WHAT'S NEW IN GENERAL SURGERY

Molecules, Cancer, and the Surgeon

A Review of Molecular Biology and Its Implications for Surgical Oncology

JEFFREY M. ARBEIT, M.D.

From the Department of Surgery, University of California, San Francisco, San Francisco, California

Interactions between molecules control intra- and intercellular physiology. Cancer is emerging as a disease in which individual molecules are either overproduced, mutated, expressed at inappropriate stages of development, or lost due to inheritance or aberrant mitotic division. The major players in this contest of cellular control are growth factors, growth factor receptors (GFRs), signal transducers, and dominant or suppressor/recessive oncogenes. The tumors most frequently removed by surgeons have been reported to have changes in one or another of these types of molecules. The concept of multistage carcinogenesis, whereby malignancy arises after a sequence of changes that are cumulative, and passed from progenitor to daughter cells, is also being defined as a sequence of molecular, genetic, and chromosomal alterations. Molecular antineoplastic therapy is in early stages of development at the laboratory bench. The future may see patients screened for cancer susceptibility, evaluated for adjuvant therapy, and chosen for particular treatment based on molecular analysis. The types of cancer operations and the scope of surgical resection may change as molecular techniques enhance oncologic treatment.

INCREASINGLY CANCER IS perceived by modern oncologists as a disease characterized by changes in specific molecules. These changes include alterations in the structure, regulation or quantity of growth factors and their receptors, signal transducers, and the proteins encoded by dominant or suppressor/recessive oncogenes. This review will first introduce an overview of normal cellular molecular physiology, followed by examples of changes in each of the above types of molecules in malignancy. Then the current information about the involvement of these molecular or genetic changes in cancers frequently removed by the surgical oncologist will be presented. These ideas will be brought together to formulate a molecular definition of multistage carcinogenesis, a concept that is central to understanding the formation of epithelial cancers. Finally the potential of this knowledge of molecular and genetic function to identify patients at risk to develop cancer or to experience recurrence after curative resection, or to facilitate the development of novel, specific, antineoplastic therapies will be discussed. This evolving understanding of the molecular physiology of cancer will ultimately change classical surgical oncology.

Accepted for publication September 29, 1989

Address reprint requests to Jeffrey Arbeit, M.D., Department of Surgery, University of San Francisco, 400 Parnasus Ave., Room M590, San Francisco, CA 94143.

Organization of Molecules in the Cell

Mammalian cells are organized into tissues and organs with coordinated functions. This organization is controlled by certain classes of molecules. Growth factors are secreted by nearly all cells and act either on the cell of origin, autocrine function, nearby cells, paracrine function, or distant cells in classical endocrine fashion.[1] Each type of growth factor binds to a specific receptor anchored in the cell membrane. This binding causes the receptor to send a signal into the interior of the cell, activating molecules both in the cytoplasm and the nucleus.[2]

The nucleus contains deoxyribonucleic acid (DNA), which contains the genetic code for the constituents of the cell. In addition to DNA, the nucleus also contains ribonucleic acid (RNA), which has a number of functions predominantly dealing with protein synthesis, *i.e.*, ribosomal RNA, transfer RNA, and messenger RNA (mRNA). The latter molecule contains the genetic information that directly codes for specific proteins. In addition the nucleus also contains proteins that are intimately associated with DNA. These proteins replicate the DNA molecules leading to cell proliferation, repair the DNA molecules to prevent mutations, or read the genetic code, *i.e.*, transcribe DNA, to produce the RNA molecules that ultimately result in the synthesis of other proteins.[3] There is a give-and-take relationship of the nuclear constituents with the cell periphery. Signals, mRNA, can emanate from

the nucleus to produce new growth factors or novel growth factor receptors (GFRs). Growth factors or their receptors send signals to the nucleus that alter proteins associated with DNA and change either the rate of DNA synthesis or the rate of transcription of certain genes, *i.e.*, gene expression.[2]

The molecular changes seen in cancer cells are either exaggerations of normal physiology, or inappropriate expression of juvenile developmental patterns. Many changes seen in cancer cells have counterparts in the normal physiology of embryogenesis and organ development, lymphocyte blastogenesis, surgical wound healing, metabolic response to trauma and sepsis, and organ regeneration or hypertrophy after injury or resection.[4] The common molecular threads in this diverse tapestry of normal physiologic functions are controlled proliferation and a regulated response to specific growth factors. In cancers proliferation and growth factor signalling are deregulated and amplified, respectively, which facilitates persistent tumor growth. The goal of modern oncology is to fully characterize the molecular mechanisms of this unrestrained malignant growth to develop specific therapies directed against the aberrant molecular physiology of the cancer cell.

Growth Factors

Growth factors are small peptides that bind to specific cell membrane receptors. This binding produces an intracellular response that results in growth and proliferation, but may also be inhibitory. The growth factors of oncologic importance[5] are platelet-derived growth factor (PDGF), epidermal growth factor (EGF), transforming growth factors α and β (TGFα, TGFβ), fibroblast growth factor (FGF), insulinlike growth factor I and II (IGF-I, IGF-II), interleukin-2 (IL-2), and colony-stimulating factors (CSF-1, CSF-2, multi-CSF). Because the last three factors are predominantly important in hematologic malignancy and immunology, they will not be further discussed in this review.

