Proximity versus allostery: the role of regulated protein dimerization in biology

Regulated dimerization of proteins is increasingly understood to be important in many cellular processes, including signaling, transcription and protein degradation. Organic molecules that induce dimerization may offer as much potential to regulate biological processes as those that allosterically induce conformational change.

Chemistry & Biology November 1994, **1**:131–136

Several different mechanisms are used to transfer information in biological systems. The most familiar one is that of induced conformational change, relying on the ability of ligands to induce an allosteric change in their receptors. But a second mechanism for information transfer, only relatively recently uncovered by cell biologists, is equally important. This is the regulated association of specific proteins, 'protein dimerization'. A new class of organic molecules that can induce the association of specific proteins can be envisaged; such molecules may allow the regulation of biological systems, like the classical allosteric agents that have been a primary focus of biological and medical research for many years.

Cell-surface receptors

The two processes of allosteric change and receptor dimerization are clearly contrasted in two types of cell surface receptors that activate intracellular signaling pathways, the G-protein coupled receptors and the growth factor or growth hormone receptors (Fig. 1). Both activate signaling pathways in response to external binding events, using distinctly different mechanisms. The G-protein coupled receptors are allosterically activated when ligands bind to their transmembrane domain. This ligand-induced conformational change allows the cytoplasmic loops to activate an associated GTP-binding protein (G-protein). Receptor-associated G-proteins are trimeric, consisting of an α subunit that binds guanine nucleotides and β and γ subunits; the α and $\beta\gamma$ subunits have independent signaling functions. The ligand-induced change in receptor conformation is thought to promote the release of GDP so that GTP can bind in its place. Once GTP is bound, the activated Gα-GTP complex and the $\beta\gamma$ subunits separate from each other and dissociate from the membrane, then propagate the signal further by interacting with downstream effectors [1].

The growth-factor-type receptors are activated by a ligand-induced protein dimerization. Hormones and cytokines such as erythropoietin, granulocyte colony-stimulating factor and human growth hormone (hGH) can cross-link two receptors, resulting in the juxtaposition of two cytoplasmic tails. Many of the dimerization-activated receptors (for example, the receptor tyrosine kinases) have protein kinase domains within their cytoplasmic tails that phosphorylate the neighboring tail upon dimerization. In related receptors, such as the cytokine receptors, the cytoplasmic tails lack intrinsic kinase activity but are functionally similar since they associate with protein kinases. In both cases, the phosphorylation results in the activation of a signaling pathway (see below).

In the case of hGH and its receptor, structural studies have revealed the basis for the dimerization of the extracellular domain, which of course causes the dimerization of the intracellular domain. For this receptor, dimerization brings together the associated JAK protein tyrosine kinases. hGH

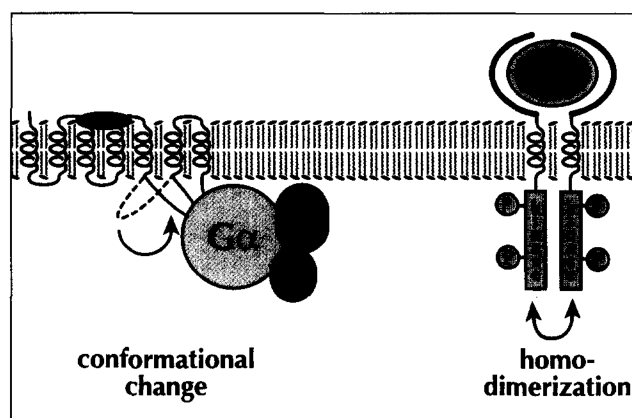


Fig. 1. Allostery and proximity. The allosteric effect of a low molecular weight ligand such as a neurotransmitter on its G-protein-coupled (serpentine or seven-transmembrane spanning) receptor and the dimerizing effect of an extracellular protein such as a growth factor, cytokine or hormone on its normally monomeric receptor. When a G-protein-coupled receptor binds its ligand, a conformational change in the intracellular cytoplasmic loops of the receptor results in the activation of a trimeric G-protein, resulting in GDP/GTP exchange and the dissociation of the Gα subunit/GTP complex from the $\beta\gamma$ subunits. When a growth factor binds to its receptor, however, two receptors are crosslinked. If the cytoplasmic domain of the receptor contains a kinase domain, the two receptor tails will phosphorylate each other.

