Yuan, J., Horvitz, H.R. 1991. Mechanisms and functions of cell death. Annu. Rev. Cell Biol. 7: 663-698.

Elsholtz, H.P., Albert, V.R., Treacy, M.N., Rosenfeld, M.G. 1990. A two-base change in a POU factor binding site switches pituitary-specific to lymphoid-specific gene expression. Genes Dev. 4: 43-51.

Engel, L.W. and Young, N.A. 1978. Human breast carcinoma cells in continuous culture: a review. Cancer Res. 38: 4327-4339.

Engel, L.W., Young, N.A., Tralka, T.S., Lippman, M.E., O'Brien, S.J., Joyce, M.J. 1978. Establishment amd characterization of three new continuous cell lines derived from human breast carcinomas. Cancer Res. 38: 3352-3364.

Epstein, R.J. 1988. The clinical biology of hormone-responsive breast cancer. Cancer Treat. Rev. 15: 33-51.

Erikson, J., Finger, L., Sun, L., ar-Rushidi, A., Nishikura, K., Minowada, J., Finan, J., Emanuel, B.S., Nowell, P.C., Croce, C.M. 1986. Deregulation of c-*myc* by translocation of the alpha-locus of the T-cell receptor in T-cell leukemias. Science 232: 884-886.

Escot, C., Thiellet, C., Lidereau, R. Spyratos, F., Champene, M.H., Gest, J., Calahan, R. 1986. Genetic alterations of the c-*myc* proto-oncogene (Myc) in human primary breast carcinomas. Proc. Natl. Acad. Sci. 83: 4834-4838.

Evan, G.I. and Hancock, D.C. 1985. Studies on the interaction of the human c-*myc* protein with cell nuclei: p62$^{c\text{-}myc}$ as a member of a discrete subset of nuclear proteins. Cell 43: 253-261.

Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., Hancock, D.C. 1992. Induction of apoptosis in fibroblasts by c-*myc* protein. Cell 69: 119-128.

Evans, R.M. 1988. The steroid and thyroid hormone receptor superfamily. Science 240: 889-895.

Faisst, S. and Meyer, S. 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. 20: 3-26.

Falvey, E. and Schibler, U. 1991. How are the regulators regulated? FASEB J. 5: 309-314.

Fanidi, A., Harrington, E.A., Evan, G.I. 1992. Cooperative interaction between c-*myc* and *bcl-2* proto-oncogenes. Nature 359: 554-556.

Fawell, S.E., White, R., Hoare, S., Sydenham, M., Page, M., Parker, M.G. 1990. Inhibition of estrogen-receptor DNA-binding by the pure antiestrogen ICI-164,384 appears to be mediated by impaired receptor dimerization. Proc. Natl. Acad. Sci. 87: 6883-6887.

Felsenfeld, G. 1992. Chromatin as an essential part of the transcriptional mechanism. Nature 355: 219-224.

Feo, S., Harvey, R., Showe, L., Croce, C.M. 1986. Regulation of translocated c-*myc* genes transfected into plasmacytoma cells. Proc. Natl. Acad. Sci. 83: 706-709.

Ffrench-Constant, C. 1992. A division too far? Cur. Biol. 2(11): 577-579.

Fisher, B. and Fisher, E.R. 1966. The interrelationship of hematogenous and lymphatic tumor cell dissemination. Surg. Gynecol. Obstet. 122: 791-798.

Fisher, D.E., Parent, L.A., Sharp, P.A. 1991. TFEB has DNA-binding and oligomerization properties of a unique helix-loop-helix / leucine zipper family. Genes Dev. 5: 2342-2352.

Fisher, D.E., Parent, L.A., Sharp, P.A. 1993. High affinity DNA-binding Myc analogs: recognition by an α-helix. Cell 72: 467-476.

Forman, B.M., Casanova, J., Raaka, B.M., Ghysdael, J., Samuels, H.H. 1992. Half-site spacing and orientation determines whether thyroid hormone and retinoic acid receptors and related factors bind to DNA response

elements as monomers, homodimers, or heterodimers. Mol. Endo. 6: 429-442.

Foster, B.D., Cavener, D.R., Parl, F.F. 1991. Binding analysis of the estrogen receptor to its specific DNA target site in human breast cancer. Cancer Res. 51: 3405-3410.

Foulkes, W., Black, D., Solomon, E., Trowsdale, J., 1991. Allele loss on chromosome-17q in sporadic ovarian-cancer. Lancet 338: 444-445.

Frankel, A.D. and Kim, P.S. 1991. Modular structure of transcription factors: implications for gene regulation. Cell 65: 717-719.

Franks, S., Kiwi, R., Nabarro, J.D.N. 1977. Pregnancy and lactation after pituitary surgery. Br. Med. J. 1: 882-885.

Freytag, S.O. 1988. Enforced expression of the c-*myc* oncogene inhibits cell differentiation by precluding entry into a distinct predifferentiation state in G_0/G_1. Mol. Cell. Biol. 8(4): 1614-1624.

Friesen, H.G. 1973. Human prolactin in clinical endocrinology: The impact of radioimmunoassays. Metabolism 22: 1039-1045.

Fritsch, M., Welch, R.D., Murdoch, F.E., Anderson, I., Gorski, J. 1992. DNA allosterically modulates the steroid binding domain of the estrogen receptor. J. Biol. Chem. 267: 1823-1828.

Fuqua, S.A.W., Fitzgerald, S.D., Chamness, G.C., Tandon, A.K., McDonnell, D.P., Nawaz, Z., O'Malley, B.W., McGuire, W.L. 1991. Variant human breast tumor estrogen receptor with constitutive transcriptional activity. Cancer Res. 51: 105-109.

Furr, B.J.A. and Jordan, V.C. 1984. The pharmacology and clinical uses of tamoxifen. Pharmac. Ther. 25: 127-205.

Futreal, P.A., Soderkvist, P., Marks, J.R., Iglehart, J.D., Cochran, C., Barrett, J.C., Wiseman, R.W. 1992. Detection of frequent allelic loss on

proximal chromosome 17q in sporadic breast carcinoma using microsatellite length polymorphisms. Cancer Res. 52: 2624-2627.

Gaffney, E.V. 1982. A cell line (HBL-100) established from human breast milk. Cell Tissue Res. 227: 563-568.

Garcia, I., Dietrich, P.Y., Aapro, M., Vauthier, G., Vadas, L., Engel, E. 1989. Genetic alterations of c-*myc*, c-*erbB-2* and c-*Ha-ras* proto-oncogenes and clinical associations in human breast carcinomas. Cancer Res. 49: 6675-6679.

Garrard, W.T. 1991. Histone H1 and the conformation of transcriptionally active chromatin. BioEssays 13(2): 87-88.

Gasser, S.M. and Laemmli, U.K. 1987. A glimpse at chromosomal order. Trends Genet. 3: 16-22.

Gaub, M.-P., Bellard, M., Scheuer, I., Chambon, P., Sassone-Corsi, P. 1990. Activation of the ovalbumin gene by the estrogen receptor involves the Fos-Jun complex. Cell 63: 1267-1276.

Gazin, C., Dupont de Dinechin, S., Hampe, A., Masson, J.-M., Martin, P., Stehelin, D., Galibert, F. 1984. Nucleotide sequence of the human c-*myc* locus: provocative open reading frame within the first exon. EMBO J. 3(2): 383-387.

Gazin, C., Rigolet, M., Briand, J.P., Van Regenmortel, M.H., Galibert, F. 1986. Immunochemical detection of proteins related to the human c-*myc* exon 1. EMBO J. 5 (9): 2241-2250.

Getzenberg, R.H. and Coffey, D.S. 1990. Tissue specificity of the hormonal response in sex accessory tissues is associated with nuclear matrix protein patterns. Mol. Endo. 4: 1336-1342.

Getzenberg, R.H., Pienta, K.J., Ward, W.S., Coffey, D.S. 1991. Nuclear structure and the three-dimensional organization of DNA. J. Cell. Biochem. 47: 289-299.

Gibson, J.S. and Croker, B.P. 1992. Chromatin structure changes suggest a compensatory response to c-*myc* gene amplification in malignant histiocytoma. J. Cell. Biochem. 49: 148-156.

Gilbert, D.M., Losson, R., Chambon, P. 1992. Ligand dependence of estrogen receptor induced changes in chromatin structure. Nucleic Acids Res. 20: 4525-4531.

Gillespie, D.A.F. and Eisenman, R.N. 1989. Detection of a Myc-associated protein by chemical cross-linking. Mol. Cell. Biol. 9(2): 865-868.

Girvitz, S.C., Bacchetti, S., Rainbow, A.J., Graham, F.L. 1980. A rapid and efficient procedure for the purification of DNA from agarose gels. Anal. Biochem. 106: 492-496.

Glass, C.K., Holloway, J.M., Devary, O.V., Rosenfeld, M.G. 1988. The thyroid hormone receptor binds with the opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. Cell 54: 313-323.

Goldie, J.H. and Coldman, A.J. 1984. The genetic origin of drug resistance in neoplasms: implication for systemic therapy. Cancer Res. 44: 3643-3653.

Goldman, M.A. 1988. The chromatin domain as a unit of gene regulation. BioEssays 9: 50-55.

Gonzalez, I.L., Gorski, J.L., Campen, T.J., Dorney, D.J., Erickson, J.M., Sylvester, J.E., Schmickel, R.D. 1985. Variation among human 28s ribosomal RNA genes. Proc. Natl. Acad. Sci. 82: 7666-7670.

Goodrich, D.W. and Lee, W-H. 1992. Abrogation by c-*myc* of G1 phase arrest induced by RB protein but not p53. Nature 360: 177-179.

Gorski, J., Furlow, D., Murdoch, F.E., Fritsch, M., Kaneko, K., Ying, C., Malayer, J.R. 1993. Perturbations in the model of estrogen receptor regulation of gene expression. Biology of Reproduction 48: 8-14.

Green, S. and Chambon, P. 1986. A superfamily of potentially oncogenic hormone receptors. Nature 324: 615-617.

Green, S., Kumar, V., Krust, A., Walter, P., Chambon, P. 1986. Structural and functional domains of the estrogen receptor. Cold Spring Harbor Symposia on Quantitative Biology 51: 751-758.

Greene, G.L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., Shine, J. 1986. Sequence and expression of human estrogen receptor complementary DNA. Science 231: 1150-1154.

Gregor, P.D., Sawadogo, M., Roeder, R.C. 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. Genes Dev. 4: 1730-1740.

Grignani, F., Lombardi, L., Inghirami, G., Sternas, L., Cechova, K., Dalla-Favera, R. 1990. Negative autoregulation of c-*myc* gene expression is inactivated in transformed cells. EMBO J. 9: 3913-3922.

Gronemeyer, H. 1991. Transcription activation by estrogen and progesterone receptors. Annu. Rev. Genet. 25: 89-123.

Gross, D.S. and Garrard, W.T. 1987. Poising chromatin for transcription. Trends Biol. Sci. 12: 293-297.

Gross, D.S. and Garrard. W.T. 1988. Nuclease hypersensitive sites in chromatin. Ann. Rev. Biochem. 57: 159-197.

Grosso, L.E. and Pitot, H.C. 1985. Chromatin structure of the c-*myc* gene in HL-60 cells during alteration of transcriptional activity. Cancer Res. 45: 5035-5041.

Grosveld, F., Van Assendelft, G.B., Greaves, D.R., Kollias, G. 1987. Position-independent, high-level expression of human ß-globin gene expression in transgenic mice. Cell 51: 975-985.

Groudine, M., Eisenman, R., Weintraub, H. 1981. Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. Nature 292: 311-317.

Grunstein, M. 1990. Nucleosomes: regulators of transcription. Trends Genet. 6: 395-400

Gupta, S., Seth, A., Davis, R.J. 1993. Transactivation of gene expression by Myc is inhibited by mutation at the phosphorylation sites Thr-58 and Ser-62. Proc. Natl. Acad. Sci. 90: 3216-3220.

Hahn, S. 1993. Structure (?) and function of acidic transcription activators. Cell 72: 481-483.

Halazonetis, T.D. and Kandil, A.N. 1991. Determination of the c-Myc DNA-binding site. Proc. Natl. Acad. Sci. 88: 6162-6166.

Haldar, S., Beatty, C., Tsujimoto, Y., Croce, C.M. 1989. The *bcl-2* gene encodes a novel G protein. Nature 342: 195-198.

Hall, D.J. 1990. Regulation of c-*myc* transcription *in vitro* : dependence on the guanine-rich promoter element ME1a1. Oncogene 5: 47-54.

Hall, J. M., Lee, M.K., Newman, B., Morrow, J.E., Anderson, L.A., Huey, B., King, M.-C. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250: 1684-1689.

Ham, J., Steger, G. Yaniv, M. 1992. How do eukaryotic activator proteins stimulate the rate of transcription by RNA polymerase II? FEBS L. 307: 81-86.