Platelet-derived Growth Factor

The principal source of PDGF is from platelets.[6] The molecule was the first growth factor isolated and was identified by the ability of serum, in contrast to plasma, to stimulate the growth of cultured cells. When PDGF is added to nontumorigenic, long-term fibroblast lines such as NIH 3T3 cells, these cells are 'transformed,' *i.e.*, they behave like neoplastic cells, their growth rate increases, and they lose contact inhibition *in vitro*, and form tumors when injected into nude mice *in vivo*.[2] This cellular proliferation is the result of changes in gene expression induced by PDGF (see below). Platelet-derived growth factor exists either as a homodimer composed of two α or two β chains or as an α/β heterodimer. Platelet-derived growth factor is produced and primarily stimulates the growth of mesenchymal cells[6]; however the growth factor is also secreted in a paracrine fashion by breast cancer cells.[7]

Epidermal Growth Factor

Epidermal growth factor is released from and acts on epithelial cells.[5] It is an autocrine factor for cultured squamous carcinoma cells. Epidermal growth factor may also be important in the growth of colon cancer cells.[8] Clinically EGF and the presence of the EGF receptor have been used to predict recurrences in breast cancer. The presence of the EGF receptor in human breast cancers has been correlated with a higher incidence of lymph node metastasis at initial presentation and decreased duration of disease-free survival.[9] In node-negative patients the presence of the tumor EGF receptor can identify patients who are likely to experience recurrence, and thus possibly benefit from adjuvant chemotherapy.

Transforming Growth Factor-α

Transforming growth factor-α was originally isolated from sarcoma cell-conditioned media and was operationally defined by its ability to transform 3T3 cells.[10,11] Transforming growth factor-α is 50% homologous to EGF and binds to the same receptor.[11] Transforming growth factor-α is produced by a number of different cell lines, but its most interesting physiology has been associated with breast cancer.[7] Hormone-sensitive breast cancer cells have been found to secrete TGF-α (and also PDGF and IGF-1) as they proliferate in response to estrogen.[7] Tamoxifen, which inhibits the growth of these cells, also blocks TGF-α secretion. Hormone-insensitive breast cancer cells can produce TGF-α without estrogen. This TGF-α secretion is also unresponsive to tamoxifen.

Transforming Growth Factor-β

Transforming growth factor-β was originally identified in the same sarcoma-conditioned media along with TGF-α, but has a completely different structure and physiology.[11,12] Transforming growth factor-β binds to its own receptor on the cell membrane distinct from the EGF receptor. Transforming growth factor-β can cooperate with EGF or TGF-α in stimulating transformation and growth 3T3 cells. The effects of TGF-β are complex in that the molecule inhibits the growth of many types of malignant epithelial cells.[12] In contrast, other epthelial cell lines, *i.e.*, colon cancer cells produce TGF-β.[13] Paracrine secretion of TGF-β by epithelial cells may function to inhibit the growth of normal cells in the proliferating basal layer, facilitating a selective overgrowth of neoplastic cells.

Fibroblast Growth Factor

There are several FGF-related growth factors, but the molecule exists in two predominant forms, acidic and basic. The former is found exclusively in the central nervous system, while the latter is ubiquitous and is the principal form of FGF in the body.[8,14] Fibroblast growth factor produced by fibroblast and macrophages also is stimulatory to these cells. Fibroblast growth factor-b is strongly angiogenic and may be the principal angiogenesis factor *in vivo*, despite the recent cloning of a separate angiogenin gene.[8,15,16] The physiology of FGF is interesting because, unlike other secreted cellular proteins, its post-translational precursor lacks a signal peptide that would normally direct the molecule to exocytotic vesicles. Fibroblast growth factor presumably enters the extracellular space by destruction of cells and lytic release. Fibroblast growth factor binds avidly to heparin and heparan sulfate, which are important extracellular matrix proteoglycans.[14] Thus the extracellular matrix may be an important reservoir of FGF, and the lysis of this matrix by invasive malignant cells may result in liberation of FGF with secondary growth stimulation and angiogenesis.

Insulinlike Growth Factor-I and -II

Insulinlike growth factor-I (somatomedin-A), IGF-II (somatomedin-C), and insulin all belong to the same peptide family.[17] They have a minimal amino acid sequence homology, but a similar tertiary structure.[17] The similar tertiary structure enables the molecules to bind to each other's receptors, albeit at reduced affinity. Several types of malignant cells, *i.e.*, breast cancer[7] and sarcomas,[18] have been shown to overproduce IGF-I or IGF-II, which may result in enhanced autocrine growth stimulation.