binds to the extracellular portion of the hGH-receptor with 1:2 stoichiometry by a homodimerization mechanism [2]. The determination of the structure of hGH bound to two extracellular domains provides a fascinating insight into the workings of a natural homodimerizer (Fig. 2) [3]. Although the two receptor molecules that are brought together by a single molecule of hGH are identical, there is no symmetry in the hGH molecule. The two binding events can be clearly distinguished, and must occur in the correct order. The first binding event, although it does not change the shape of the hGH molecule, creates a composite surface that can now bind to the second hGH receptor. Thus, binding at site 1 is essential to allow binding at site 2. By analogy with hGH, we define a dimerizer as an agent that has its effect by bringing two molecules together by binding to both of them. Although it may form only part of the binding site for the second molecule, it does not, or at least need not, change the shape of either molecule to have its effect.

Fig. 3. Bacteriorhodopsin's seven-helical bundle, allosterically regulated by all-*trans*-retinal (left) and 13-*cis*-retinal (right). These are theoretical models based on atomic coordinates determined by electron diffraction and subsequent computational energy optimizations [4]. It is believed that the seven-transmembrane receptors function in similar ways. Generated using GRASP [20].

Allosteric conformational change is important in cell-membrane signaling pathways, the control of transcription, and the activity of several enzymes, to name just a few examples. But regulated protein dimerization is important in at least as many different classes of cellular processes, including diverse signaling pathways, immunological recognition, transcription and the control of protein degradation. We will first examine a few examples of the importance of allosteric change before turning to areas in which regulated dimerization is important.

Allostery

An allosteric conformational change is one that is initiated by the binding of a ligand (even a small molecule) to a protein or macromolecule, but takes place at a location distant from the binding site. In most cases this ligand-induced conformational change influences the activity or function of the protein or macromolecule. Nature and synthetic chemists have both provided many low molecular-weight molecules that can induce allosteric changes in their receptors with dramatic cellular consequences.

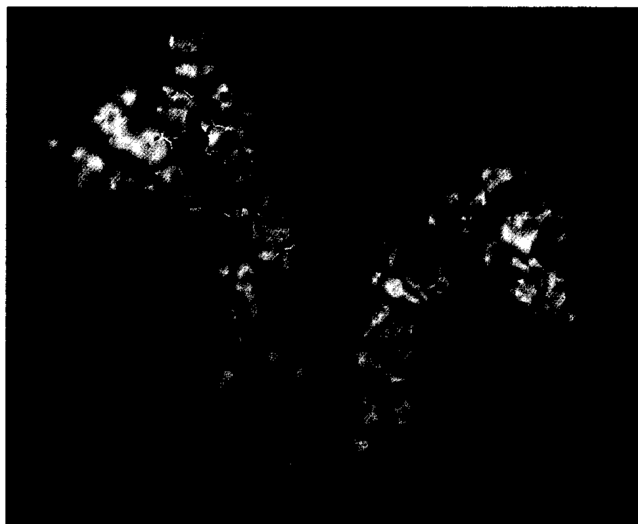


Fig. 2. The homodimerizer human growth hormone (hGH, red) bound to two identical extracellular domains of its receptor (blue). The receptor domains are shown as molecular surfaces, and a sampling of the hGH protein side chains is also shown. A remarkable feature of the hGH homodimerizer is that it has two distinct receptor-binding surfaces that each bind the hGH receptor differently, yet the binding site used by the receptor is essentially the same in both cases. Generated using the program GRASP [20].