Hann, S.R. and Eisenman, R.N. 1984. Proteins encoded by the human c-*myc* oncogene: differential expression in neoplastic cells. Mol. Cell. Biol. 4: 2486-2497.

Hann, S.R., King, M.W., Bentley, D.L., Anderson, C., Eisenman, R.N. 1988. A non-AUG translational initiation in c-*myc* exon 1 generates a N-terminally distinct polypeptide whose synthesis is disrupted in Burkitt's lymphoma. Cell 52: 185-195.

Hann, S.R., Thompson, C.B., Eisenman, R.N. 1985. c-*myc* oncogene protein synthesis is independent of the cell cycle in human and avian cells. Nature 314: 366-369.

Hansen, J.C. and Ausio, J. 1992. Chromatin dynamics and the modulation of genetic activity. Trends Biochem. Sci. 17: 187-191.

Harrington, M.A., Jones, P.A., Imagawa, M., Karin, M. 1988. Cytosine methylation does not affect binding of transcription factor-Sp1. Proc. Natl. Acad. Sci. 85: 2066-2070.

Harvey, H.A., Lipton, A., Max, D.T., Pearlman, H.G., Diaz-Perches, R., de la Garza, J. 1985. Medical castration produced by the GnRH analogue leuprolide to treat metastatic breast cancer. J. Clin. Oncol. 3: 1068-1072.

Hay, N., Bishop, J.M., Levens, D. 1987. Regulatory elements that modulate expression of human c-*myc*. Genes Dev. 1: 659-671.

Hay, N., Takimoto, M., Bishop, J.M. 1989. A *Fos* protein is present in a complex that binds a negative regulator of *Myc*. Genes Dev. 3: 293-303.

Hayday, A.C., Gillies, S.D., Saito, H., Wood, C., Wiman, K., Hayward, W.S., Tonegawa, S. 1984. Activation of a translocated human c-*myc* gene by an enhancer in the immunoglobulin heavy-chain locus. Nature 307: 334-340.

Hayes, J.J., Clark, D.J., Wolffe, A.P. 1991. Histone contributions to the structure of DNA in the nucleosome. Proc. Natl. Acad. Sci. 88: 6829-6833.

Hayes, J.J. and Wolffe, A.P. 1992. The interaction of transcription factors with nucleosomal DNA. BioEssays 14(9): 597-603.

Hayward, W.S., Neel, B.G., Astrin, S.M. 1981. Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290: 475-480.

Heikkila, R., Schwab, G., Wickstrom, E., Loke, S.L., Plluznik, D.H., Watt, R., Neckers, L.M. 1987. A c-*myc* antisense oligodeoxynucleotide inhibits entry into S phase but not progress from G_0 to G_1. Nature 336: 445-449.

Henderson, B.E., Ross, R.K., Pike, M.C. 1993. Hormonal chemoprevention of cancer in women. Science 259: 633-638.

Henderson, M. 1993. Current approaches to breast cancer prevention. Science 259: 630-632.

Herschlag, D. and Johnson, F.B. 1993. Synergism in transcriptional activation: a kinetic view. Genes Dev. 7: 173-179.

Hiebert, S.W., Chellappan, S.P., Horowitz, J.M., Nevins, J.R. 1992. The interaction of Rb with E2F coincides with an inhibition of the transcriptional activity of E2F. Genes Dev. 6: 177-185.

Hiebert, S.W., Lipp, M., Nevins, J.R. 1989. E1A-dependent *trans*-activation of the human Myc promoter is mediated by the E2F factor. Proc. Natl. Acad. Sci. 86: 3594-3598.

Hoey, T., Weinzierl, R.O.J., Gill, G., Chen, J.-L., Dynlacht, B.D., Tjian, R. 1993. Molecular-cloning and functional-analysis of *Drosophila* Taf110 reveal properties expected of coactivators. Cell 72: 247-260.

Hollenberg, S.M. and Evans, R. 1988. Multiple and cooperative *trans*-activation domains of the human glucocorticoid receptor. Cell 55: 897-906.

Holmes, F.A., Fritsche, H.A., Loewy, J.W., Geitner, A.M., Sutton, R.C., Buzdar, A.U., Hortobagyi, G.N. 1990. Measurement of estrogen and progesterone receptors in human breast tumors: enzyme immunoassay versus binding assay. J. Clin. Oncol. 8: 1025-1035.

Hope, I.A. and Strühl, K. 1985. GCN4 protein, synthesized *in vitro*, binds HIS3 regulatory sequences: implications for general control of amino acid biosynthetic genes in yeast. Cell 43: 177-188.

Horikoshi, M., Hai, T., Lin, Y.-S., Green, M.R., Roeder, R.G. 1988. Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. Cell 54: 1033-1042.

Horwitz, K.B. and McGuire, W.L. 1978. Estrogen control of progesterone receptor in human breast cancer. J. Biol. Chem. 253: 2223-2228.

Howard, A. and Pelc, S.R. 1951. Nuclear incorporation of ^{32}P as demonstrated by autoradiographs. Expl. Cell Res. 2: 178-187.

Hu, Y.-F., Lüscher, B., Admon, A., Mermod, N., Tjian, R. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes Dev. 4: 1741-1752.

Hunter, T. and Karin, M. 1992. The regulation of transcription by phosphorylation. Cell 70: 375-387.

Hutchison, N, and Weintraub, H. 1985. Localization of DNase I-sensitive sequences to specific regions of interphase nuclei. Cell 43: 471-482.

Iguchi-Ariga, S.M.M., Itani, T., Kiji, Y., Ariga, H. 1987. Function of the c-*myc* product - promotion of cellular DNA-replication. EMBO J. 6: 2365-2371.

Iguchi-Ariga, S.M.M., Okazaki, T., Itani, T., Ogata, M., Sato, Y., Ariga, H. 1988. An initiation site of DNA replication with transcriptional enhancer activity present upstream of the c-*myc* gene. EMBO J. 7(10): 3135-3142.

Ikegaki, N., Minna, J., Kennett, R.H. 1989. The human L-*myc* gene is expressed as two forms of protein in small cell lung carcinoma cell lines: detection by monoclonal antibodies specific to two *myc* homology box sequences. EMBO J. 8: 1793-1799.

Imagawa, M., Chiu, R., Karin, M. 1987. Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell 51: 251-260.

Inghirami, G., Grignani, F., Sternas, L., Lombardi, L., Knowles, D., Dalla-Favera, R. 1990. Down-regulation of LFA-1 adhesion receptors by c-*myc* oncogene in human B lymphoblastoid cells. Science 250: 682-686.

Ingvarsson, S., Asker, C., Axelson, H., Klein, G., Sumegi, J. 1988. Structure and expression of B-*myc*, a new member of the *myc* gene family. Mol. Cell. Biol. 8: 3168-3174.

Isotalo, H., Tyrggvason, K., Vierikko, P., Kauppila, A., Vihko, R. 1983. Plasminogen activators and steroid receptor concentrations in normal

benign, and malignant breast and ovarian tissues. Anticancer Res. 3: 331-335.

Jackson, D.A. 1991. Structure-function relationships in eukaryotic nuclei. BioEssays 13: 1-10.

Jacobs, I.J., Smith, S.A., Wiseman, R.W., Futreal, P.A., Harrington, T., Osborne, R.J., Leech, V., Molyneux, A., Berchuck, A., Ponder, B.A.J., Bast, R.C. Jr. 1993. A deletion unit on chromosome 17q in epithelial ovarian tumors distal to the familial breast/ovarian cancer locus. Cancer Res. 53: 1218-1221.

Jakobovits, A., Schwab, M., Bishop, J.M., Martin, G.R. 1985. Expression of N-*myc* in teratocarcinoma stem cells and mouse embryos. Nature 318: 188-191.

Jansen-Dürr, P., Meichle, A., Steiner, P., Pagano, M., Finke, K., Botz, J., Wessbecher, J., Draetta, G., Eilers, M. 1993. Differential modulation of cyclin gene expression by Myc. Proc. Natl. Acad. Sci. 90: 3685-3689.

Jantzen, H.-M., Strähle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R., Schütz, G. 1987. Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell 49: 29-38.

Jenks, S. 1992. RU 486 enters limited clinical trials. J. Natl. Cancer Inst. 84: 563.

Jiang, S.-Y. and Jordan, V.C. 1992. Growth regulation of estrogen receptor negative breast cancer cells transfected with complementary DNAs for estrogen receptor. J. Natl. Cancer Inst. 84: 580-591.

Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J., Tjian, R. 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell 48: 79-89.

Jones, P.A. and Buckley, J.D. 1990. The role of DNA methylation in cancer. Adv. Cancer Res. 54: 1-23.

Jones, T.R. and Cole, M.D. 1987. Rapid cytoplasmic turnover of c-*myc* mRNA: requirement of the 3' untranslated sequences. Mol. Cell. Biol. 7: 4513-4521.

Jordan, V.C. and Murphy, C.S. 1990. Endocrine pharmacology of antiestrogens as antitumor agents. Endocrine Rev. 11: 578-610.

Kaczmarek, L., Hyland, J.K., Watt, R., Rosenberg, M., Baserga, R. 1985. Microinjected c-Myc as a competence factor. Science 228: 1313-1315.

Kadonaga, J.T., Jones, K.A., Tjian, R. 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biol. Sci. 11: 20-23.

Kadonaga, J.T., Courey, A.J., Ladika, J., Tjian, R. 1988. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science 242: 1566-1570.

Kadonaga, J.T. 1990. Gene transcription: basal and regulated transcription by RNA polymerase II. Cur. Opin. Cell Biol. 2: 496-501.

Kakkis, E., Prehn, J., Calame, K. 1986. An active chromatin structure acquired by translocated c-*myc* genes. Mol. Cell. Biol. 6: 1357-1361.

Kakkis, E. and Calame, K. 1987. A plasmacytoma-specific factor binds the c-*myc* promoter region. Proc. Natl. Acad. Sci. 84: 7031-7035.

Kakkis, E., Riggs, K.J., Gillespie, W., Calame, K. 1989. A transcriptional repressor of c-*myc*. Nature 339: 718-721.

Kageyama, R., Merlino, G.T., Pastan, I. 1988. Epidermal growth factor (EGF) receptor gene transcription: requirement for Sp1 and a receptor specific factor. J. Biol. Chem. 263: 6329-6336.

Kageyama, R. and Pastan, I., 1989. Molecular cloning and characterization of a human DNA binding factor that represses transcription. Cell 59: 815-825.

Kamakaka, R.T. and Thomas, J.O. 1990. Chromatin structure of transcritptionally competent and repressed genes. EMBO J. 9: 3997-4006.

Kaplan, C.R. and Schenken, R.S. 1990. Endocrinology of the Breast. In: The female breast and its disorders. pp. 22-44. Mitchell, G.W. and Basset, L.W. eds. Williams & Wilkins, Baltimore, USA.

Karn, J., Watson, J.V., Lowe, A.D., Green, S.M., Vedeckis, W. 1989. Regulation of cell cycle duration by c-*myc* levels. Oncogene 4: 773-787.

Kato, G.J. and Dang, C.V. 1992. Function of the c-Myc oncoprotein. FASEB J. 6: 3065-3072.

Kato, G.J., Barrett, J., Villa-Garcia, M., Dang, C.V. 1989. An amino-terminal c-Myc domain required for neoplastic transformation activates transcription. Mol. Cell. Biol. 10: 5914-5920.

Kato, G.J., Lee, W.M.F., Chen, L., Dang, C.V. 1992a. Max: functional domains and interaction with c-Myc. Genes Dev. 6: 81-92.

Kato, S., Tora, L., Yamauchi, J., Masushige, S., Bellard, M., Chambon, P. 1992b. A far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. Cell 68: 731-742.

Katzenellenbogen, B.S., Kendra, K.L., Norman, M.J., Berthois, Y. 1987. Proliferation, hormonal responsiveness, and estrogen receptor content

of MCF-7 human breast cancer cells grown in the short-term and long-term absence of estrogens. Cancer Res. 47: 4355-4360.

Kaye, J.S., Pratt-Kaye, S., Bellard, M., Dretzen, G., Bellard, F., Chambon, P. 1986. Steroid hormone dependence of four DNase I hypersensitive regions located within the 7000 bp 5' flanking segment of the ovalbumin gene. EMBO J. 5: 277-285.

Kelly, K., Cochran, B.H., Stiles, C.D., Leder, P. 1983. Cell specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. Cell 35: 603-610.

Kerkhoff, E., Bister, K., Klempnauer, K.-H. 1991. Sequence-specific DNA-binding by Myc proteins. Proc. Natl. Acad. Sci. 88: 4324-4327.