Bombesin

Bombesin is predominantly found in the gut and in the brain functioning as a neurotransmitter. The molecule has been shown to transform 3T3 cells similar to PDGF.[2] Bombesin is produced by small cell lung carcinoma (SCLC) cells and increased numbers of bombesin receptors on these cells make up a autocrine stimulatory loop.[19] The potential importance of bombesin in the proliferation of SCLC cells was demonstrated in experiments in which antibodies to the bombesin receptor were shown to inhibit the growth of cells in culture, and of tumors in nude mice.[20]

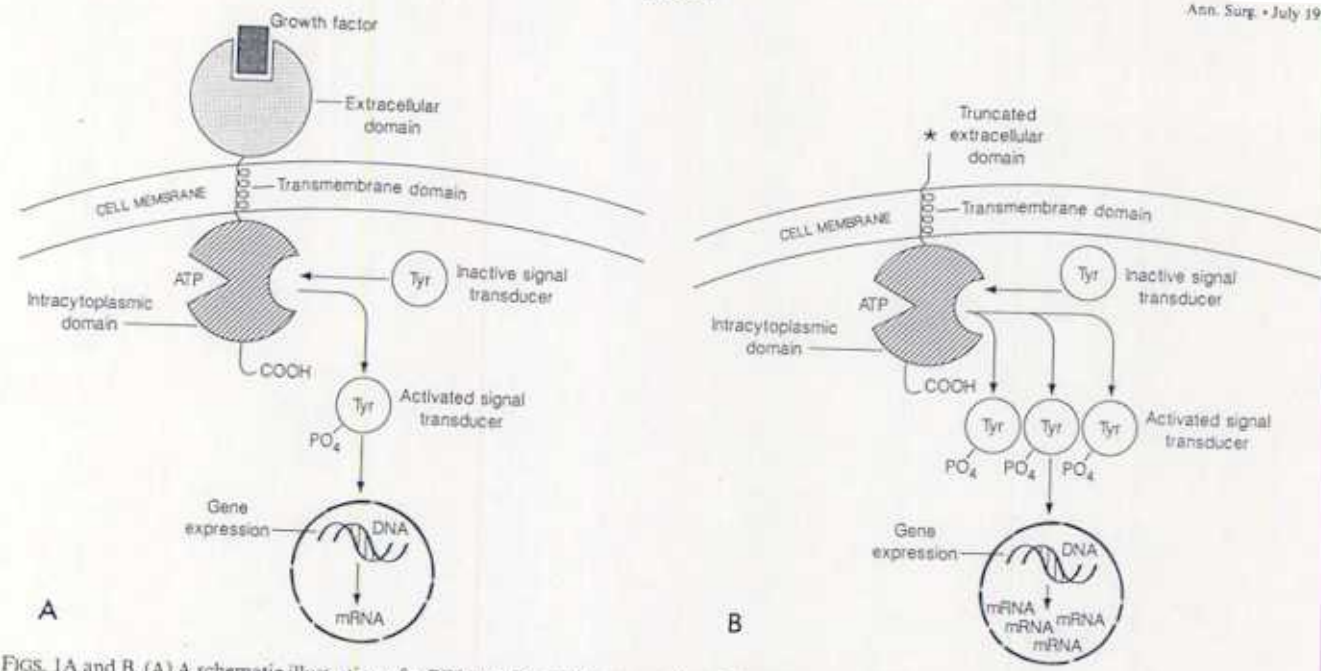
The proliferative effects of growth factors appear to be due to alterations in gene expression consequent to their binding to their receptors. The two proteins that are markedly increased after growth factor administration are the products of the c-*fos* and c-*myc* proto-oncogenes.[2,21,22] Within minutes after growth factor addition to the media of cultured cells, there is a burst of c-*fos* expression that decreases back to baseline with 1 to 2 hours with a slower sustained rise in c-*myc* expression.[23] Both c-*fos* and c-*myc* are potent nuclear oncogenes controlling cellular proliferation.[24] In addition there are a number of other proteins of unknown function that are induced by growth factors.[25] The molecular mechanism of increased proliferation in response to growth factors will increase in complexity as the functions of these proteins and the genes encoding them are elucidated.

Growth Factor Receptors

Growth factor receptors (GFR) have both a similar structural organization and a similar intracellular physiology. They are composed of an extracellular domain, an alpha helical transmembrane segment, and an intracellular domain[26] (Fig. 1). Each domain has discrete functions. The extracellular domain binds ligand (growth factors, or other unknown molecules), the transmembrane helix anchors the protein and may be involved in signal modulation or transduction, and the intracellular domain binds ATP and effector proteins, resulting in protein phosphorylation on tyrosine, *i.e.*, tyrosine kinase activity.[26]

Growth factor receptors can be grouped into three subclasses.[26] Subclass I includes the receptors EGF/TGF-α and the oncoproteins of v-*erb*-B and HER-2/*neu*. Subclass II is composed of the insulin and IGF receptors. Subclass III includes the PDGF and the FGF receptor and the proto-oncogenes c-*fms*/CSF-1 and c-*kit*. The different subclasses are distinguished by the structure of the extracellular and intracellular domains.[26]

