New Intracellular Targets for Therapeutic Drug Design

Joan S. Brugge

The traditional approach to the development of therapeutic drugs involves screening natural products or synthetic compounds for substances that interfere with biological events associated with disease. In the past, this random approach was necessitated by our limited knowledge of the molecular etiology of human disease. During the last decade, however, the explosive development of molecular genetic techniques has led to the identification of key proteins that regulate normal biological processes, and, in some instances, has elucidated the cellular pathways responsible for disease. This information can now be exploited for the design of therapeutic drugs that specifically target these pathways.

Signal Transduction

The activity of cells is controlled by external signals that stimulate or inhibit intracellular events. The process by which an external signal is transmitted into and within a cell to elicit an intracellular response is referred to as signal transduction. Signal transduction is generally initiated by the interaction of extracellular factors (for example, hormones, adhesion molecules, neurotransmitters) with membrane receptors on the cell surface. These extracellular signals are transduced to the inner face of the cell membrane, where the cytoplasmic domains of receptor molecules make contact with intracellular targets. The initial receptor-target interactions stimulate a cascade of additional molecular interactions involving multiple intracellular pathways that disseminate the signal throughout the cell. These complex, branching pathways coordinate the multifunctional cellular programs that trigger changes in cell behavior.

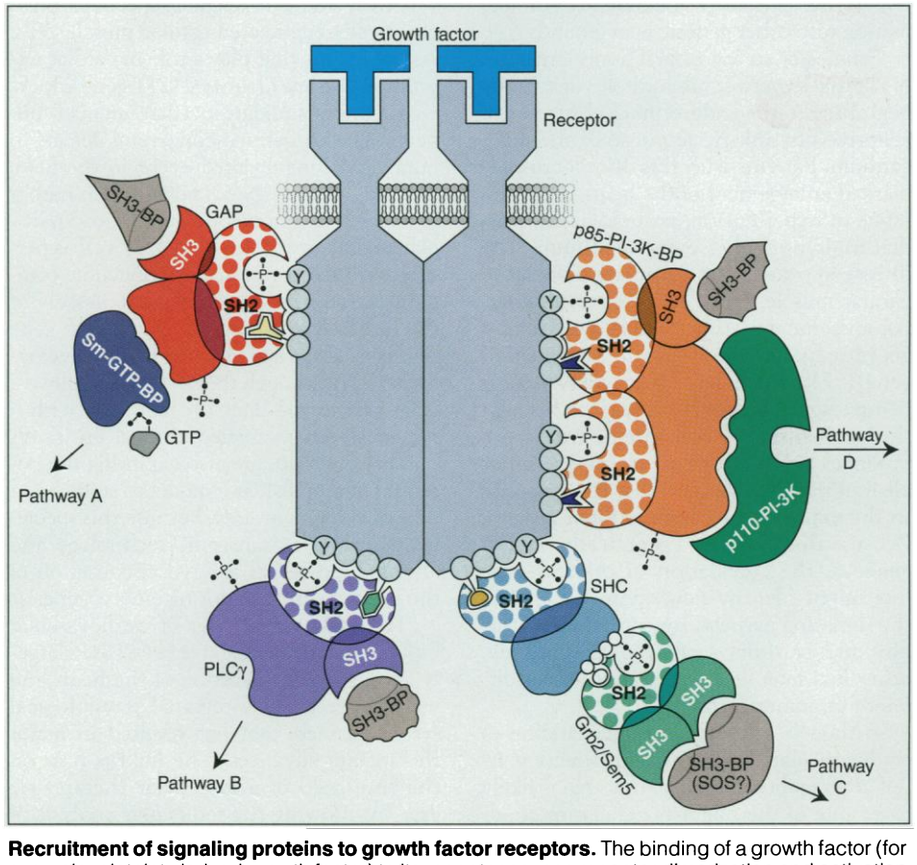
How are all of these intracellular pathways coordinated through the cytoplasm and nucleus, which each contain thousands of distinct proteins? How do individual proteins that function in a pathway find each other to relay the signal down the pathway? The orchestration of diverse proteins in finely tuned intracellular pathways appears to require transient "compartmentalization" of the proteins into complexes. Through a series of inducible and reversible protein-protein interactions, regulatory proteins are recruited from soluble cell material to form short-lived protein complexes that relay signals throughout the cell.

The author is at ARIAD Pharmaceuticals, Inc., 26 Landsdowne Street, Cambridge, MA 02139–4234.

The structural nature of these protein interactions is emerging through the identification of the individual proteins that participate in each signal transduction pathway, the elucidation of the temporal order in which these proteins interact, and the definition of the sites of contact between the proteins. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy studies have provided detailed structural information on a few of these interactive protein domains.