Kessler, D.J., Spicer, D.B., La Rosa, F.A., Sonenshein, G.E. 1992. A novel NF-κB element within exon 1 of the murine c-*myc* gene. Oncogene 7: 2447-2453.

Key, T.J.A. and Pike, M.C. 1988. The role of estrogens and progestagens in the epidemiology and prevention of breast-cancer. Eur. J. Cancer Clin. Oncol. 24: 29-43.

Keydar, I., Chen, L., Karby, S., Weiss, F.R., Delarea, J., Radu, M., Chaitcik, S., Brenner, H.J. 1979. Establishment and characterization of a cell line of human breast carcinoma origin. Eur. J. Cancer. 15: 659-670.

Kim, S.-J., Onwuta, U.S., Lee, Y.I., Li, R., Botchan, M.R., Robbins, P.D. 1992. The retinoblastoma gene product regulates Sp1-mediated transcription. Mol. Cell. Biol. 12: 2455-2463.

King, M.W., Blackwood, E.M., Eisenman, R.N. 1993. Expression of two distinct homologues of *Xenopus* Max during early development. Cell Growth Diff. 4: 85-92.

King, M.W., Roberts, J.M., Eisenman, R.N. 1986. Expression of the c-*myc* proto-oncogene during development of *Xenopus laevis*. Mol. Cell. Biol. 6: 4499-4508.

King, W.J. and Greene, G.L. 1984. Monoclonal antibodies localize estrogen receptor in the nuclei of target-cells. Nature 307: 745-747.

Kirschner, M. 1992. The cell-cycle then and now. Trends Biochem. Sci. 17: 281-285.

Kitaura, H., Galli, I., Taira, T., Iguchi-Ariga, S.M.M., Ariga, H. 1991. Activation of c-*myc* promoter by c-*myc* protein in serum starved cells. FEBS L. 290: 147-152.

Klein-Hitpass, L., Ryffel, G.U., Heitlinger, E., Cato, A.C.B. 1988a. A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. Nucleic Acids Res. 16: 647-657.

Klein-Hitpass, L., Kaling, M., Ryffel, G.U. 1988b. Synergism of closely adjacent estrogen-responsive elements increases their regulatory potential. J. Mol. Biol. 201: 537-544.

Klein-Hitpass, L., Tsai, S.Y., Greene, G.L., Clark, J.H., Tsai, M.-J., O'Malley, B.W. 1989. Specific binding of estrogen receptor to the estrogen response element. Mol. Cell. Biol. 9: 43-49.

Klein-Hitpass, L., Tsai, S.Y., Weigel, N.L., Allan, G.F., Riley, D., Rodriguez, R., Schrader, W.T., Tsai, M.-J., O'Malley, B.W. 1990. The progesterone receptor stimulates cell-free transcription by enhancing the formulation of a stable preinitiation complex. Cell 60: 247-257.

Klempnauer, K.-H., Arnold, H., Beidenkapp, H. 1989. Activation of transcription by v-Myb - evidence for two different mechanisms. Genes Dev. 3: 1582-1589.

Klock, G., Strähle, U., Schutz, G. 1987. Oestrogen and glucocorticoid responsive elements are closely related but distinct. Nature 329: 734-736.

Klug, A. and Rhodes, D. 1987. 'Zinc fingers': a novel protein motif for nucleic acid recognition. Trends Biochem. Sci. 12: 464-469.

Kohl, N.E. and Ruley, H.R. 1987. Role of c-*myc* in the transformation of REF52 cells by viral and cellular oncogenes. Oncogene 2: 41-48.

Kohl, N.E., Kanda, N., Schreck, R. R., Bruns, G., Latt, S.A., Gilbert, F., Alt, F.W. 1983. Transposition and amplification of oncogene-related sequences in human neuroblastomas. Cell 35: 359-367.

Kohl, N.E., Legouy, E., DePinho, R.A., Nisen, P.D., Smith, R.K., Gee, C.E., Alt, F.W. 1986. Human N-*myc* is closely related in organization and nucleotide sequence to c-*myc*. Nature 319: 73-77.

Kornberg, R. 1981. The location of nucleosomes in chromatin - specific or statistical. Nature 292: 579-580.

Kornberg, R.D. and Lorch, Y. 1991. Irresistible force meets immovable object: transcription and the nucleosome. Cell 67: 833-836.

Kozcor, D. and Croce, C.M. 1984. Amplification of the c-*myc* oncogene in one of five human breast carcinoma cell lines. Cancer Res. 44: 438-441.

Kretzner, L., Blackwood, E.M., Eisenman, R.N. 1992. Myc and Max proteins possess distinct transcriptional activities. Nature 359: 426-429.

Krumm, A., Meulia, T., Brunvand, M., Groudine, M. 1992. The block to transcriptional elongation within the human c-*myc* gene is determined in the promoter-proximal region. Genes. Dev. 6: 2201-2213.

Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J.M., Chambon, P. 1986. The chicken oestrogen receptor sequence:

homology with v-erbA and the human oestrogen and glucocorticoid receptors. EMBO J. 5: 891-897.

Krystal, G.W., Armstrong, B.C., Battey, J.F. 1990. N-*myc* mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. Mol. Cell. Biol. 10: 4180-4191.

Kumar, S. and Leffak, M. 1989. DNA topology of the ordered chromatin domain 5' to the human c-*myc* gene. Nucl. Acids Res. 17: 2819-2833.

Kumar, S. and Leffak, M. 1991. Conserved chromatin structure in c-*myc* 5'-flanking DNA after viral transduction. J. Mol. Biol. 222: 45-57.

Kumar, V. and Chambon, P. 1988. The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. Cell 55: 145-156.

Kumar, V., Green, S., Stack, G., Berry, M., Jin, J.-R., Chambon, P. 1987. Functional domains of the human estrogen receptor. Cell 51: 941-951.

Kushner, P.J., Hort, E., Shine, J., Baxter, J.D., Greene, G.L. 1990. Construction of cell lines that express high levels of the human estrogen receptor and are killed by estrogens. Mol. Endocrinol. 4: 1465-1473.

Kutoh, E., Stromstedt, P.-E., Poellinger, L. 1992. Functional interference between the ubiquitous and constitutive octamer transcription factor 1 (OTF-1) and the glucocorticoid receptor by direct protein-protein interaction involving the homeo subdomain of OTF-1. Mol. Cell. Biol. 12: 4960-4969.

Kyprianou, N., English, H.F., Davidson, N.E., Isaacs, J.T. 1991. Programmed cell death during regression of the MCF-7 human breast cancer following estrogen ablation. Cancer Res. 51: 162-166.

Laird-Offringa, I.A., Eifferich, P., Knaken, H.J., de Ruiter, J., van der Eb, A.J. 1989. Analysis of polyadenylation site usage of the c-*myc* oncogene. Nucleic Acids Res. 17: 6499-6513.

Land, H., Parada, L., Weinberg, R.A. 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 304: 596-602.

Landschulz, W.H., Lohnson, P.F., McKnight, S.L. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240: 1759-1764.

Lang, J.C., Whitelaw, B., Talbot, S. and Wilkie, N.M. 1988. Transcriptional regulation of the human c-*myc* gene. Br. J. Cancer 58: Suppl IX, 62-66.

Lang, J.C., Wilkie, N.M., Clark, A.M., Chudleigh, A., Talbot, S., Whitelaw, B., Frame, M.C. 1991. Regulatory domains within the P_0 promoter of human c-*myc*. Oncogene 6 (11): 2067-2075.

Lasfargues, E.Y. and Ozzello, L. 1958. Cultivation of human breast carcinomas. J. Natl. Cancer Inst. 21: 1131-1147.

Laybourn, P.J. and Kadonaga, J.T. 1991. Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science 254: 238-245.

Leder, A., Pattengale, P.K., Kuo, A., Stewart, T.A., Leder, P. 1986. Consequence of widespread deregulation of the c-*myc* gene in transgenic mice - multiple neoplasms and normal development. Cell 45: 485-495.

Leder, P., Battey, J., Lenoir, G., Moulding, C., Murphy, W., Potter, H., Stewart, T., Taub, R. 1983. Translocations among the antibody genes in human cancer. Science 222: 765-771.

Lee, M.-S. and Garrard, W.T. 1991. Transcription-induced nucleosome 'splitting': an underlying structure for DNase I sensitive chromatin. EMBO J. 10: 607-615.

Lee, M.-S. and Garrard, W.T. 1992. Uncoupling gene activity from chromatin structure: promoter mutations can inactivate transcription of the yeast *HSP82* gene without eliminating nucleosome-free regions. Proc. Natl. Acad. Sci. 89: 9166-9170.

Lee, W.-H., Murphree, A.L., Benedict, W.F. 1984. Expression and amplification of the N-*myc* gene in primary retinoblastoma. Nature 309: 458-460.

Lee, W.M.F., Schwab, M., Westaway, D., Varmus, H.E. 1985. Augmented expression of normal c-*myc* is sufficient for cotransformation of rat embryo cells with a mutant *ras* gene. Mol. Cell. Biol. 5: 3345-3356.

Legouy, E., DePinho, R., Zimmerman , K., Collum, R., Yancopoulos, G., Mitsock, R., Kriz, R., Alt, F.W. 1987. Structure and expression of the murine L-*myc* gene. EMBO J. 6: 3359-3366.

Lehrach, H., Diamond, D., Wozney, J.M., Boedtker, H. 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical re-examination. Biochem. 16: 4743-4751.

Lenardo, M., Rustgi, A., Schievella, A., Bernards, R. 1989. Suppression of MHC class I gene expression by N-*myc* through enhancer inactivation. EMBO J. 8: 3351-3355.

Levine, A., Momand, J., Finlay, C.A. 1991. The p53 tumor suppressor gene. Nature 351: 453-456.

Lewin, B. 1990. Commitment and activation at Pol II promoters - a tail of protein-protein interactions. Cell 61: 1161-1164.

Li, R., Knight, J.D., Jackson, S.P., Tjian, R., Botchan, M.R. 1991. Direct interaction between Sp1 and the BPV enhancer E2-protein mediates synergistic activation of transcription. Cell 65: 493-505.

Lieberman, M.E., Maurer, R.A., Gorski, J. 1978. Estrogen control of prolactin synthesis *in vitro*. Proc. Natl. Acad. Sci. 74: 5946-5949.

Lin, Y.-S., Carey, M., Ptashne, M., Green, M.R. 1990. How different eukaryotic transcriptional activators can cooperate promiscuously. Nature 345: 359-361.

Lin, Y.-S. and Green, M.R. 1991. Mechanism of action of an acidic transcriptional activator *in vitro*. Cell 64: 971-981.

Linial, M. and Groudine, M. 1985. Transcription of three c-*myc* exons is enhanced in chicken bursal lymphoma cell-lines. Proc. Natl. Acad. Sci. 82: 53-57.

Lipp, M., Schilling, R., Bernhardt, G. 1989. *Trans*-activation of human Myc: The second promoter is target for the stimulation by adenovirus E1a proteins. Oncogene 4: 535-541.

Lipp, M., Schilling, R., Weist, S., Laux, G., Bornkamm, G.W. 1987. Target sequences for *cis*-acting regulation within the dual promoter of the human c-*myc* gene. Mol. Cell. Biol. 7(4): 1393-1400.

Lippman, M.E. 1993. The development of biological therapies for breast cancer. Science 259: 631-632.

Lippman, M., Bolan, G., Huff, K. 1976. Effects of estrogens and antiestrogens on hormone-responsive human breast-cancer in long term tissue-culture. Cancer Res. 36: 4595-4601.

Littlewood, T.D., Amati, B., Land, H., Evan, G.I. 1992. Max and c-Myc/Max DNA binding activities in cell extracts. Oncogene 7: 1783-1792.

Liu, E., Santos, G., Lee, W.M.F., Osborne, C.K., Benz, C.C. 1989. Effects of c-*myc* overexpression on the growth characteristics of MCF-7 human breast cancer cells. Oncogene 4: 979-984.

Lobanenkov, V.V., Nicolas, R.H., Plumb, M.A., Wright, C.A., Goodwin, G.H. 1986. Sequence specific DNA-binding proteins which interact with (G+C)-rich sequences flanking the chicken c-*myc* gene. Eur. J. Biochem. 159: 181-188.

Lobanenkov, V.V., Nicolas, R.H., Adler, V.V., Paterson, H., Klenova, E.M., Polotskaja, A.V., Goodwin, G.H. 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-*myc* gene. Oncogene 5: 1743-1753.

Logan, W.P. 1975. Cancer of the breast: no decline in mortality. WHO Chronicle 29: 462-471.

Luisi, B.F., Xu, W.X., Otwinowski,Z., Freedman, L.P., Yamamoto, K.R., Sigler, P.B. 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 352: 497-505.

Lumpkin, M.D., Samson, W.K., McCann, S.M. 1983. Hypothalamic and pituitary sites of action of oxytocin to alter prolactin secretion in the rat. Endocrinol. 112: 1711-1717.