When ligand binds to the GFR, the activation of the tyrosine kinase results in a cascade of proteins phosphorylated on tyrosine. Some of the proteins phosphorylated due to growth-factor activation include second messengers and nuclear DNA binding proteins.[27] This protein phosphorylation alters the activity of enzymes, resulting in either enhanced or decreased substrate affinity or catalytic efficiency. A key second messenger that is phosphorylated by activated GFRs is phospholipase-C (PLC),[28] which results in release of diacylglycerol (DAG), inosine triphosphate (ITP_3), and elevation of prostaglandins and leukotrienes. As a result of PLC activation, protein kinase-C (PKC) is stimulated, which produces the burst of c-*fos* and c-*myc* expression.[29] In addition activated PKC phosphorylates GFRs on threonine, which decreases their response to their growth factors, thus effectively shutting them off.[30] Also the GFR auto-phosphorylates itself,[31] which also turns the GFR off because phosphorylation stimulates internalization of the receptor, *i.e.*, 'down regulation.'[26] Growth factor receptor signal transduction appears to require that two or more receptors bind the growth



FIGS. 1A and B. (A) A schematic illustration of a GFR, and the cellular events that follow ligand binding. The GFR has three domains, extracytoplasmic, transmembrane alpha helical, and intracytoplasmic. The extracytoplasmic domain binds the growth factor, the transmembrane helix anchors the protein in the membrane, and the intracytoplasmic domain binds ATP and is a tyrosine kinase for bound substrate proteins. The transmembrane domain was thought to transmit the signal from the bound growth factor to the intracytoplasmic tryosine kinase, but the conformational change in the protein is probably secondary to dimerization of one receptor with another. Tyrosine phosphorylation activates certain proteins; genes that are subsequently transcribed when these proteins reach the nucleus include protooncogenes such as c-*fos* and c-*myc*. (B) Mutated GFRs may lose part of the extracellular domain, *i.e.*, the oncogene v-*erb*-B. This truncation of the extracellular domain leaves the receptor continually activated, and a constant proliferative signal is sent from the cell membrane to the nucleus.

factor as dimers or oligomers.[32] This dimerization of the GFR molecules is thought to produce a conformational change in the protein that activates the intracytoplasmic tyrosine kinase. Receptor dimerization may also facilitate receptor autophosphorylation. Activated receptors also transphosphorylate other neighboring receptors, altering their affinities for their respective growth factors.[33] Receptors encoded by oncogenes frequently have truncated extracellular domains or altered transmembrane helices.[26] These structural changes presumably lock the receptor in a continuous 'on' mode (Fig. 1). This constitutive activation sends a continuous proliferative signal to the nucleus, which may support uncontrolled malignant cell growth.

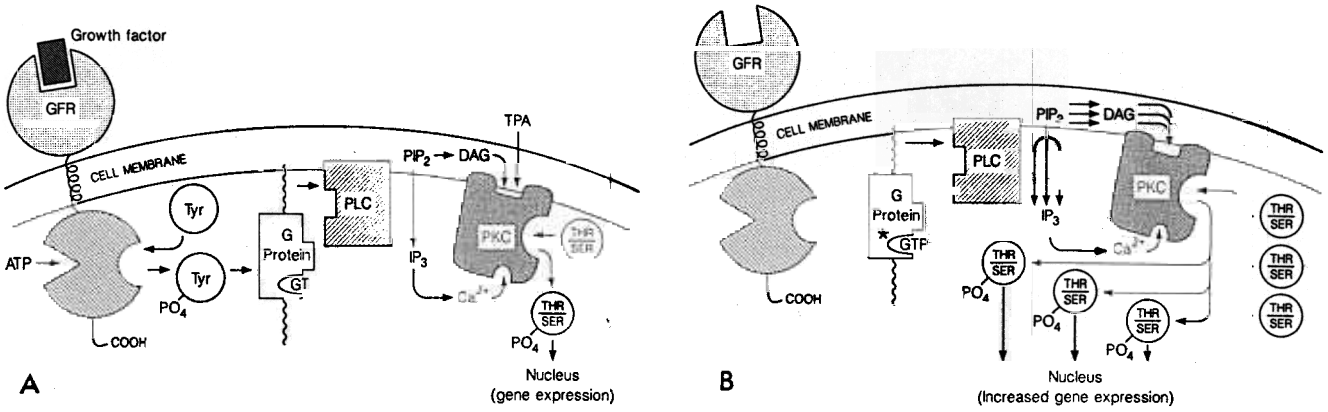
While complex, the physiology of GFRs and neighboring membrane proteins is an example of the multiple interactions that occur when a signal is transmitted from the cell membrane to the nucleus. This complexity offers the hope that part of this molecular pathway in cancers will be amenable to therapeutic manipulation. In fact the emerging understanding of GFR physiology has led to potential novel antineoplastic therapies. Monoclonal antibodies to GFRs have been shown to decrease tumor growth.[20,34] A new class of compounds, 'tyrphostins,' are specific, competitive inhibitors of the intracytoplasmic tyrosine kinase domain of the GFR.[35] These drugs can inhibit the proliferation and growth rate of malignant cells *in vitro*.[35]

Signal Transduction

The signal transduction pathway is composed of molecules that are localized in the membranes that transmit signals from GFRs to secondary effector proteins in the cytoplasm and possibly the nucleus.[27] The signal transduction pathway has three components that are the phosphotidylinositol/ITP_3/DAG/Ca^{++} system,[29] protein kinase-C,[36] and the GTP binding proteins, to which ras proteins belong.[37] Figure 2 illustrates the relationships of each of these components and shows how mutant proteins may facilitate cell proliferation.