Retinal is an example of a molecular switch acting on the halobacterial proton pump, bacteriorhodopsin. Bacteriorhodopsin consists of a seven-helix transmembrane protein, but is not a receptor and is not linked to a trimeric G-protein. Instead, it is covalently bound to all-*trans*-retinal, which upon photolysis isomerizes to 13-*cis*-retinal. This event results in a conformational change in the protein's helical bundle, opening an ion channel and inducing a proton concentration gradient across the cellular membrane (Fig. 3; [4]). This gradient drives ATP synthesis in the bacterium.

An elegant example of allosteric protein activation is the metabolic self-regulation of tryptophan biosynthesis. The *trp* repressor is a dimeric protein that regulates the transcription of the *trp* operon, which encodes the tryptophan biosynthesis genes. In the absence of tryptophan, the protein is inactive and the *trp* operon genes are transcribed, allowing the biosynthesis of tryptophan. Once the protein binds to tryptophan, it changes its conformation, and becomes activated (Fig. 4; [5,6]). The active conformation of the repressor allows it to bind specifically to the operator region of the *trp* operon. When the operator site is bound, the promoter region is blocked, preventing transcription. Therefore, in the presence of excess tryptophan, tryptophan biosynthesis is turned off.

The anti-HIV agent nevaripine [7] is a synthetic molecule that inhibits reverse transcription of viral RNA mediated by HIV-1 reverse transcriptase (RT). The proposed mode of action is based partly on the crystal

structure of Steitz and coworkers (Fig. 5; [8]), which shows that the RNA template that RT copies into DNA binds in a large cleft with a shape resembling that of a right hand. The subdomains of the cleft are termed the 'fingers', 'palm' and 'thumb', and the function of the enzyme appears to involve a conformational change of the thumb with respect to the palm and fingers. The nevaripine-bound conformation of HIV-1 reverse transcriptase is thought to inhibit polymerase activity by crimping or restricting the motion of the 'thumb'.

Thus allosteric change occurs in diverse types of proteins, is induced by diverse ligands, and has a variety of effects. Such allosteric changes have often been the target of drug development efforts, and agents that affect them will no doubt continue to be important in improving our understanding of cell biology as well as in medicine.

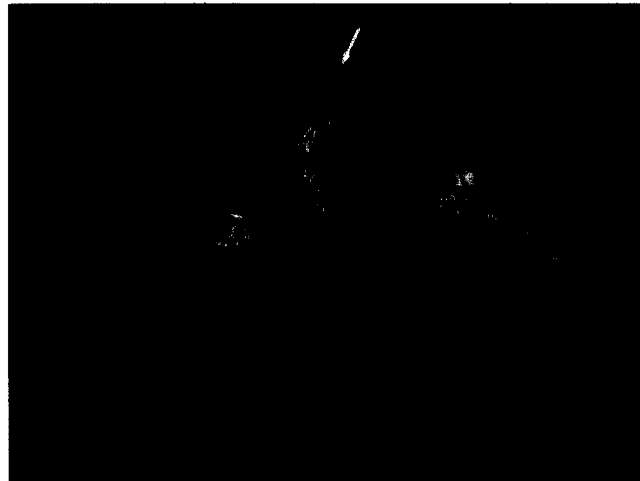


Fig. 5. Nevaripine (yellow) bound to the p66 domain (blue) of HIV-1 reverse transcriptase. The p51 domain (red) serves as a scaffolding for the p66 domain and the RNase H domain (white). The orange spheres represent the polymerase active site and the lavender balls show the RNase H active site. Modeled into the structure is single-stranded RNA (green) and single-stranded DNA (white). Courtesy of Dr Julian Adams, Myogenics Inc., Cambridge, MA. The action of the 'thumb' (arrow) is thought to be essential for DNA polymerization activity, and to be blocked by nevaripine.