Lüscher, B. and Eisenman, R.N. 1990. New light on Myc and Myb. Part 1. Myc. Genes and Dev. 4: 2025-2035.

Lüscher, B., Kuenzel, E.A., Kriebs, E.G., Eisenman, R.N. 1989. Myc oncoproteins are phosphorylated by casein kinase II. EMBO J. 8: 1111-1119.

Lyons, W.R. 1958. Hormonal synergism in mammary growth. Proc. Roy. Soc. London. 149: 303-325.

Mader, S., Kumar, V., de Verneuil, H., Chambon, P. 1989. Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element. Nature 338: 271-274.

Maheswaran, S., McCormack, J.E., Sonenshein, G.E. 1991. Changes in phosphorylation of *myc* oncogene and *RB* antioncogene protein products during growth arrest of the murine lymphoma WEHI 231 cell line. Oncogene 6: 1965-1971.

Mäkelä, T.P., Koskinen, P.J., Västrik, I., Alitalo, K. 1992. Alternative forms of Max as enhancers or suppressors of Myc-Ras cotransformation. Science 256: 373-377.

Maminta, M.L.D. , Molteni, A., Rosen, S.T. 1991. Stable expression of the human estrogen recptor in HeLa cells by infection: effect of estrogen on cell proliferation and c-*myc* expression. Mol. Cell. Endocrinol. 78: 61-69.

Mango, S.E., Schuler, G.D., Steele, M.E., Cole, M.D. 1989. Germ-line c-*myc* is not down-regulated by loss or exclusion of activating factors in *myc*-induced macrophage tumors. Mol. Cell. Biol. 9: 3482-3490.

Marcu, K.B., Asselin, C., Nepveu, A., Weisinger, G., Yang, J.Q. 1988. Negative control elements within and near the murine c-*myc* gene. Curr. Topics Microbiol. Immunol. 141: 253-263.

Marcu, K.B., Bossone, S.A., Patel, A.J. 1992. *myc* function and regulation. Annu. Rev. Biochem. 61: 809-860.

Mariani-Costantini, R., Escot, C., Theillet, C., Gentile, A., Merlo, G., Lidereau, R., Callahan, R. 1988. *In situ* c-*myc* expression and genomic status of the c-*myc* locus in infiltrating ductal carcinoma. Cancer Res. 48: 199-205.

Marshall, E. 1993a. The politics of breast cancer. Science 259: 616-617.

Marshall, E. 1993b. Search for a killer: focus shifts from fat to hormones. Science 259: 618-621.

Martin, K.J., Lillie, W.J., Green, M.R. 1990. Evidence for interaction of different eukaryotic transcriptional activators with distinct cellular targets. Nature 346: 147-152.

Martin, K.J. 1991. The interactions of transcription factors and their adaptors, coactivators and accessory proteins. BioEssays 13: 499-503.

Martinez, E. and Wahli, W. 1989. Cooperative binding of estrogen receptor to imperfect estrogen-responsive DNA elements correlates with their synergistic hormone-dependent enhancer activity. EMBO J. 8: 3781-3791.

Martinez, E., Dusserre, Y., Wahli, W., Mermod, N. 1991. Synergistic transcriptional activation by CTF/NF-I and the estrogen receptor involves stabilized interactions with a limiting target factor. Mol. Cell. Biol. 11: 2937-2945.

Marx, J. 1993. Cellular changes on the route to metastasis. Science 259: 626-629.

Maxam, A. and Gilbert, W. 1980. Sequencing end-labelled DNA with base specific chemical cleavages. Method. Enzymol. 65: 499-560.

McGuire, W.L., Carbone, P.P., Vollmer, R.P. (eds). 1975. Estrogen receptors in human breast cancer. Raven Press, New York.

McGuire, W.L. 1980. Steroid hormone receptors in breast cancer treatment strategy. Rec. Prog. Horm. Res. 36: 135-145.

McGuire, W.L., Lippman, M.E., Osborne, C.K., Thompson, E.B. 1987. Resistance to endocrine therapy, a panel discussion. Breast Cancer Res. Treat. 9: 165-173.

McMichael, A.J. and Giles, G.G. 1988. Cancer in migrants to Australia - extending the descriptive epidemiological data. Cancer Res. 48: 751-756.

McWhinney, C. amd Leffak, M. 1990. Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-*myc* gene. Nucleic Acids Res. 18: 1233-1242.

Medici, N., Nigro, V., Abbondanza, C., Moncharmont, B., Molinari, A.M., Puca, G.A. 1991. *In vitro* binding of the purified subunit of the estrogen receptor to oligonucleotides containing natural or modified sequences of an estrogen-responsive element. Mol. Endo. 5: 555-563.

Mermod, N., O'Neill, E.A., Kelly, T.J., Tjian, R. 1989. The proline-rich transcriptional activator of CTF/NF1 is distinct from the replication and DNA binding domain. Cell 58: 741-753.

Meyer, M.-E. Gronmeyer, H., Turcotte, B., Bocquel, M.-T., Tasset, D., Chambon, P. 1989. Steroid hormone receptors compete for factors that mediate their enhancer function, Cell 57: 433-442.

Miller, B.A., Feuer, E.J., Hankey, B.F. 1991. The increasing incidence of breast cancer since 1982: relevance of early detection. Cancer Causes and Control 2: 67-74.

Miller, D.R., Rosenberg, L., Kaufman, D.W., Stolley, P., Shapiro, S. 1989a. Breast cancer before age 45 and oral contraceptive use: new findings. Am. J. Epidemiol. 129: 269-280.

Miller, H., Asselin, C., Dufort, D., Yang, J.-Q., Gupta., K., Marcu, K.B., Nepveu, A. 1989b. A *cis*-acting element in the promoter region of the murine c-*myc* gene is necessary for transcriptional block. Mol. Cell. Biol. 9: 5340-5349.

Miller, J., McLachlan, A.D., Klug, A. 1985. Repetitive zinc binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. 4: 1609-1614.

Miller, T.L, Huzel, N.J, Murphy, L.C., Davie, J.R. 1993. c-*myc* gene chromatin of estrogen receptor positive and negative breast cancer cells. Mol. Cell Endo. 91: 83-39.

Min, S. and Taparowsky, E.J. 1992. v-Myc, but not Max, possesses domains that function in both transcription activation and cellular transformation. Oncogene 7: 1531-1540.

Miner, J.H. and Wold, B.J. 1991. c-*myc* inhibition of MyoD and Myogenin initiated myogenic differentiation. Mol. Cell Biol. 11(5): 2842-2851.

Mirkovitch, J., Mirault, M.-E., Laemmli, U.K. 1984. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell 39: 223-232.

Mitchell, P.J. and Tjian, R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins. Science 245: 371-378.

Miyagi, M., Inazawa, J., Takita, K.-I., Nakamura, Y. 1992. Cloning and characterization of an interstitial deletion at chromosome 11p15 in a sporadic breast cancer. Hum. Mol. Gen. 1(9): 705-708.

Moberg, K.H., Logan, T.J., Tyndall, W.A., Hall, D.J. 1992. Three distinct elements within the murine c-*myc* promoter are required for transcription. Oncogene 7: 411-421.

Moberg, K.H., Tyndall, W.A., Pyrc, J., Hall, D.J. 1991. Analysis of the c-*myc* P_2 promoter. J. Cell. Physiol. 148: 75-84.

Modak, S.P., Principaud, E., Spohr, G. 1993. Regulation of *Xenopus* c-*myc* promoter activity in oocytes and embryos. Oncogene 8: 645-654.

Mudryj, M., Devoto, S.H., Hiebert, S.W., Hunter, T., Pines, J., Nevins, J.R. 1991. Cell cycle regulation of the E2F transcritpion factor involves an interaction with cyclin A. Cell 65: 1243-1253.

Mugrauer, G. and Ekblom, P. 1991. Contrasting expression patterns of three members of the *myc* family of protooncogenes in the developing and adult mouse kidney. J. Cell Biol. 112(1): 13-25.

Münzel, P., Marx, D., Köchel, H., Schauer, A., Bock, K.W. 1991. Genomic alterations of the c-*myc* protooncogene in relation to the overexpression of c-*erbB2* and *Ki-67* in human breast and cervix carcinomas. J. Cancer Res. Clin. Oncol. 117: 603-607.

Murphy, L.C. and Dotzlaw, H. 1989a. Variant estrogen receptor mRNA species detected in human breast cancer biopsy samples. Mol. Endocrinol. 3: 687-693.

Murphy, L.C. and Dotzlaw, H. 1989b. Endogenous growth factor expression in T-47D, human breast cancer cells, associated with reduced sensitivity to antiproliferative effects of progestins and antiestrogens. Cancer Res. 49: 599-604.

Murphy, L.C. and Dotzlaw, H. 1989c. Regulation of transforming growth factor alpha and transforming growth factor beta messenger ribonucleic acid abundance in T-47D, human breast cancer cells. Mol. Endocrinol. 3: 611-617.

Murphy, L.C., Dotzlaw, H., Wong, M.S.J., Miller, T., Murphy, L.J. 1991. Mechanisms involved in the evolution of progestin resistance in human breast cancer cells. Cancer Res. 51: 2051-2057.

Murphy, L.C., Murphy, L.J., Tsuyuki, D., Duckworth, M.L., Shiu, R.P. 1988. Cloning and characterization of a cDNA encoding a highly conserved,

putative calcium binding protein, identified by an anti-prolactin receptor antiserum. J. Biol. Chem. 263: 2397-2401.

Murphy L.J., Murphy L.C., Friesen H.G. 1987. Estrogen induction of N-*myc* and c-*myc* proto-oncogene expression in the rat uterus. Endocrinology 120: 1882-1888.

Murré, C., McCaw, P.S., Baltimore, D. 1989a. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, myoD, and myc proteins. Cell 56: 777-784.

Murré, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., Weintraub, H., Baltimore, D. 1989b. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 58: 537-544.

Näär, A.M., Boutin, J.-M., Lipkin, S.M., Yu, V.C., Holloway, J.M., Glass, C.K., Rosenfeld, M.G. 1991. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell 65: 1267-1279.

Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Kaprov, V.L., Ebralidse, K.K., Mirzabekov, A.D. 1989. Change in the pattern of histone binding to DNA upon transcriptional activation. Cell 58: 27-36.

Nakagoshi, H., Kanei-Ishii, C., Sawazaki, T., Mizuguchi, G., Ishii, S. 1992. Transcriptional activation of the c-*myc* gene by the c-*myb* and B-*myb* gene products. Oncogene 7: 1233-1240.

Nandi, S. 1958. Endocrine control of mammary-gland development and function in the C3H/He Crgl mouse. J. Natl. Cancer Inst. 21: 1039-1063.

Narod, S.A., Feunteun, J., Lynch, H.T., Watson, P., Conway, T., Lynch, J., Lenoir, G.M. 1991. Familial breast-ovarian cancer locus on chromosome 17q12-23. Lancet 338: 82-83.

Nau, M.M., Brooks, B.J., Battey, J., Sausville, E., Gazdar, A.F., Kirsch, I., McBride, O.W., Bertness, W., Hollis, G.F., Minna, J.D. 1985. L-*myc*, a new *myc*-related gene amplified and expressed in human small cell lung cancer. Nature 318: 69-73.

Nau, M.M., Brooks, B.J., Carney, D.N., Gazdar, A.F., Battey, J.F., et al. 1986. Human small-cell lung cancers show amplification and expression of the N-*myc* gene. Proc. Natl. Acad. Sci. 83: 1092-1096.

Nayfield, S.G., Karp, J.E., Ford, L.G., Dorr, F.A., Kramer, B.S. 1991. Potential role of tamoxifen in prevention of breast cancer. J. Natl. Cancer Inst. 83: 1450-1459.

Neilson, P.V., Pederson, H., Kampmann, E.-M. 1979. Absence of human placental lactogen in an otherwise uneventful pregnancy. Am. J. Obstet. Gyneco. 135: 322-330.

Nelson, C., Albert, V.R., Elsholtz, H.P., Lu, L. I-W., Rosenfeld, M.G. 1988. Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. Science 239: 1400-1405.

Nelson, C., Crenshaw III, E.B., Franco, R., Lira, S.A., Albert, V.R., Evans, R.M., Rosenfeld, M.G. 1986. Discrete *cis*-active genomic sequences dictate the pituitary cell type-specific expression of rat prolactin and growth hormone genes. Nature 322: 557-562.

Nenci, I., Marchetti, E., Querzoli, P. 1988. Commentary on human mammary preneoplasia. The estrogen receptor-promotion hypothesis. J. Steroid. Biochem. 30: 105-106.

Nepveu, A. and Marcu, K.B. 1986. Intragenic pausing and antisense transcription within the murine c-*myc* locus. EMBO J. 5: 2859-2865.