Phosphotidylinositol System

Phosphotidylinositol is quantitatively a minor constituent of membrane phospholipids. In response to a number of stimuli, including ligand binding to GFRs, this phospholipid is split by PLC into two second messengers that have profound intracellular effects, *i.e.*, ITP_3 and DAG.[29] Inosine triphosphate releases Ca^{++} from intracellular vesicles, which directly activates both critical metabolic enzymes and structural or mobility proteins such as actin and myosin.[38] Inosine triphosphate also directly stimulates



FIGS. 2A and B. (A) The multiple proteins involved in signal transduction. After a growth factor binds to its receptor a cytoplasmic G protein binds GTP and attaches to the membrane *via* a fatty acid tail. The G protein then stimulates PLC, which splits phosphoinositol-bis-phosphate (PIP_2) into ITP_3 and DAG. Diacylglycerol diffuses in the membrane and activates PKC in conjunction with calcium released from cytoplasmic vesicles by ITP_3. The tumor promoter 12-O-tetradecanoylphorbol-13-acetate directly activates PKC without calcium. Activated PKC phosphorylates proteins on threonine or serine and produces expression of proto-oncogenes such as c-*fos* and c-*myc*. (B) A G protein with a mutation at its GTP binding site, such as an activated ras oncoprotein, does not hydrolyze bound GTP and as a result is continuously activated. Phospholipase-C and subsequently PKC are continuously stimulated, without a growth factor binding to a receptor, producing a constant proliferative signal from the cell membrane to the nucleus.

a number of membrane proteins including GFRs and protein kinase-C. Inosine triphosphate is metabolized by two divergent pathways: one is rapid degradation to inosine phosphate and resynthesis to phosphotidylinositol, the other is progressive phosphorylation to molecules such as inositol tetrakisphosphate (ITP_4), ITP_5, ITP_6, and so on.[29] These molecules may be additional intracellular signals. Diacylglycerol also has two divergent functions. First it activates protein kinase-C, which results in a myriad of intracellular responses.[36] Second it is broken down to arachidonic acid and phosphatidic acid. The arachidonate is further metabolized to prostaglandins and/or leukotrienes or thromboxanes that mediate many intra- and intercellular effects.

Protein Kinase-C

Protein kinase-C is the hub of the membrane signal transduction pathway for many of the GFRs. Protein kinase-C is a cytoplasmic protein that translocates into the membrane after the release of Ca^{++}. It is activated by a combination of DAG, ITP_3, and Ca^{++}. The tumor promoter 12-O-tetradecanoylphorbol-13-acetate (TPA or phorbol myristate acetate, PMA) is structurally similar to DAG and is able to directly activate PKC without ITP_3 or Ca^{++}.[36] Activated PKC phosphorylates proteins on serine and threonine, which either stimulates or inhibits their activity. Protein kinase-C transphosphorylates GFRs, which results in their down regulation. Stimulation of PKC is responsible for the increase in c-*fos* and c-*myc* expression seen after many growth factors bind to their GFRs.[5] The intermediate steps in the transduction of the PKC stimulus to the nucleus are unknown. It appears that GFRs stimulate PKC either by directly activating PLC, which raises membrane DAG levels, or by activating G proteins, which subsequently stimulate PLC and increase membrane DAG (Fig. 2).

G-(ras) Proteins

Proteins that bind GTP are classified as G proteins.[37] The bound GTP stimulates the enzyme. The stimulus decays when the GTP is converted to GDP. The system can be regulated by modulating the length of time GTP is bound before it is degraded to GDP. The paradigm for the function and mechanism of action of the G proteins is the adenylate cyclase G-protein complex that transduces the stimulus from the catecholamine receptor to adenyl cyclase.[37] The principle G proteins that may be involved in signal transduction after GFR stimulation belong to the ras family. The ras family are proto-oncogenes encoded by c-Ha (Harvey)-*ras*, c-Ki (Kirsten)-*ras*, or N-*ras*.[39] Ras proteins are globular with linear amino acid terminal and carboxy terminal tails. The current model is that after ligand binds to GFR, short-chain fatty acids bind to either the amino or carboxy terminal tails of the ras protein, GTP binds, then molecule translocates into the membrane.[40] The molecule then shuttles to PLC, or other as-yet unidentified effectors, activates them, and they in turn activate PKC (Fig. 2). The signal decays as the ras protein hydrolyses its bound GTP. The ras proteins can be constitutively activated by characteristic point mutations (in codons 12, 13, or 61).[39] These mutations alter critical amino acids at the GTP-binding site, resulting in delayed hydrolysis, or no hydrolysis, of the bound nucleotide. These mutated proteins would send continuous prolif-

erative signals to the nucleus because of constant activation of PKC (Fig. 2). This continuous signal for proliferation may facilitate malignant transformation. Recent work has suggested that drugs that inhibit cholesterol metabolism such as compactin or lovastatin inhibit binding of the fatty acids to the tails of the ras protein, preventing their movement into the membrane and their subsequent activation of second messenger enzymes.[41] Because activated ras oncogenes are found in many human cancers,[42] these drugs may have a novel therapeutic potential. The role of activated ras oncogenes in the genesis of solid malignancies is outlined below.