Induced proximity

The homodimerizing action of growth factors and hormones, of which hGH is one example, sets in motion intracellular signaling pathways that also depend on the proximity induced by regulated protein dimerization. Many growth-factor receptors activate the well-studied Ras pathway, in which several contingent protein heterodimerizations result in the translocation of signaling proteins to the inner leaflet of the plasma membrane (Fig. 6; [9]). The first step in the pathway initiated by homodimerization of the epidermal growth factor (EGF) receptor, for example, is *trans*-phosphorylation of tyrosines in the tail of the membrane receptor. The phosphorylated receptor tail can now bind the protein Grb2, which contains one SH2 domain that binds phosphotyrosine residues in the receptor tail [9]. The molecular details of this association are now understood (Fig. 7; [10]). Grb2 is a heterodimerizing agent; as well as the SH2 domain that permits association with phosphorylated receptor tails, it also contains two SH3 domains that bind a number of proteins including an activator of Ras named Sos. Ras is a GTP-binding protein, but is not one of the class of heterotrimeric G proteins that is directly activated by the seven-transmembrane spanning receptors shown in Fig. 1; however, it is activated in the same way, by the exchange of GDP for GTP. Ras carries farnesyl and palmitoyl groups, and therefore resides in the inner membrane; when Grb2 moves to the membrane with Sos, it thus brings Sos into proximity with Ras. Sos can then promote the exchange of GTP for GDP, activating Ras. Unlike the G-protein coupled receptors, which require allosteric activation to promote GDP release from the heterotrimeric G proteins, Sos only needs to be brought close to its target to initiate signalling. Thus all that is necessary for signaling is recruitment of Sos to the membrane by a heterodimerizing agent, in this case Grb2 [11].

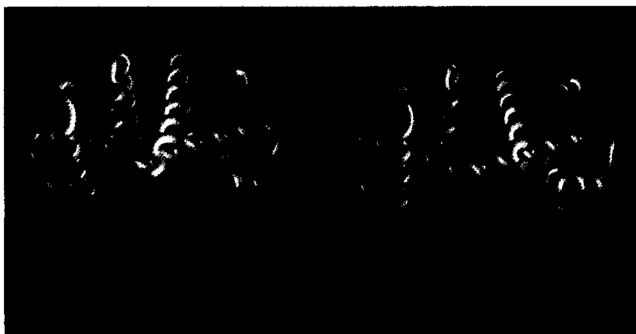


Fig. 4. Turning off transcription. The *trp* aporepressor (left) and the altered conformation of the activated repressor (right) bound to tryptophan (green) and duplex DNA. Generated using GRASP [20].

Organic heterodimerizers

Once Ras is bound to GTP, yet another inducible protein-dimerization event results. Activated Ras binds to the serine/threonine kinase Raf and recruits it to the inner membrane; at present, it is unclear how GTP binding allows Ras to bind to Raf. One possibility is that GTP acts like a heterodimerizer, binding simultaneously to both Ras and Raf.

An analogous situation exists with the low-molecular-weight, organic heterodimerizers cyclosporin A (CsA) and FK506 (Fig. 8) [12]. By binding tightly to the soluble, cytosolic protein FKBP and forming a composite surface, FK506 recruits it to the protein phosphatase calcineurin, which is associated with intracellular membrane components via its myristic acid moiety. CsA similarly binds cyclophilin and calcineurin simultaneously, and the result in both cases is the inhibition of calcineurin function, which is required for intracellular transmission of the signal that emanates from the activated T cell receptor. Because of the immunosuppressive effects of their ligands, FKBP and cyclophilin are known as immunophilins.