Nepveu, A., Levine, R.A., Campisi, J., Greenberg, M.E., Ziff, E.B., Marcu, K.B. 1987. Alternative modes of c-*myc* regulation in growth factor-stimulated and differentiating cells. Oncogene 1: 243-250.

Nishikura, K. 1986. Sequences involved in accurate and efficient transcription of human c-*myc* genes microinjected into frog oocytes. Mol. Cell. Biol. 6: 4093-4098.

Nishizuka, Y. 1984. The role of protein kinase C in cell surface signal transduction and tumor promotion. Nature 308: 693-698.

Nottenberg, C. and Varmus, H.E. 1986. Features of the chicken c-*myc* gene that influence the structure of c-*myc* RNA in normal cells and bursal lymphomas. Mol. Cell. Biol. 6: 2800-2806.

Ohmori, Y., Tanabe, J., Takada, S., Lee, W.M.F., Obinata, M. 1993. Functional domains of c-Myc involved in the commitment and differentiation of murine erythroleukemia cells. Oncogene 8: 379-386.

Oka, T., Perry, J.W., Topper, Y.J. 1974. Changes in insulin responsiveness during development of mammary epithelium. J. Cell Biol. 62: 550-556.

O'Neil, K.T., Hoess, R.H., DeGrado, W.F. 1990. Design of DNA-binding peptides based on the leucine zipper motif. Science 249: 774-778.

O'Shea, E.K., Rutkowski, R.H., Stafford III, W.S., Kim, P.S. 1989. Evidence that the leucine zipper is a coiled coil. Science 243: 538-541.

Ono, T., Takahashi, N., Okada, S. 1989. Age-associated changes in DNA methylation and messenger-RNA level of the c-*myc* gene in spleen and liver of mice. Mutat. Res. 219: 39-50.

Papoulas, O., Williams, N.G., Kingston, R.E. 1992. DNA binding activities of c-*myc* purified from eukaryotic cells. J. Biol. Chem. 267: 470-480.

Pardee, A.B. 1974. A restriction point for control of normal animal cell proliferation. Proc. Natl. Acad. Sci. 71: 1286-1290.

Parker, M.G. 1988. The expanding family of nuclear hormone receptors. J. Endocr. 119: 175-177.

Partin, A.W., Getzenberg, R.H., CarMichael, M.J., Vindivich, D., Yoo, J., Epstein, J.I., Coffey, D.S. 1993. Nuclear matrix protein patterns in human prostatic hyperplasia and prostate cancer. Cancer Res. 53: 744-746.

Pascal, E. and Tjian, R. 1991. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev. 5: 1646-1656.

Paulson, J.R. and Laemmli, U.K. 1977. The structure of histone-depleted metaphase chromosomes. Cell 12: 817-828.

Pavelic, Z.P., Pavelic, K., Carter, C.P., Pavelic, L. 1992. Heterogeneity of c-*myc* expression in histologically similar infiltrating ductal carcinomas of the breast. J. Cancer Res. Clin. Oncol. 118: 16-22.

Payne, G.S., Bishop, J.M., Varmus, H.E. 1982. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. Nature 295: 209-214.

Peaker, M. 1980. The effect of raised intra-mammary pressure on mammary function in the goat in relation to the cessation of lactation. J. Physiol. 301: 415-428.

Pei, R. and Calame, K. 1988. Differential stability of c-*myc* mRNAs in a cell-free system. Mol. Cell. Biol. 8: 2860-2868.

Pelicci, P., Knowles, D.M., Magrath, I., Dalla-Favera, R. 1986. Chromosomal breakpoints and structural alterations in the c-*myc* locus differ in endemic and sporadic forms of Burkitt lymphoma. Proc. Natl. Acad. Sci. 83: 2984-2988.

Penn, L.J.Z., Brooks, M.W., Laufer, E.M., Littlewood, T.D., Morgenstern, J.P., Evan, G.I., Lee, W.M., Land, H. 1990a. Domains of human c-*myc* protein required for autosuppression and cooperation with *ras* oncogenes are overlapping. Mol. Cell. Biol. 10(9): 4961-4966.

Penn, L.J.Z., Laufer, E.M., Land, H. 1990b. c-*myc*: Evidence for multiple regulatory functions. Sem. Cancer Biol. 1: 69-80.

Penn, L.J.Z., Brooks, M.W., Laufer, E.M., Land, H. 1990c. Negative autoregulation of c-*myc* transcription. EMBO J. 9: 1113-1121.

Perlmann, T. and Wrange, O. 1988. Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. EMBO J. 7: 3073-3079.

Perry, V.H., Henderson, Z., Linden, R. 1983. Postnatal changes in retinal ganglion cell and optic axon populations in the pigmented rat. J. Comp. Neurol. 219: 356-368.

Persson, H. and Leder, P. 1984. Nuclear localization and DNA binding properties of a protein expressed by human c-*myc* oncogene. Science 225: 718-720.

Persson, H., Hennighausen, L., Taub, R., DeGrado, W., Leder, P. 1984. Antibodies to human c-*myc* oncogene product: evidence of an evolutionarily conserved protein induced during cell proliferation. Science 225: 687-693.

Persson, I., Yuen, J. Bergkvist, L., Adami, H.O., Hoover, R., Schairer, C. 1992. Combined estrogen-progestogen replacement and breast cancer risk. Lancet 340: 1044.

Peterson, O.W., Rønnov-Jessen, L., Howlett, A.R., Bissell, M.J. 1992. Interaction with basement membrane serves to rapidly distinguish growth and differentiation patterns of normal and malignant human breast epithelial cells. Proc. Natl. Acad. Sci. 89: 9064-9068.

Pettersson, S., Cook, G.P., Brüggerman, M., Williams, G.T., Neuberger, M.S. 1990. A second B-cell-specific enhancer 3' of the immunoglobulin heavy-chain locus Nature 344: 165-168.

Pfeifer-Ohlsson, S., Rydnert, J., Goustin, A.S., Larsson, E., Betsholz, C., Ohlsson, R. 1985. Cell-type specific pattern of *myc* proto-oncogene expression in developing human embryos. Proc. Natl. Acad. Sci. 82: 5050-5054.

Pham, T.A., Elliston, J.F., Nawaz, Z., McDonnell, D.P., Tsai, M.-J., O'Malley, B.W. 1991a. Antiestrogen can establish nonproductive receptor complexes and alter chromatin structure at target enhancers. Proc. Natl. Acad. Sci. 88: 3125-3129.

Pham, T.A., Hwung, Y.-P., McDonnell, D.P., O'Malley. B.W. 1991b. Transactivation functions facilitate the disruption of chromatin structure by estrogen receptor derivatives *in vivo*. J. Biol. Chem. 266: 18179-18187.

Phi-Van, L. and Strätling, W.H. 1988. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J. 7: 655-664.

Phi-Van, L., von Kries, J.P., Ostertag, W., Strätling, W.H. 1990. The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol. Cell. Biol. 10(9): 2302-2307.

Piechaczyk, M., Yang, J.-Q., Blanchard, J.-M., Janteur, P., Marcu, K.B. 1985. Post-transcriptional mechanisms are responsible for accumulation of truncated c-*myc* mRNAs in murine plasma cell tumors. Cell 42: 589-597.

Piechaczyk, M., Blanchard, J.-M., Jeanteur, P. 1987. c-*myc* gene regulation still holds its secret. Trends Genet. 3: 47-51.

Pissott, L.E. and Nandi, S. 1961. Experimental induction of mammogenesis and lactogenesis in the dwarf mouse. Acta Endocrinol. 37: 161-175.

Plet, A., Tourkine, N., Mechti, N., Jeanteur, P., Blanchard, J.-M. 1992. *In vivo* footprints between the murine c-*myc* P_1 and P_2 promoters. Oncogene 7: 1847-1851.

Postel, E.H., Mango, S.E., Flint, S.J. 1989. A nuclease-hypersensitive element of the human c-*myc* promoter interacts with a transcription initiation factor. Mol. Cell. Biol. 9(11): 5123-5133.

Postel, E.H., Berberich, S.J., Flint, S.J., Ferrone, C.A. 1993. Human c-*myc* transcription factor PuF identified as NM23-H2 nucleoside diphophate kinase, a candidate suppressor of tumor metastasis. Science 261: 478-480.

Ponglikitmongkol, M., Green, S., Chambon, P. 1988. Genomic organization of the human estrogen receptor gene. EMBO J. 7: 3385-3388.

Ponglikitmongkol, M., White, J.H., Chambon, P. 1990. Synergistic activation of transcription by the human estrogen receptor bound to tandem responsive elements. EMBO J. 9: 2221-2231.

Prendergast, G.C. and Cole, M.D. 1989. Post-transcriptional regulation of cellular gene expression by the c-*myc* oncogene. Mol. Cell. Biol. 9: 124-134.

Prendergast, G.C. and Ziff, E.B. 1991. Methylation sensitive sequence specific DNA binding by the c-Myc basic region. Science 251: 186-189.

Prendergast, G.C. and Ziff, E.B. 1992. A new bind for Myc. Trends Genet. 8(3): 91-96.

Prendergast, G.C., Hopewell, R., Gorham, B.J., Ziff, E.B. 1992. Biphasic effect of Max on Myc cotransformation activity and dependence on amino-terminal and carboxy-terminal Max functions. Gene Dev. 6: 2429-2439.

Prendergast, G.C., Lawe, D., Ziff, E.B. 1991. Association of Myn, the murine homolog of Max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell 65: 395-407.

Preston-Marin, S., Pike, M.C., Ross, R.K., Jones, P.A., Henderson, B.E. 1990. Increased cell division as a cause of human cancer. Cancer Res. 50: 7415-7421.

Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. Nature 322: 697-701.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. Nature 335: 683-689.

Ptashne, M. and Gann, A.F. 1990. Activators and targets. Nature 346: 329-331.

Pugh, B.F. and Tjian, R. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell 61: 1187-1197.

Pyrc, J.J., Moberg, K.H., Hall, D.J. 1992. Isolation of a novel cDNA encoding a zinc-finger protein that binds to two sites within the c-*myc* promoter. Biochem. 31: 4102-4110.

Rabbitts, T.H., Forster, A., Hamlyn, P., Baer, R. 1984. Effect of somatic mutation within translocated c-*myc* genes in Burkitt's lymphoma. Nature 309: 592-597.

Rabbitts, T.H., Hamlyn, P.H., Baer, R. 1983. Altered nucleotide sequence of a translocated c-*myc* gene in Burkitt-lymphoma. Nature 306: 760-765.

Radloff, R., Bauer, W., Vinograd, J. 1967. A dye buoyant density method for the detection and isolation of closed circular duplex DNA: closed circular DNA in HeLa cells. Proc. Natl. Acad. Sci. 57: 1514-1521.

Ramsay, G., Evan, G.I., Bishop, J.M. 1984. The protein encoded by the human proto-oncogene *myc*. Proc. Natl. Acad. Sci. 81: 7742-7746.

Ramsay, G., Stanton, L., Schwab, M., Bishop, J.M. 1986. Human proto-oncogene N-*myc* encodes nuclear proteins that bind DNA. Mol. Cell. Biol. 6(12): 4450-4457.

Ramzy, I. 1990. Pathology of malignant neoplasms. In: The female breast and its disorders. pp 100-119. Mitchell, G.W. and Bassett, L.W. eds. Williams & Wilkins. Baltimore, USA.

Ray, D. and Robert-Lézénès, J. 1989. Coexistence of a c-*myc* messenger RNA initiated in intron-1 with the normal c-*myc* messenger RNA and similar regulation of both transcripts in mammalian cells. Oncog. Res. 5: 73-78.

Ray, R. and Miller, D.M. 1991. Cloning and characterization of a human c-*myc* promoter-binding protein. Mol. Cell. Biol. 9: 5123-5133.

Raynaud, A. 1950. Recherches expérimentales sur le développement de l'appareil génital et le fonctionnement des glandes endocrines des foetus de souris et de mulot. Arch. Anat. Microsc. Morphol. Exp. 39: 518-576.

Razin, S.V., Kekelidze, M.G., Lukanidin, E.M., Scherrer, K., Georgiev, G.P. 1986. Replication origins are attached to the nuclear skeleton. Nucleic Acids Res. 14: 8189-8207.

Reddel, R.R., Alexander, I.E., Koga, M., Shine, J., Sutherland, R.L. 1988. Genetic instability and the development of steroid hormone insensitivity in cultured T 47D human breast cancer cells. Cancer Res. 48: 4340-4347.

Reeves, R. 1984. Transcriptionally active chromatin. Biochem. Biophys. Acta 782: 343-393.

Reik, A., Schütz, G., Stewart, A.F. 1991. Glucocorticoids are required for establishment and maintenance of an alteration in chromatin structure: induction leads to a reversible disruption of nucleosomes over an enhancer. EMBO J. 10: 2569-2576.