Many of the proteins involved in signal transduction consist of multiple domains, some of which have enzymatic activity and some of which bind to other cellular proteins, DNA regulatory elements, calcium, nucleotides, or lipid mediators. This "modular" structure is ideally suited for signaling events that require both binding interactions (to relay and compartmentalize signals) and enzymatic functions (to effect changes in cellular functions). Remarkably, the DNA sequences encoding these modular domains have been shuffled rather extensively throughout evolution. As a result, one specific domain can be found in 30 to 100 or more distinct proteins. The amino acids within each shuffled domain have diverged, however, and this divergence appears to have subtle effects on the binding specificity of the domains.

With this new wealth of information on the molecular interactions that regulate cellular responses, it is now feasible to develop rationally designed drugs that specifically interfere with the critical molecular interactions that underlie disease processes. The legitimacy of targeting signaling pathways for therapeutic intervention has been demonstrated by the immunosuppressive drugs, cyclosporin A and FK506 (*1*). Both of these drugs bind to calcineurin, a critical intracellular component of the T cell activation pathway that regulates the T cell–specific transcrip-



Recruitment of signaling proteins to growth factor receptors. The binding of a growth factor (for example, platelet-derived growth factor) to its receptor causes receptor dimerization and activation of the tyrosine protein kinase activity of the receptor. These events lead in turn to autophosphorylation at several sites on the receptor. The phosphorylated sites serve as binding sites for proteins with SH2 domains [shown are $p21^{ras}$ guanosine triphosphatase–activating protein (GAP), p85 phosphatidylinositol 3' kinase (PI-3K) binding protein, SHC, and phospholipase C-γ (PLC-γ)]. Each of these proteins has a modular structure, consisting of several protein domains. In some cases, the phosphorylation sites on these SH2-containing proteins serve as binding sites for further SH2-binding interactions (for example, for Grb2/Sem5 binding to SHC). The SH3 domains of these proteins presumably contain sites for distinct binding proteins [SH3-binding proteins (BP); (*2*)].

tion factor, NF-AT. Although these drugs were discovered fortuitously, they prove that intervention in intracellular signaling pathways can effectively alter the course of a disease.

Intracellular Protein-Protein Interactions

Protein-protein interactions are involved in all stages of the intracellular signal transduction process—at the plasma membrane, where the signal is initiated in the cytoplasm by receptor recruitment of other cellular proteins, in the cytoplasm where the signals are disseminated to different cellular locations, and in the nucleus where proteins involved in transcriptional control congregate to turn on or turn off gene expression. Although the structural features of each interaction are uniquely adapted for the specific functions regulated by the interaction, some general concepts can be illustrated through the specific example of growth factor receptors (see figure).

The binding of growth factors to their membrane receptors activates a cascade of intracellular pathways that regulate phospholipid metabolism, arachidonate metabolism, protein phosphorylation, calcium mobilization and transport, and transcriptional regulation. These signaling events can induce changes in cell shape, mobility, and adhesiveness, or stimulate DNA synthesis. Aberrations in these growth factor–induced events are associated with a variety of hyperproliferative diseases ranging from cancer to psoriasis.

Growth factor receptors contain distinct binding sites that serve to recruit multiple signaling molecules through protein-protein interactions. Receptor engagement by growth factors stimulates their protein tyrosine kinase activity and subsequent autophosphorylation on multiple tyrosine residues. These phosphorylated tyrosine residues and surrounding amino acids serve as high-affinity binding sites for cellular proteins that carry certain recognition domains, referred to as src homology-2 (SH2) domains (*2*). These SH2-phosphopeptide interactions recruit other signaling molecules to the receptor, where they can be phosphorylated by the receptor. Both the recruitment and phosphorylation of these receptor targets appear to play important roles in signal transduction; for example, if the target is an enzyme, the recruitment brings the enzyme to the cellular location of its substrates and the tyrosine phosphorylation can stimulate or inhibit its catalytic activity.

Each of the receptor binding proteins controls a cellular pathway involved in the biological response to the growth factor. Activation of a particular pathway could be inhibited by designing a small molecule that specifically disrupts one of the receptor-target interactions. Such an inhibitor could be developed by modern screening methods that are based on natural products from fermentation broths or randomly generated "libraries" of synthetic peptides or other organic compounds. Alternatively, the inhibitor could be developed by structure-based strategies that involve the design of small organic molecules that mimic the structure of the phosphopeptide binding site (peptide mimetics). Either strategy could potentially identify molecules that would bind to a specific SH2 domain and block its interaction with receptor molecules. The advantage offered by structure-based approach is that lead compounds could be optimized by iterated cycles of structural analysis, computational modeling, and synthetic chemistry.