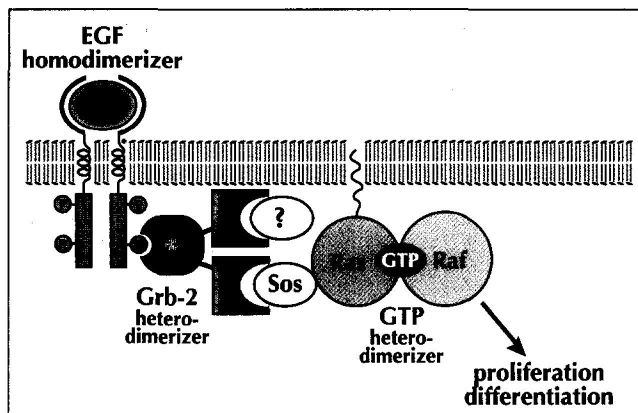


Fig. 6. The early steps of the Ras pathway, illustrating the role of homodimerizers and heterodimerizers. Once the cytoplasmic domain of the growth factor receptor has been phosphorylated (see Fig. 1), it can bind the heterodimerizer Grb-2. This in turn binds Sos, a guanine nucleotide exchange factor for Ras, bringing it to the membrane and into proximity with Ras, which is held near the membrane by farnesyl and palmitoyl groups. The exchange of GDP for GTP creates a binding site for Raf; thus, GTP in this context can be considered to be a heterodimerizing agent.

CsA and FK506 clearly make contact both with the immunophilin and with calcineurin. Like the hGH receptor homodimerization induced by hGH, however, heterodimerizer binding to the second protein is dependent on the first binding event. Thus, CsA does not bind to calcineurin unless it has already associated with cyclophilin, nor can FK506 bind to calcineurin in the absence of FKBP. In both cases, a composite surface composed of both the immunophilin and its organic ligand is formed that is recognized by calcineurin. These natural products and others that, like them, can have a biological effect by bringing two proteins into close proximity, fall into a new class of 'chemical inducers of dimerization' (CID) that seems certain to grow in the future.

The view of these natural products as small molecule heterodimerizers, equipped with two protein-binding surfaces, reveals another fascinating aspect of natural dimerizing agents. They can cause a single protein to bind to multiple protein partners, with specificity determined by the dimerizer. An example is seen in the case of FKBP12. This immunophilin binds to another natural product and heterodimerizer, rapamycin. Rapamycin binding creates a composite surface distinct from the one formed with FK506, resulting in the heterodimerization of another signaling protein, the putative lipid kinase FKBP–rapamycin-associated protein (FRAP) [13,14]. FKBP12 binds neither calcineurin nor FRAP, but in the presence of the heterodimerizers FK506 or rapamycin it can form two distinct receptor–ligand–receptor complexes (Fig. 9), FKBP12–FK506–calcineurin and FKBP12–rapamycin–FRAP. The former complex prevents resting cells from entering into the cell cycle, while the latter prevents cells from progressing through the first gap phase (G1) of the cell cycle.

Peptide dimerizers

The most remarkable illustration of the ability of dimerizers to increase the number of targets an individual protein can bind to is surely that of the MHC class I and II antigen-presenting molecules. Any given member of this family can bind many different peptides; these peptides are routinely produced by degradation of proteins within the cell. In most cases the peptide is derived from a self protein, but in some cases it results from the degradation of a viral or bacterial peptide. MHC class I and II molecules select peptides from this mixture of self and foreign peptides according to the preferences of their binding sites, and present them to T cells for identification. Depending on which peptide has been selected, the surface of the MHC-peptide complex will vary slightly in shape, which in turn will determine which of the many T cell receptors available will bind to the complex [15]. The antigenic peptides can be viewed as heterodimerizers that elicit a biological response by forming a composite surface with their MHC receptor, thereby inducing a specific protein–protein interaction that otherwise would not occur.

The importance of dimerizers and inducible proximity is not restricted to the early and intermediate events of signaling pathways. The culmination of many of these pathways is the activation of transcription in the nucleus, and here, again, regulated heterodimerization is important. Indeed, transcriptional activators themselves can be viewed as heterodimerizers composed of DNA-binding and protein-binding surfaces (Fig. 10). Class II nuclear genes have an element within their promoter region known as the TATA box, a short sequence composed entirely of T-A base-pairs. The TATA-box binding protein (TBP) is required for initiation of transcription of all such genes. TBP alone does not efficiently activate transcription, however; it requires a number of transcription factors and TBP-associated factors (TAFs).

Fig. 7. Crystal structure of the SH2 domain found in the non-receptor tyrosine kinase Src, bound to the phosphorylated peptide ligand pYEEI. The heterodimerizer Grb-2 is expected to bind its phosphorylated substrate in a similar way. This structure thus gives insight into one of the two protein–ligand interactions required for a natural dimerizer to function. Generated using GRASP [20].