Reisman, D., Elkind, N.B., Roy, B., Beamon, J., Rotter, V. 1993. c-Myc *trans*-activates the *p53* promoter through a required downstream CACGTG motif. Cell Growth Diff. 4: 57-65.

Remmers, E.F., Yang, J.-Q., Marcu, K.B. 1986. A negative transcriptional control element located upstream of the murine c-*myc* gene. EMBO J. 5: 899-904.

Resar, L.M.S., Dolde, C., Barrett, J.F., Dang, C.V. 1993. B-Myc inhibits neoplastic transformation and transcriptional activation by c-Myc. Mol. Cell. Biol. 13(2): 1130-1136.

Richard, S. and Zingg, H.H. 1990. The human oxytocin gene promoter is regulated by estrogens, J. Biol. Chem. 265: 6098-6103.

Richardson, F.L. 1955. The relative growth of the mammary glands in normal, ovariectomized and adrenalectomized female mice. Anat. Rec. 123: 279-289.

Richman, A. and Hayday, A. 1989. Serum inducible expression of transfected human c-myc genes. Mol. Cell. Biol. 9: 4962-4969.

Rigaud, G., Roux, J., Pictet, R., Grange, T. 1991. *In vivo* footprinting of rat TAT Gene - dynamic interplay between the glucocorticoid receptor and a liver-specific factor. Cell 67: 977-986.

Riggs, K.J., Merrell, K.T., Wilson, G., Calame, K. 1991. Common factor 1 is a transcriptional activator which binds in the c-*myc* promoter, the skeletal α-actin promoter, and the immunoglobulin heavy-chain enhancer. Mol. Cell. Biol. 11: 1765-1769.

Ridsdale, J.A., Hendzel, M.J., Delcuve, G.P., Davie, J.R. 1990. Histone acetylation alters the capacity of the H1 histones to condense transcriptionally active / competent chromatin. J. Biol. Chem. 265: 5150-5156.

Roberts, L. 1993. Zeroing in on a breast cancer susceptibility gene. Science 259: 622-625.

Romieu, I., Berlin, J.A., Colditz, G. 1990. Oral contraceptives and breast cancer: review and meta-analysis. Cancer 66: 2253-2263.

Rosenfeld, M.G. 1991. POU-domain transcription factors: pou-er-ful developmental regulators. Genes Dev. 5: 897-907.

Roy, A.L., Meisterernst, M., Pognonec, P., Roeder, R.G. 1991. Cooperative interaction of an initiator-binding transcription initiation-factor and the helix loop helix activator USF. Nature 354: 245-248.

Ruley, H.E. 1983. Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. Nature 304: 602-606.

Rustgi, A.K., Dyson, N., Bernards, R. 1991. Amino-terminal domains of c-*myc* and N-*myc* proteins mediate binding to the retinoblastoma gene product. Nature 352: 541-544.

Saito, H., Hayday, A.C., Wiman, K., Hayward, W.S., Tonegawa, S. 1983. Activation of the c-*myc* gene by translocation: a mode for translational control. Proc. Natl. Acad. Sci. 80: 7476-7480.

Sakai, D.D., Helms, S., Carlstedt-Duke, J., Gustafsson, J.-Å., Rottman, F.M., Yamamoto, K.R. 1988. Hormone-mediated repression: a negative glucocorticoid response element from the bovine prolactin gene. Genes Dev. 2: 1144-1154.

Saksela, K., Koskinen, P., Alitalo, K. 1991. Binding of a nuclear factor to the upstream region of the c-*myc* gene. Oncogene Res. 6: 73-76.

Saltzman, A.G. and Weinmann, R. 1989. Promoter specificity and modulation of RNA polymerase II transcription. FASEB J. 3: 1723-1733.

Saluz, H.P., Jiricny, J., Jost, J.P. 1988. Genomic sequencing and *in vivo* genomic footprinting of an expression-specific DNase I hypersensitive site of avian vitellogenin II promoter reveal a demethylation of a mCpG and a change in specific interactions of proteins with DNA. Proc. Natl. Acad. Sci. 85: 6697-6700.

Saluz, H.P. and Jost, J.P., 1989. A simple high-resolution procedure to study DNA methylation and *in vivo* DNA-protein interactions on a single-copy gene level in higher eukaryotes. Proc. Natl. Acad. Sci. 86: 2602-2606.

Sambrook, J., Fritsch, E.F., Maniatis, T. (eds). 1989. Molecular Cloning - a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press, New York.

Santos, G.F., Scott, G.K., Lee, W.M.F., Liu, E., Benz, C. 1988. Estrogen-induced post-transcriptional modulation of c-*myc* proto-oncogene expression in human breast cancer cells. J. Biol. Chem. 263: 9565-9568.

Schauer, M., Chalepakis, G., Wilman, T., Beato, M. 1989. Binding of hormone accelerates the kinetics of glucocorticoid and progesterone-receptor binding to DNA. Proc. Natl. Acad. Sci. 86: 1123-1127.

Schaufele, F., West, B.L., Reudelhuber, T. 1990. Somatotroph- and lactotroph-specific interactions with the homeobox protein binding sites in the rat growth hormone gene promoter. Nucleic Acids Res. 18(17): 5235-5243.

Schild, C., Claret, F.-X., Wahli, W., Wolffe, A.P. 1993. A nucleosome-dependent static loop potentiates estrogen-regulated transcription from *Xenopus* vitellogenin B1 promoter *in vitro*. EMBO J. 12: 423-433.

Schmid, P., Schulz, W.A., Hameister, H. 1989. Dynamic expression pattern of the *myc* protooncogene in midgestation mouse embryos. Science 243: 226-229.

Schöenenberger, C.-A., Andres, A.-C., Groner, B., van der Valk, M., LeMeur, M., Gerlinger, P. 1988. Targeted c-*myc* gene expression in mammary glands of transgenic mice induces mammary tumours with constitutive milk protein gene transcription. EMBO J. 7: 169-175.

Schreiber, A.B., Winkler, M.E., Derynck, R. 1986. Transforming growth factor-α: a more potent angiogenesis mediator than epidermal growth factor. Science 232: 1250-1254.

Schüle, R., Muller, M., Kaltschmidt, C., Renkawitz, R. 1988a. Many transcription factors interact synergistically with steroid receptors. Science 242: 1418-1420.

Schüle, R., Muller, Otsuka-Murakami, H., Renkawitz, R. 1988b. Cooperativity of the glucocorticoid receptor and the CACCC-box binding factor. Nature 332: 87-90.

Schüle, R., Rangarajan, P., Yang, N., Kleeiewer, S., Ransone, L.J., Bolado, J., Verma, I.M., Evans, R. 1991. Retinoic acid is a negative regulator of AP-1 responsive genes. Proc. Natl. Acad. Sci 88: 6092-6096.

Schwab, M., Alitalo, K., Klempnauer, K.-H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., Trent, J. 1983. Amplified DNA with limited homology to the *myc* cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature 305: 245-248.

Schwab, M., Varmus, H.E., Bishop, J.M. 1985. Human N-*myc* gene contributes to neoplastic transformation of mammalian cells in culture. Nature 316: 160-162.

Schwabe, J.W., Chapman, L., Finch, J.T., Rhodes, D. 1993. The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell 75: 567-578.

Schwabe, J.W.R. and Rhodes, D. 1991. Beyond zinc fingers: steroid hormone receptors have a novel structural motif for DNA recognition. Trends Biochem. Sci. 16: 291-296.

Schweers, L.A. and Sanders, M.M. 1991. A protein with a binding specificity similar to NF-κB binds to a steroid-dependent regulatory element in the ovalbumin gene. J. Biol. Chem. 266: 10490-10497.

Scott, G.K., Kushner, P., Vigne, J.-L., Benz, C.C. 1991. Truncated forms of DNA-binding estrogen receptors in human breast cancer. J. Clin. Invest. 88: 700-706.

Seipel, K., Georgiev, O., Schaffner, W. 1992. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. EMBO J. 11: 4961-4968.

Serfling, E., Jasin, M., Schaffner, W. 1985. Enhancers and eukaryotic gene transcription. Trends Genet. 8: 224-230.

Seyfred, M.A. and Gorski, J. 1990. An interaction between the 5'flanking distal and proximal regulatory domains of the rat prolactin gene is required for transcriptional activation by estrogens. Mol. Endocrinol. 4: 1226-1234.

Sheiness, D. and Bishop, J.M. 1979. DNA and RNA from uninfected vertebrate cells contain nucleotide sequences related to the putative transforming gene of avian myelocytomatosis virus. J. Virol. 31: 514-521.

Shemshedini, L., Knauthe, R., Sassone-Corsi, P., Pornon, A., Gronemeyer, H. 1991. Cell-specific inhibitory and stimulatory effects of Fos and Jun on transcription activation by nuclear receptors. EMBO J. 10: 3839-3849.

Shemshedini, L., Ji, J., Brou, C., Chambon, P., Gronemeyer, H. 1992. *In vitro* activity of the transcription activation functions of the progesterone receptor: evidence for intermediary factors. J. Biol. Chem. 267: 1834-1839.

Shichiri, M., Hanson, K.D., Sedivy, J.M. 1993. Effects of c-*myc* expression on proliferation, quiescence, and the G_0 to G_1 transition in nontransformed cells. Cell Growth Diff. 4: 93-104.

Shih, C.-K., Linial, M., Goodenow, M.M., Hayward, W.S. 1984. Nucleotide sequence 5' of the chicken c-*myc* coding region: localization of a noncoding exon that is absent from *myc* transcripts in most avian leukosis virus induced lymphomas. Proc. Natl. Acad. Sci. 81: 4697-4701.

Showe, L.C. and Croce, C.M. 1987. The role of chromosomal translocations in B cell and T cell neoplasia. Annu. Rev. Immunol. 5: 253-277.

Shupnik, M.A., Weinmann, C.M., Notides, A.C., Chin, W.W. 1989. An upstream region of the rat leutenizing hormone ß gene binds estrogen receptor and confers estrogen responsiveness. J. Biol. Chem. 264: 80-86.

Sibley, D.R., Benovic, J.L., Caron, M.G., Lefkowitz, R.J. 1988. Phosphorylation of cell surface receptors: a mechanism for regulating signal transduction pathways. Endocrine Rev. 9: 38-56.

Siebenlist, U., Bressler, P., Kelly, K. 1988. Two distinct mechanisms of transcriptional control operate on c-*myc* during differentiation of HL60 cells. Mol. Cell. Biol. 8: 867-874.

Siebenlist, U., Henninghausen, L., Battey, J., Leder, P. 1984. Chromatin structure and protein binding in the putative regulatory region of the c-*myc* gene in Burkitt lymphoma. Cell 37: 381-391.

Slamon, D.J., Boone, T.C., Seeger, R.C., Keith, D.E., Cazin, V., Lee, H.C., Souza, L.M. 1986. Identification and characterization of the protein encoded by the human N-*myc* oncogene. Science 232: 768-777.

Slater, E.P., Redeuihl, G., Theis, K., Suske, G., Beato, M. 1990. The uteroglobin promoter contains a noncanonical estrogen responsive element. Mol. Endocrinol. 4: 604-610.

Small, M.B., Hay, N., Schwab, M., Bishop, J.M. 1987. Neoplastic transformation by the human N-*myc* gene. Mol. Cell. Biol. 7: 1638-1645.

Smith, G.E. and Summers, M.D., 1980. The bidirectional transfer of DNA to nitrocellulose or diazobenzyloxymethyl paper. Anal. Biochem. 109: 123-129.

Snyder, R.C., Ray, R., Blume, S., Miller, D.M. 1991. Mithramycin blocks transcriptional initiation of the c-*myc* P1 and P2 promoters. Biochemistry 30: 4290-4297.

Sorrentino, V., Drozdoff, V., McKinney, M.D., Zeitz, L., Fleissner, E. 1986. Potentiation of growth factor activity by exogenous c-*myc* expression. Proc. Natl. Acad. Sci. 83: 8167-8171.

Soule, H.D. and McGrath, C.M. 1980. Estrogen responsive proliferation of clonal human breast carcinoma cells in athymic nude mice. Cancer Lett. 10: 177-189.

Soule, H.D., Vazquez, A., Long, A., Albert, S., Brennan, M.A. 1973. Human cell line from a pleural effusion derived from a breast carcinoma. J. Natl. Cancer Inst. 51: 1409-1413.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503.

Spencer, C.A. and Groudine, M. 1991. Control of c-*myc* regulation in normal and neoplastic cells. Adv. Cancer Res. 56: 1-48.

Spencer, C.A., LeStrange, R.C., Hayward, W.S., Novak, U., Groudine, M. 1990. The block to transcription elongation is promoter dependent in normal and Burkitt's lymphoma c-*myc* alleles. Genes Dev. 4: 75-88.