Oncogenes

Oncogenes were originally identified as being the genes responsible for tumor formation by RNA tumor viruses.[43] Later it was discovered that these genes had been acquired (transduced) from the host genome during the process of viral excision.[44] During transduction the structure of these host genes are frequently rearranged or mutated.[44] The normal host genes are called proto-oncogenes while the mutated forms are called oncogenes. When these genes are isolated from viruses, they have a prefix 'v,' *i.e.*, v-Ha-*ras*; when they are recovered from host DNA, they have a 'c' prefix, *i.e.*, c-Ha-*ras*. The italics refer to the gene, whereas normal script is used to describe the protein (oncoprotein) encoded by the gene. Proto-oncogenes are involved in proliferation of cells and are expressed at defined periods in the development of organisms ranging from yeast to mammals.[43] Certain of these genes are reactivated during organ regeneration, wound healing, lymphocyte blastogenesis, *i.e.*, whenever a cellular proliferative response is evoked.[4] Oncogenes are best understood by classifying these molecules into dominant and recessive/suppressor oncogenes, and then subclassifying them according to their intracellular location, and, if known, the function of their encoded oncoproteins[45] (Table 1). The dominant genes can cause malignant transformation by overproduction of either a normal or mutated oncoprotein. The recessive oncogenes suppress tumor formation by antagonizing the effects of dominant oncogenes.

TABLE 1. *Location, Classification, and Functions of Oncogenes*

			Nuclear	
Ligands	Membrane Receptors	Signal Transducers	Transcription Factors	Cell Cycle
sis	*erb*-B (neu)	N-*ras*	*fos*	*myc*
	fms	Ha-*ras*	*jun*	N-*myc*
	kit	Ki-*ras*	*erb*-A	L-*myc*
	ros	*src*		
	trk	*yes*		
	ret	*fgr*		
	sea	*abl*		
	met	*fes*		
		A-*raf*		
		B-*raf*		
		mos		

Dominant Oncogenes

Membrane Oncogenes

This class of oncogenes encode oncoproteins that belong to one of three subclasses of GFRs outlined above. Important proteins in each of these classes include c-erb-B2 (HER-2/neu), c-kit, c-fms, and others. The former two are in the EGFR family, while the latter two belong to the PDGF family.[26] The protoyypical example has been the v-erb-B (avian erythroblastosis) protein. This oncoprotein has an intact transmembrane and cytoplasmic domain but a truncated extracytoplasmic domain.[26] The truncation of the extracytoplasmic domain is presumed to result in constitutive activation of the GFR (Fig. 1).

Growth factorlike oncogenes. The only documented example of this type of oncoprotein is the product of the *sis* (simian sarcoma virus) oncogene. The *sis* product is a variant of PDGF.[5] Normal PDGF is a heterodimer composed of α and β chains, the *sis* product is a β-chain homodimer.[6] Constitutive production of this oncoprotein can cause malignant transformation. Receptor down regulation is ineffective in decreasing the stimulus because the PDGF receptor is activated by the sis oncoprotein while still inside intracytoplasmic secretory vesicles.

Signal transduction oncogenes. The ras proteins are the predominant members of this group. Their mechanism of action has been outlined above. The *src* oncogene and its phosphorylated product $pp60^{src}$ belong to this family. Other oncogenes encoding proteins with tyrosine kinase activity and a cytoplasmic location are listed in Table 1. These oncogenes were found in human tumors, and identified by their ability to transform NIH 3T3 cells. The substrates for these proteins and their role in normal cellular signal transduction remains to be clarified. To date proteins such as PKC or PLC have not had activated oncoprotein counterparts identified, although proto-oncogenes such as c-*raf* and c-*mos* encode oncoproteins that have threonine/serine kinase activity, similar to PKC (Table 1).[46]

Nuclear oncogenes. The prominent members of this group of oncogenes include c-*fos*, c-*erb*-A, c-*jun*, c-*myc*, and c-*myb*.[24] The group represented by c-*fos*, c-*jun*, and c-*erb*-A encode proteins that are transcription factors.[47] These oncoproteins bind to DNA and attract other proteins sequentially, eventually producing a multiprotein machine that moves along the DNA molecule and transcribes the genetic code into messenger, transfer, and ribonuclear RNA.[48] c-*jun* can potentially bind to DNA directly as a homodimer, but it preferentially binds as a heterodimer with c-*fos*.[49] The jun/fos heterodimer has an