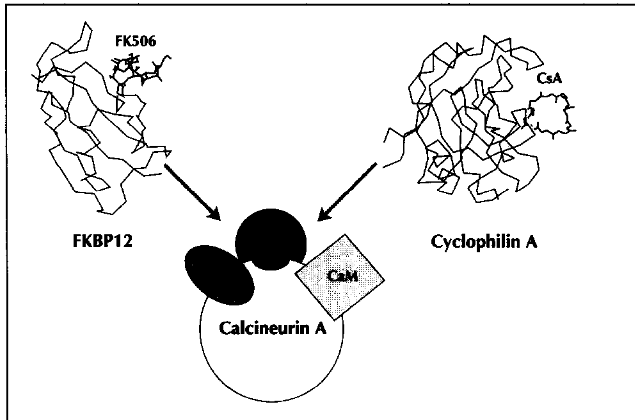


Fig. 8. FK506 and CsA both function as heterodimerizers by causing immunophilins to interact with the signaling protein calcineurin. These ligands create a composite surface following binding to their immunophilin receptors that results in the formation of high-affinity ligands to calcineurin. CNB, calcineurin subunit B; CaM, calmodulin; Imm, immunophilin; L, ligand.

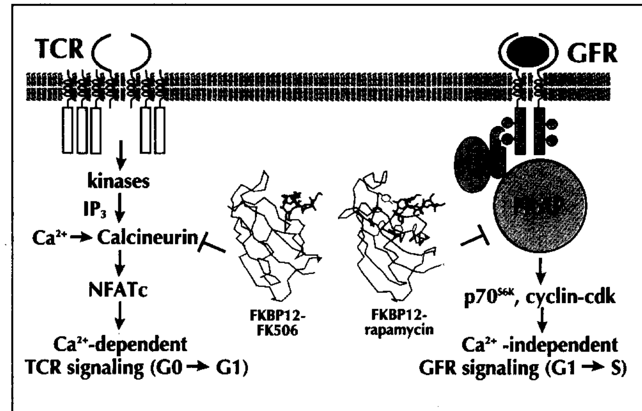


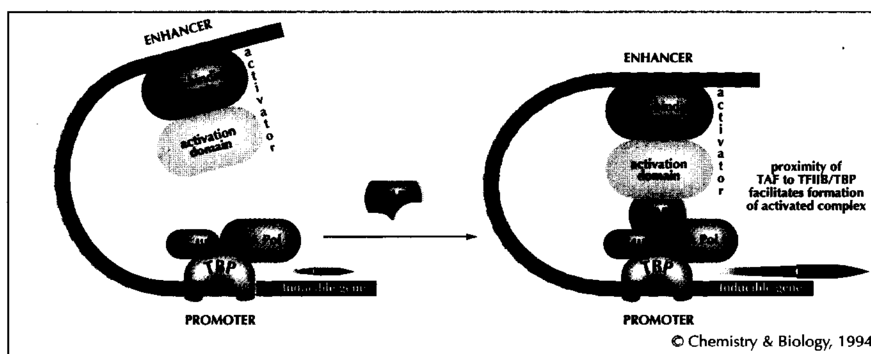
Fig. 9. FK506 and rapamycin both bind FKBP12, but each forms a different composite surface resulting in the heterodimerization of the immunophilin with different signaling proteins thereby resulting in different biological outcomes. In the T cell receptor (TCR) signaling pathway, engagement of the TCR by an appropriate MHC–peptide complex initiates signaling via inositol triphosphate (IP_3) and calcineurin, which eventually leads to the activation of the cytoplasmic component of the transcription factor NFAT ($NFAT_C$), and entry of cells into the cell cycle. The FKBP12–rapamycin complex, on the other hand, blocks the signaling pathway initiated by the IL-2 receptor and several other cytokine and growth factor receptors (GFRs) by binding to FKBP–rapamycin-assocated protein (FRAP). The FRAP signaling pathway activates the protein $p70^{S6K}$ and cyclin-dependent kinases (cdks), and causes the cell to make the transition from the first gap phase of the cell cycle (G1) to the synthesis (S) phase.