Spicer, D., Shoupe, D., Pike, M.C. 1991. GnRH agonists as contraceptive agents - predicted significantly reduced risk of breast cancer. Contraception. 44: 289-310.

Spicer, D.B. and Sonenshein, G.E. 1992. An antisense promoter of the murine c-*myc* gene is localized within intron 2. Mol. Cell. Biol. 12(3): 1324-1329.

Spotts, G.D. and Hann, S.R. 1990. Enhanced translation and increased turnover of c-*myc* proteins occur during differentiation of murine erythroleukemia cells. Mol. Cell. Biol. 10: 3952-3964.

Stanton, B.R., Perkins, A.S., Tessarollo, L., Sassoon, D.A., Parada, L.F. 1992. Loss of N-*myc* function results in embryonic lethality and failure of the epithelial component of the embryo to develop. Genes and Dev. 6: 2235-2247.

Stanton, L.W., Watt, R., Marcu, K.B. 1983. Translocation, breakage and truncated transcripts of c-*myc* oncogene in murine plasmacytomas. Nature 303: 401-406.

Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A., Baltimore, D. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature 323: 640-643.

Stern, D., Roberts, A., Roche, N.S., Sporn, M.D., Weinberg, R. 1986. Differential responsiveness of *myc* and *ras*-transfected cells to growth factors: selective stimulation of *myc*-transfected cells by epidermal growth factor. Mol. Cell. Biol. 6: 870-877.

Stewart, T.A., Pattengale, P.K., Leder, P. 1984. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/*myc* fusion genes. Cell 38: 627-637.

Stone, J., de Lange, T., Ramsay, G., Jakobvits, E., Bishop, J.M., Varmus, H.E., Lee, W.M.F. 1987. Definition of regions in human c-*myc* that are involved in transformation and nuclear localization. Mol. Cell. Biol. 7: 1697-1709.

Strähle, U., Schmid, W., Schütz, G. 1988. Synergistic action of the glucocorticoid receptor with transcription factors. EMBO J. 7: 3389-3395.

Street, A.J., Blackwood, E., Lüscher, B., Eisenman, R.N. 1990. Mutational analysis of the carboxy-terminal casein kinase II phosphorylation site in human c-*myc*. Curr. Topics Microbiol. Immunol. 166: 251-258.

Strühl, K. 1989. Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. Trends Biol. Sci. 14: 137-140.

Sturm, R.A. and Herr, W. 1988. The POU domain is a bipartite DNA binding structure. Nature 366: 601-604.

Su, W., Jackson, S., Tjian, R., Echols, H. 1991. DNA looping between sites for transcriptional activation - self-association of DNA-bound Sp1. Genes Dev. 5: 820-826.

Suetake, I., Tajima, S., Asano, A. 1993. Identification of two novel mouse nuclear proteins that bind selectively to a methylated c-Myc recognizing sequence. Nuc. Acids Res. 21(9): 2125-2130.

Sugiyama, A., Kume, A., Nemoto, K., Lee, S.Y., Agami, Y., Nemoto, F., Nishimura, S., Kuchino, Y. 1989. Isolation and characterization of s-*myc*, a member of the rat *myc* gene family. Proc. Natl. Acad. Sci. 86: 9144-9148.

Sun, J-M., Penner, C.G., Davie, J.R. 1992. Analysis of erythroid nuclear proteins binding to the promoter and enhancer elements of the chicken histone H5 gene. Nucleic Acids Res. 20: 6385-6392.

Sussman, D.J., Chung, J. Leder, P. 1991. *In vitro* and *in vivo* analysis of the c-*myc* RNA polymerase III promoter. Nucleic Acids Res. 19: 5045-5052.

Suzuki, M. 1989. Spxx, a frequent sequence motif in gene regulatory proteins. J. Mol. Biol. 207: 61-84.

Svaren, J. and Chalkley, R. 1990. The structure and assembly of active chromatin. Trends. Genet. 6: 52-56.

Swartwout, S.G. and Kinniburgh, A.J. 1989. c-*myc* RNA degradation in growing and differentiating cells: possible alternate pathways. Mol. Cell. Biol. 9: 288-295.

Szajnert, M.F., Bornkamm, G.W., Kaplan, J.C., Lenoir, G.M., Saule, S., Wajcman, H. 1987. Clustered somatic mutations in and around 1st exon of non-rearranged c-*myc* in Burkitt-lymphoma with T(8-22) translocation. Nucleic Acids Res. 15: 4553-4565.

Takimoto, M., Quinn, J.P., Farina, A.R., Staudt, L.M., Levens, D. 1989. *fos/jun* and octamer-binding protein interact with a common site in a negative element of the human c-*myc* gene. J. Biol. Chem. 264: 8992-8999.

Talanian, R.V., McKnight, C.J., Kim, P.S. 1990. Sequence-specific DNA binding by a short peptide dimer. Science 249: 769-771.

Talwalker, P.K. and Meites, J. 1961. Mammary lobuloalveolar growth induced by anterior pituitary hormones in adreno-ovariectomized and adreno-ovariectomized-hypophysectomized rats. Proc. Soc. Exp. Biol Med. 107: 880-883.

T'Ang, A., Varley, J.M., Chakraborty, S., Murphree, A.L., Fung, Y.-K. T. 1988. Structural rearrangement of the retinoblastoma gene in human breast carcinoma. Science 242: 263-266.

Tasset, D., Tora, L., Fromental, C., Scheer, E., Chambon, P. 1990. Distinct classes of transcriptional activating domains function by different mechanisms. Cell 62: 1177-1187.

Teng, C.T., Liu, Y., Yang, N., Walmer, D., Panella, T. 1992. Differential molecular mechanism of the estrogen action that regulates lactoferrin gene in human and mouse. Mol. Endo. 6: 1969-1981.

Thalmeier, K., Synovzik, H., Mertz, R., Winnacker, E.-L., Lipp, M. 1989. Nuclear factor E2F mediates basic transcription and *trans*-activation by E1a of the human Myc promoter. Genes Dev. 3: 527-536.

Theill, L.E., Castrillo, J.-L., Wu, D., Karin, M. 1989. Dissection of functional domains of the pituitary-specific transcription factor GHF-1. Nature 342: 945-948.

Thisse, B., Stoetzel, C., Gorostize-Thisse, C., Perin-Schmidt, F. 1988. Sequence of the *twist* gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. EMBO J. 7: 2175-2183.

Thomas, D.B. 1986. Hormones and hormone receptors in the etiology of breast cancer. Breast Cancer Res. Treat. 7: 11-22.

Thomas, P.S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. 77: 5201-5205.

Thompson, C.B., Challoner, P.B., Neiman, P.E., Groudine, M. 1985. Levels of c-*myc* oncogene mRNA are invariant throughout the cell cycle. Nature 314: 363-366.

Tonelli, Q.J. and Soroff, S. 1980. Epidermal growth factor requirement for development of cultured mammary gland. Nature 285: 250-254.

Topper, Y.J. 1970. Multiple hormone interactions in the development of the mammary gland *in vitro*. Recent Prog. Horm. Res. 26: 287-308.

Topper, Y.J. and Freeman, C.S. 1980. Multiple hormone interactions in the developmental biology of the mammary gland. Physiol. Rev. 60: 1049-1106.

Tora, L., White, J., Brou, C., Tasset, D., Webster, N., Scheer, E., Chambon, P. 1989a. The human estrogen receptor has two independent nonacidic transcriptional activation functions. Cell 59: 477-487.

Tora, L., Mullick, A., Metzger, D., Ponglikitmongkol, M., Park, I., Chambon, P. 1989b. The cloned oestrogen receptor contains a mutation which alters its hormone binding properties. EMBO J. 8: 1981-1986.

Touitou, I., Mathieu, M., Rochefort, H. 1990. Stable transfection of the estrogen receptor cDNA into HeLa cells induces estrogen responsiveness of endogenous cathepsin D gene but not of cell growth. Biochem. Biophys. Res. Commun. 169: 109-115.

Tourkine, N., Mechti, N., Piechaczyk, M., Jeanteur, P., Blanchard, J.-M. 1989. *In situ* and *in vitro* evidence for intragenic binding of nuclear factors at the murine c-*myc* locus. Oncogene 4: 973-978.

Travers, A.A. 1987. DNA bending and nucleosome positioning. Trends Biochem. Sci. 12: 108-112.

Treisman, J., Harris, E., Wilson, D., Desplan, C. 1992. The homeodomain: a new face for the helix-turn-helix? BioEssays 14: 145-150.

Tsai, S.Y., Carlstedt-Duke, J., Weigel, N.L., Dahlman, K., Gustafsson, J.-Å., Tsai, M.-J., O'Malley, B.W. 1988. Molecular interactions of steroid

hormone receptor with its enhancer element: evidence for receptor dimer formation. Cell 55: 361-369.

Tsai, S.Y., Tsai, M.-J., O'Malley, B.W. 1989. Cooperative binding of steroid hormone receptors contributes to transcriptional synergism at target enhancer elements. Cell 57: 443-448.

Tsukamoto, A.S., Grosschedl, R., Guzman, R.C., Parslow, T., Varmus, H.E. 1988. Expression of the *int-1* gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. Cell 55: 619-625.

Tulinius, H., Egilsson, V., Olafsdottir, G.H., Sigvaldason, H. 1992. Risk of prostate, ovarian and endometrial cancer among relatives of women with breast cancer. British Medical J. 305: 855-857.

Udvadia, A.J., Rogers, K.T., Higgins, P.D.R., Murata, Y., Martin, K.H., Humphrey, P.A., Horowitz, J.M. 1993. Sp-1 binds promoter elements regulated by the RB protein and Sp-1-mediated transcription is stimulated by RB coexpression. Proc. Natl. Acad. Sci. 90: 3265-3269.

Umesono, K. and Evans, R.M. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57: 1139-1146.

Umesono, K., Murakami, K.K., Thompson, C.C., Evans, R.M. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell 65: 1255-1266.

Vallee, B.L., Coleman, J.E., Auld, D.S. 1991. Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. Proc. Natl. Acad. Sci. 88: 999-1003.

van Beneden, R.J., Watson, D.K., Chen, T.T., Lantenberger, J.A., Papas, T.S. 1986. Cellular *myc* (c-*myc*) in fish (rainbow trout): Its relationship to

other vertebrate *myc* genes and to the transforming genes of the MC29 family of viruses. Proc. Natl. Acad. Sci. 83: 3698-3702.

Vanderbilt, J.N., Bloom, K.S., Anderson, J.N. 1982. Endogenous nuclease - properties and effects on transcribed genes in chromatin. J. Biol. Chem. 257: 3009-3017.

van der Burg, B., van Selm-Miltenburg, A.J.P., de Laat, S.W., van Zoelen, E.J.J. 1989. Direct effects of estrogen on c-*fos* and c-*myc* protooncogene expression and cellular proliferation in human breast cancer cells. Mol. Cell. Endo. 64: 223-228.

Van Driel, R., Humbel, B., de Jong, L. 1991. The nucleus: A black box being opened. J. Cell. Biochem. 47: 311-316.

van Wijnen, A.J., Bidwell, J., Lian, J.B., Stein, J.L., Stein, G.S. 1992. Stairway assays: rapid localization of multiple protein/DNA interaction sites in gene-regulatory 5' regions. Biotechniques 12(3): 400-407.

van Wijnen, A.J., Bidwell, J.P., Fey, E.G., Penman, S., Lian, J.B., Stein, J.L., Stein, G.S. 1993. Nuclear matrix association of multiple sequence-specific DNA-binding activities related to Sp1, ATF, CCAAT, C/EBP, Oct-1, and Ap-1. Biochem. 32: 283-293.

Vassilev, L. and Johnson, E.M. 1990. An initiation zone of chromosomal DNA located upstream of the c-*myc* gene in proliferating HeLa cells. Mol. Cell. Biol. 10: 4899-4904.

Versteeg, R., Noordermeer, I., Kruse-Wolters, M., Ruiter, D., Schrier, P. 1988. c-*myc* down-regulates class I HLA expression in human melanomas. EMBO J. 7: 1023-1029.

Vickers, P.J., Dickson, R.B., Shoemaker, R., Cowan, K.H. 1988. A multidrug-resistant MCF-7 human breast cancer cell line which

exhibits cross-resistance to antiestrogens and hormone-independent tumor growth *in vivo*. Mol. Endocrinol. 2: 886-892.

Vignon, F., Bardon, S., Chalbos, D., Rochefort, H. 1983. Antiestrogenic effect of R5020: a synthetic progestin in human breast cancer cells in culture. J. Clin. Endocrinol. Met. 56: 1124-1130.

Villares, R. and Cabrera, C.V. 1987. The *achaete-scute* gene complex of *D. melanogaster*: Conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. Cell 50: 415-424.