Structure-Based Drug Design

Many features of the SH2-phosphotyrosine interactions make them especially attractive for structure-based drug design: (i) Phosphopeptides as small as five amino acids can interfere with these binding interactions in vitro (*3*). The small size of the binding sites and the fact that they are derived from a sequence of contiguous amino acids should facilitate the design of peptide mimetics. (ii) Although the binding affinity of these interactions is high (at least nanomolar affinity), the rate of dissociation is also high (*4*). Thus, it should be feasible for a peptide mimetic inhibitor to compete effectively for the natural ligand. (iii) SH2 domains can be expressed as soluble, independent domains and they readily form crystals, which are critical for structural analyses (*5*). (iv) SH2 domains are small enough to be analyzed by high-resolution solution NMR (*6*). This property will facilitate the optimization of lead compounds through structural analyses of the "fit" between the SH2 domain and the peptide mimetics.

Although these properties specifically apply to SH2 domains, other reiterated protein binding domains [for example, leucine zippers (*7*), TAM/ARH motifs (*8*), SH3 domains (*9*), zinc fingers (*10*)] that are involved in intracellular signaling pathways have similar structural features; that is, they contain fewer than 100 amino acids, are amenable to NMR and x-ray structural analysis, and have binding sites consisting of a short sequence of contiguous amino acids.

The generation of peptide mimetics derived from structural analysis of protein interaction sites is clearly an attractive strategy for drug design; however, the application of these mimetics to human disease therapy poses several challenges, including those encountered in drug delivery to intracellular targets and with the potential redundancy of cellular pathways. Conceivably, many of these problems could be overcome by carefully selecting targets and optimizing drug dosage and delivery. Another issue is specificity. Given that each of these domains is found in multiple cellular proteins, how will specificity be incorporated into the design of peptide mimetics? As discussed above, there has been considerable divergence in the amino acid sequences of each class of protein domains. The divergent residues of individual domains confer specificity for binding to structural variants within each ligand binding site [for SH2 domains, ligands with different amino acids surrounding the phosphotyrosine residue (*2*)]. In principle, this specificity could be built into the design of mimetics to ensure that the therapeutic agent would not block all interactions involving a class of protein domains.

The first applications of recombinant DNA technology to human disease therapy launched the present-day biotechnology industry. This technology resulted in the large-scale production of recombinant protein factors and therapeutic antibodies. Although these products have had a substantial impact on human disease, their applicability is limited to acute disease states where these short-lived drugs can be delivered directly to their site of action. "Second-generation" biopharmaceuticals, consisting of small organic molecules that can be used to treat chronic diseases, will be designed through exploitation of the fundamental information on the molecular basis of human disease, and through applications of recent advances in high-resolution structural analysis and computational modeling. The protein-protein interactions involved in intracellular signal transduction, which have been emphasized here, offer attractive targets for the design of small molecule drugs. However, other intracellular molecular interactions that involve the binding of proteins to DNA (transcription factor binding to DNA regulatory elements), to lipid mediators (diacylglycerol activation of protein kinase C), and to nucleotides [adenosine 3',5'-monophosphate (cAMP) activation of protein kinase A] are also potentially important targets.

Although there are presently only a few examples of the successful application of structure-based drug design, the continuing improvements in structural analysis and computational methods, and the identification of new targets that may be more amenable to this technology, provide a strong rationale for further efforts to bring this technology to bear on human disease.

References

1. S. L. Schreiber, *Cell* **70**, 365 (1992).
2. C. A. Koch, D. Anderson, M. F. Moran, C. Ellis, T. Pawson, *Science* **252**, 668 (1991); L. C. Cantley *et al.*, *Cell* **64**, 281 (1991); M. J. Pazin and L. T. Williams, *Trends Biochem. Sci.* **17**, 374 (1992).
3. W. J. Fantl *et al.*, *Cell* **69**, 413 (1992).
4. S. Felder *et al.*, *Mol. Cell. Biol.* **13**, 1449 (1993).
5. G. Waksman *et al.*, *Nature* **358**, 646 (1992).
6. G. W. Booker *et al.*, *ibid.*, p. 684.
7. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
8. L. Samelson and R. D. Klausner, *J. Biol. Chem.* **267**, 24913 (1992).
9. T. Pawson and G. Gish, *Cell* **71**, 359 (1992).
10. G. Jacobs and G. Michaels, *New Biol.* **2**, 583 (1990).