ulation at large so that a given individual stands a good chance of being heterozygous (a different allele from the mother and father), and close to the presumed suppressor oncogene of interest (closely linked).[66] An example of a polymorphic allele would be the HLA transplantation locus, or the minor blood group loci. The two different alleles will have a different pattern of 'restriction sites,' places where DNA digesting enzymes can cut, hence two different length fragments will be generated when human DNA is cut and analyzed with a radioactive probe specific for the polymorphic allele.[66] The genetic analysis of patients with malignancies uses peripheral blood monocytes and tumor tissue. The monocytes retain two different bands on a gel after DNA digestion and probing, while the tumor just has one band of double intensity[67] (Fig. 3). This means the tumor is hemizygous for the RFLP, and has lost heterozygosity compared to the monocytes. Because the RFLP is very near the presumed site of the suppressor oncogene, it is very likely that this locus was lost during tumorigenesis.[68] The term suppressor/recessive is derived from the ability of the cell to maintain a normal phenotype despite the loss of one parental gene. Cells are transformed into cancers when both parental copies of these genes are lost due to mutation.

Retinoblastoma. Genetic analysis of families with hereditary and sporadic retinoblastoma (Rb) have provided a paradigm for the recessive malignant transformation in humans. Patients with hereditary Rb inherit an abnormal chromosome from one of their affected parents. This inheritance is termed a germ-line mutation.[67] The condition is recessive because a normal chromosome from the unaffected parent is also present. However during the final weeks of gestation or during early infancy, another mutation occurs in a retinal cell that inactivates this remaining normal Rb gene.[67] This event is termed a sporadic somatic mutation. With the loss of both parental copies of the suppressor oncogene, malignant transformation occurs in the retinal cell (Fig. 3). In the sporadic form of this disease, both events are somatic mutations sequentially inactivating the suppressor gene locus. Restriction-fragment-length polymorphism analysis has demonstrated that the tumor tissue is hemizygous for the Rb locus while peripheral blood lymphocytes maintain heterozygosity.[67] The normal diploid chromosome number is from duplication of the remaining abnormal Rb mutant chromosome. A portion of the long arm of chromosome 13 has been found to be deleted in the tumors of these patients, which presumably is the locus for the Rb gene.[68,69]

These patients continue to display an enhanced tissue tumorigenicity, even if cured of their Rb by surgery and radiation therapy. The incidence of second primary tumors in these patients is approximately 35% after 30 years of follow-up.[70] The second primaries are frequently osteogenic or soft-tissue sarcomas that predominantly arise in the previous radiation port.[70,71] When the DNA from these tumors are probed they are also found to have lost heterozygosity of the Rb locus.[71] In adults the Rb gene has been demonstrated to be lost in breast and SCLCs.[72,73]

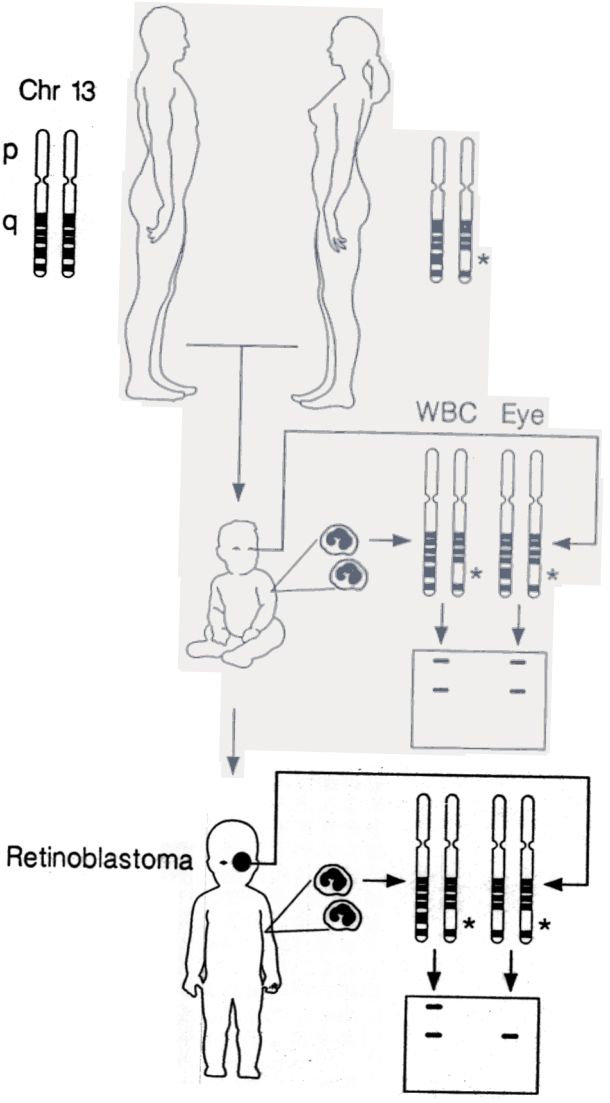


FIG. 3. The retinoblastoma (Rb) suppressor/recessive paradigm. The child inherits a chromosome 13 with a deletion in the 14th band of its long arm (13:q14). The child has a normal chromosome 13 from the father, and when monocyte DNA is digested by a restriction enzyme, two bands are seen on a gel probed with a polymorphic marker gene (restriction-fragment-length polymorphism, RFLP) close to the Rb locus. Because the chromosomal deletion is inherited, it is called a germ-line mutation. During infancy the remaining normal chromosome 13 loses its band at 14q in a retinal cell, and a tumor develops in that eye. The monocyte still has a normal chromosome, thus the monocyte retains heterozygosity while the tumor is hemizygous for the Rb gene. When both tumor and monocyte DNA are analyzed, two bands are still seen for the monocyte, while the tumor has lost one of the bands. This type of mutation is termed a sporadic somatic mutation. This genetic scheme is termed recessive because an individual with one normal and one abnormal chromosome is normal in appearance, and disease only surfaces when the normal chromosome is lost, leaving only the abnormal copy in the target tissue. Similarly the normal protein product appears to suppress the tumorigenic phenotype.