Vinson, C.R., Sigler, P.B., McKnight, S.L. 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science 246: 911-916.

Volkers, N. 1992. One in nine: the risk of a lifetime. J. Natl. Cancer Inst. 84(8): 564.

Vonderhaar, B.K. and Greco, A. 1979. Lobuloalveolar development of mouse mammary glands is regulated by thyroid hormones. Endocrinol. 104: 409-418.

Von Hippel, P.H. and Berg, O.G. 1989. Facilitated target location in biological systems. J. Biol. Chem. 264: 675-678.

von Kries, J.P., Buhrmester, H., Strätling, W.H. 1991. A matrix / scaffold attachment region binding protein: identification, purification, and mode of binding. Cell 64: 123-135.

Vorherr, H. 1980. Breast cancer: epidemiology, endocrinology, biochemistry and pathobiology. Urban & Schwarznberg, Baltimore-Munich.

Vriz, S., Taylor, M., Mechali, M. 1989. Differential expression of two *Xenopus* c-*myc* proto-oncogenes during development. EMBO J. 8: 4091-4097.

Wagner, A.J., Le Beau, M.M., Diaz, M.O., Hay, N. 1992. Expression, regulation, and chromosomal localization of the Max gene. Proc. Natl. Acad. Sci. 89: 3111-3115.

Wagner, A.J., Meyers, C., Laimins, L.A., Hay, N. 1993a. c-Myc induces the expression and activity of ornithine decarboxylase. Cell Growth Diff. 4: 879-883.

Wagner, A.J., Small, M.B., Hay, N. 1993b. Myc-mediated apoptosis is blocked by ectopic expression of Bcl-2. Mol. Cell. Biol. 13: 2432-24440.

Waitz, W. and Loide, P. 1991. Cell cycle dependent association of c-*myc* with the nuclear matrix. Oncogene 6: 29-35.

Walker, P., Germond, J.-E., Brown-Luedi, M., Givel, F., Wahli, W. 1984. Sequence homologies in the region preceding the transcription initiation site of the liver estrogen responsive vitellogenin and apo-VLDLII genes. Nucleic Acids Res. 12: 8611-8625.

Walter, P., Green, S., Greene, G., Krust, A., Bornert, J.M., Jeltsch, J.M., Staub, A., Jensen, E., Scrace, G., Waterfield, M., Chambon, P. 1985. Cloning of the human estrogen receptor cDNA. Proc. Natl. Acad. Sci. 82: 7889-7893.

Wang,W.D., Gralla, J.D., Carey, M. 1992. The acidic activator Gal4-Ah can stimulate polymerase II transcription by promoting assembly of a closed complex requiring TFIID and TFIIA. Genes Dev. 6: 1716-1727.

Waterman, M.L., Adler, S., Nelson, C., Greene, G.L., Evans, R.M., Rosenfeld, M.G. 1988. A single domain of the estrogen receptor confers deoxyribonucleic acid binding and transcriptional activation of the rat prolactin gene. Mol. Endocrinol. 2: 14-21.

Waters, C.M., Littlewood, T.D., Hancock, D.C., Moore, J.P., Evan, G.I. 1991. c-Myc protein expression in untransformed fibroblasts. Oncogene 6: 797-805.

Watson, D.K., Reddy, E.P., Duesberg, P.H., Papas, T.S. 1983. Nucleotide sequence analysis of the chicken c-*myc* gene reveals homologous and unique coding regions by comparison with the transforming gene of avian myelocytomatosis virus MC29, Delta-GAG-myc. Proc. Natl. Acad. Sci. 80: 2146-2150.

Watson, P.H., Pon, R.T., Shiu, R.P.C. 1991. Inhibition of c-*myc* expression by antisense oligodeoxynucleotides identifies a crucial role for c-*myc* in the growth of human breast cancer. Cancer Res. 51: 3996-4000.

Watt , R., Nishikura, K., Sorrentino, J., ar-Rushdi, A., Croce, C., Rovera, G. 1983. The structure and nucleotide sequence of the 5' end of the human c-*myc* oncogene. Proc. Natl. Acad. Sci. 80: 6307-6311.

Watt, R.A., Shatzman, A.R., Rosenberg, M. 1985. Expression and characterization of the human c-*myc* DNA-binding protein. Mol. Cell. Biol. 5(3): 448-456.

Webster, N.J.G., Green, S., Jin, J.R., Chambon, P. 1988. The hormone binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. Cell 54: 199-207.

Webster, N.J.G., Green, S., Tasset, D., Ponglikitmongkol, M., Chambon, P. 1989. The transcriptional activation function located in the hormone binding domain of the human oestrogen receptor is not encoded by a single exon. EMBO J. 8: 1441-1446.

Weintraub, H. 1985. Assembly and propagation of repressed and derepressed chromosomal states. Cell 42: 705-711.

Weisinger, G., Remmers, E.F., Hearing, P., Marcu, K.B. 1988. Multiple negative elements upstream of the murine c-*myc* gene share nuclear factor binding sites with SV40 and polyoma enhancers. Oncogene 3: 635-646.

Weisz, A. and Rosales, R. 1990. Identification of an estrogen response element upstream of the human c-*fos* gene that binds the estrogen receptor and the AP-1 transcription factor. Nuclei Acids Res. 18: 5097-5105.

Wenzel, A., Cziepluch, C., Hamann, U., Schürmann, J., Schwab, M. 1991. The N-Myc oncoprotein is associated *in vivo* with the phosphoprotein Max (p20/22) in human neuroblostoma cells. EMBO J. 10 (12): 3703-3712.

Whitelaw, B., Wilkie, N.M., Jones, K.A., Kadonaga, J.T., Tjian, R., Lang, J.C. 1988. Transcriptionally active domains in the 5' flanking sequence of human c-*myc*. UCLA Symposia on Molecular and Cellular Biology. 58: 337-351.

Whitlock, J.P., Jr., Rushizky, G.W., Simpson, R.T. 1977. DNase-sensitive sites in nucleosomes; their relative susceptibilities depend on nuclease used. J. Biol. Chem. 252: 3003-3006.

Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A., Harlow, E. 1988. Association between an oncogene and an antioncogene - the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature 334: 124-129.

Wilding, G., Lippman, M.E., Gelmann, E.P. 1988. Effects of steroid hormones and peptide growth factors on protooncogene c-*fos* expression in human breast cancer cells. Cancer Res. 48: 802-805.

Willett, W. 1989. The search for the causes of breast and colon cancer. Nature 338: 389-394.

Williams, G.T. 1991. Programmed cell death: apoptosis and oncogenesis. Cell 65: 1097-1098.

Williams, M.R., Todd, J.H., Ellis, I.O., Dowle, C.S., Haybittle, J.L., Elston, C.W., Nicholson, R.I., Griffiths, K., Blamey, R.W. 1987. Estrogen receptors in primary and advanced breast cancer - an 8 year review of 704 cases. Br. J. Cancer 55: 67-73.

Wingrove, T.G., Watt, R., Keng, P., Macara, I.G. 1988. Stabilization of *myc* proto-oncogene proteins during Friend murine erythroleukemia cell differentiation. J. Biol. Chem. 263: 8918-8924.

Winqvist, R., Saksela, K., Alitalo, K. 1984. The *myc* proteins are not associated with chromatin in mitotic cells. EMBO J. 3: 2947-2950.

Wittig, B., Wölfl, S., Dorbic, T., Vahrson, W., Rich, A. 1992. Transcription of human c-*myc* in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. EMBO J. 11(12): 4653-4663.

Wong, M.S.J. and Murphy, L.C. 1991. Differential regulation of c-*myc* by progestins and antiestrogens in T-47D human breast cancer cells. J. Steroid. Biochem. Molec. Biol. 39: 39-44.

Worcel, A., Han, S., Wong, M.L. 1978. Assembly of newly replicated chromatin. Cell 15: 969-977.

Workman, J.L. and Roeder, R.G. 1987. Binding of transcription factor TFIID to the major late promoter during *in vitro* nucleosome assembly potentiates subsequent initiation by RNA polymerase II. Cell 51: 613-622.

Workman, J.L., Abmayr, S.M., Cromlish, W.A., Roeder, R.G. 1988. Transcriptional regulation by the immediate early protein of pseudorabies virus during *in vitro* nucleosome assembly. Cell 55: 211-219.

Workman, J.L., Roeder, R.G., Kingston, R.E. 1990. An upstream transcription factor, USF (MLTF), facilitates the formation of preinitiation complexes during *in vitro* chromatin assembly. EMBO J. 1299-1308.

Workman, J.L., Taylor, I.C.A., Kingston, R.E. 1991. Activation domains of stably bound Gal4 derivatives alleviate repression of promoters by nucleosomes. Cell 64: 533-544.

Workman, J.L. and Buchman, A.R. 1993. Multiple functions of nucleosomes and regulatory factors in transcription. Trends Biochem. Sci. 18: 90-95.

Wright, S., Mirels, L.F., Calayag, M.C.B., Bishop, J.M. 1991. Premature termination of transcription from the P_1 promoter of the mouse c-*myc* gene. Proc. Natl. Acad. Sci. 88: 11383-11387.

Wu, X., Pugliese, T., Dickerman, H.W. 1991. Abstract 951. In: The Endocrine Society's Program and Abstracts. 73rd Annual Meeting, Washington, D.C.

Wyllie, A.H., Kerr, J.F.R., Currie, A.R. 1980. Cell death: the significance of apoptosis. Int. Rev. Cytol. 68: 251-305.

Yancopoulos, G.D., Nisen, P.D., Tesfaye, A., Kohl, N.E., Goldfarb, M.P., Alt, F.W. 1985. N-*myc* can cooperate with *ras* to transform normal cells in culture. Proc. Natl. Acad. Sci. 82: 5455-5459.

Yang, B.-S., Geddes, T., Pogulis, R., de Crombrugghe, B., Freytag, S. 1991. Transcriptional suppression of cellular gene expression by c-Myc. Mol. Cell. Biol. 11: 2291-2295.

Yang, B.-S., Gilbert, J.D., Freytag, S.O. 1993. Overexpression of Myc suppresses CCAAT Transcription Factor/Nuclear Factor 1-dependent promoters *in vivo*. Mol. Cell. Biol. 13(5): 3093-3102.

Yang, J.-Q., Remmers, E.F., Marcu, K.B. 1986. The first exon of the c-*myc* proto-oncogene contains a novel positive control element. EMBO J. 5(13): 3553-3562.

Yang-Yen, H.-F., Chambard, J.-C., Sun, Y.-L., Smeal, T., Schmidt, T.J., Drouin, J., Karin, M. 1990. Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. Cell 62: 1205-1215.

Yaniv, M. and Cereghini, S. 1986. Structure of transcriptionally active chromatin. CRC Crit. Rev. Biochem. 21: 1-26.

Yokota, J., Tsunetsugu-Yokota, Y., Battifora, H., LeFevre, C., Cline, M.J. 1986. Alterations of *myc*, *myb* and ras^{Ha} proto-oncogenes in cancers are frequent and show clinical correlation. Science 231: 261-264.

Yokoyama, K. and Imamoto, F. 1987. Transcriptional control of the endogenous *MYC* protooncogene by antisense RNA. Proc. Natl. Acad. Sci. 84: 7363-7367.

Zajac-Kaye, M., Gelmann, E.P., Levens, D. 1988. A point mutation in the c-*myc* locus of a Burkitt Lymphoma abolishes binding of a nuclear protein. Science 240: 1776-1780.

Zajac-Kaye and Levens, D. 1990. Phosphorylation-dependent binding of a 138-kDa *myc* intron factor to a regulatory element in the 1st intron of the c-*myc* gene. J. Biol. Chem. 265: 4547-4551.

Zervos, A.S., Gyuris, J., Brent, R. 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. Cell 72: 223-232.

Zhang, X.-K., Wills, K.N., Husmann, M., Hermann, T., Pfahl, M. 1991. Novel pathways for thyroid hormone receptor action through interaction with *jun* and *fos* oncogene activities. Mol. Cell. Biol. 11: 6016-6025.

Zimmerman, K.A., Yancopoulos, G.D., Collum, R.G., Smith, R.K., Kohl, N.E., Denis, K.A., Nau, M.M., Witte, O.N., Toran-Allerand, D., Gee, C.E., Minna, J.D., Alt, F.W. 1986. Differential expression of *myc* family genes during murine development. Nature 319: 780-783.

Zobel, A., Kalkbrenner, F., Vorbrueggen, G., Moelling, K. 1992. Transactivation of the human c-*myc* gene by c-Myb. Biochem. Biophys. Res. Commun. 186: 715-722.

Zuppan, P., Hall, J.M., Lee, M.K., Ponglikitmongkol, M., King, M.C. 1991. Possible linkage of the estrogen receptor gene to breast cancer in a family with late onset disease. Am. J. Hum. Genet. 48: 1065-1068.