p53 oncogene. This suppressor oncogene was originally thought to be a dominant oncogene because its overproduction could transform cells.[74] In retrospect these studies were found to be performed with a mutant version of the wild type p53 gene. Recently experiments have shown that the normal p53 gene can prevent malignant transformation of cultured cells.[75] Sequestration of the p53 gene product has been found to play an important role in DNA tumor virus carcinogenesis.[76] Clinically mutant p53 genes have been found in osteogenic sarcomas and colon cancers.[77,78]

Wilm's Tumor

These tumors have been shown to have a deletion of the short arm of chromosome 11, which is the locus for the Wilm's tumor gene.[79] As in Rb, the peripheral blood lymphocytes of these patients are heterozygous for the Wilm's locus, while tumors are hemizygous. Minicell fusion, which introduces a normal chromosome 11 into Wilm's tumor cell lines, can reverse their tumorigenicity.[80]

Lung Cancer

Both SCLC and squamous cell carcinoma of the lung have been found to consistently contain deletions of a part of the short arm of chromosome 3.[81]

Colon Cancer

Hereditary familial adenomatous polyposis (FAP) has been associated with deletions of chromosome 5.[82] Deletions of the FAP locus have been found in 40% of sporadic cancers.[54,83] In addition 40% of dysplastic polyps are also found to contain deletions of this gene.[54] The loss of this locus in polyps suggests that this is an early event in colonic tumorigenic progression. Additional chromosomal deletions have been found on chromosomes 18 and 17, but only in cancers, which suggests that these are late tumorigenic events.[54,84] The portion of chromosome 17 that is lost in these cancers is thought to contain the p53 gene.[78]

Multistep Carcinogenesis and Mechanisms of Tumorigenic Progression

The concept that cancers proceed through discrete, multiple stages before becoming fully malignant was first suggested by epidemiologic studies[85,86] and then from work with chemical carcinogens.[87] The latter studies suggested that two stages, initiation and promotion, were sufficient to produce malignant transformation. However careful studies of the behavior of cells *in vitro* have suggested that there are multiple stages before invasive metastatic malignancy.[88] These stages include subtle acceleration of growth kinetics, subsequent activation of oncogenes, autocrine growth factor production, and sequential chromosome losses and duplications.

A similar scenario can be constructed for cancers *in vivo*.[89] An initial genetic change may result from an inherited susceptibility to a cancer, frequently a recessive suppressor mutation, *i.e.*, Rb or the FAP mutation, or from environmental carcinogens that activate specific oncogenes such as c-Ha-*ras*. Activation of a single dominant oncogene usually produces increased cellular proliferation. The increased DNA synthesis can result in an increased frequency of mistakes during replication, which may go unrepaired and become permanent mutations. If these mutations occur in additional oncogenes there is a further stimulus to continued DNA replication. During mitosis whole chromosomes or parts of chromatids may be may be lost or rearranged. An augmented rate of DNA synthesis may also predispose to this type of replication error. If a critical suppressor oncogene is lost, the last cellular restraint over continual cell growth may be removed, leading to malignancy. Further mutations may produce cells capable of metastasis.[90] The importance of the multistage concept is that one single oncogene or lost suppressor oncoprotein may be necessary but not sufficient to produce the complete malignant phenotype.[89]

The Future

The theme of this review is that malignant transformation and the subsequent maintenance of the malignant state is orchestrated by molecules with specific functions. Cancer research is uncovering the intracellular, extracellular, and physiologic mechanisms of each of these molecules. In addition the molecular sequence of multistep carcinogenesis is being elucidated. The benefits of this research will surface as improved screening to detect high-risk individuals, more accurate assessment of the likelihood of occult residual disease, and more potent, specific antineoplastic therapies.

In the future a routine physical examination may include DNA analysis of colonic mucosa from colonoscopy specimens or from benign breast biopsies to uncover RFLP patterns, which indicate an inherited susceptibility to carcinoma in those organs. Early dietary modification or hormonal modulation could then be instituted. Resected cancers will be examined for mRNA oncogene expression patterns and RFLP analysis looking for lost suppressor oncogenes to predict those tumors that will locally or distantly recur and also those organs that will be the likely 'fertile soil' for metastasis. Systemic or local infusional drug therapy to susceptible target organs could be administered soon after resection. Even radical adjuvant therapy would be justified if the eventual site of metastasis were predictable with precision. The first beginnings of molecular antineoplastic therapy has arrived, with