Enhancers that specifically regulate the transcription of these genes do so by binding to an enhancer sequence some distance from the site of initiation and recruiting a TBP-associated factor (TAF) to the vicinity of the TATA box, facilitating the complete formation of an initiation complex. Thus, the activator uses its ability to bind both DNA and TAF to bring the latter into close proximity with the neighboring promoter sequence [16].

Although signaling pathways are perhaps the richest source of examples of the induction of proximity using molecular dimerizers, this mechanism of information transfer is not limited to signaling. We have already seen that proteins can be caused to translocate to various parts of the cell by dimerizers. The recruitment of signaling molecules to the inner membrane serves to illustrate how this process can activate the protein, by bringing it into close proximity with its substrate. Proteins can also be induced to translocate to a multiprotein complex known as the proteasome, which degrades cellular proteins (providing, as well as a mechanism of protein removal, a source of peptides for presentation by MHC molecules). An interesting example of this process involves the human papilloma viral (HPV) protein E6 [17]. The function of E6 appears to be to direct the tumor suppressor protein p53, produced by the host cell, to the proteasome, where it is proteolytically degraded. This seems to be an important factor in the ability of high-risk strains of HPV to cause cervical cancer. E6 binds to a cellular protein, the E6-associated protein (E6AP); the complex then binds to p53 and directs it to a set of enzymes involved in the ubiquitination of lysine sidechains (Fig. 11). Ubiquitinated proteins are 'tagged' for degradation by the proteasome. E6 is thus acting as a heterodimerizing agent, bringing p53 close to the machinery that identifies proteins for removal. The structural aspects of this overall process have not yet been studied in detail, but the overall process is another clear illustration of a proximity effect. Proteins do not wander stochastically on the way to their fateful encounter with the proteasome, but are induced to become targets for it by the action of molecular dimerizers.

Synthetic dimerizers: new opportunities for intervention

Our familarity with the concepts of conformational change induced by allosteric agents has facilitated the discovery of



© Chemistry & Biology, 1994

Fig. 10. Mechanism of transcriptional activation by transcriptional activators involving the recruitment of general transcription factors to a promoter sequence.

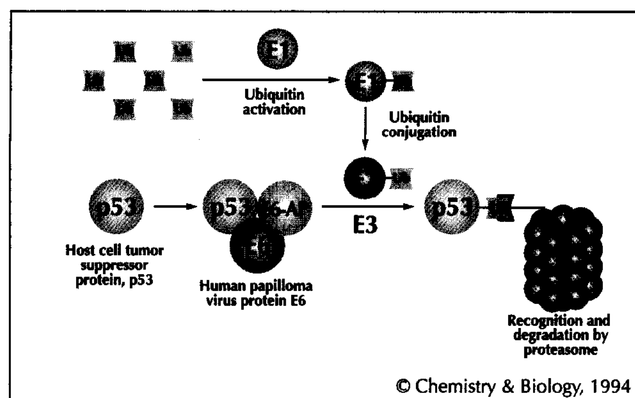


Fig. 11. Human papilloma virus-mediated ubiquitination of the tumor suppressor p53. Ubiquitination of the target protein occurs through initial conjugation of the ubiquitin-activating enzyme (E1) with the carboxy-terminal glycine of ubiquitin (Ub) via a cysteine residue, followed by a cysteine residue transfer from E1 to ubiquitin-conjugating enzyme (E2/E3). E6 and the E6-associated protein (E6-AP) bind specifically to the tumor suppressor p53, and, acting as an adaptor, work with E2 to facilitate the degradation of p53 via the ubiquitination pathway.

synthetic versions of these molecules. Many examples of synthetic allosteric agents can be found in medicine, and some of these have facilitated the study of basic cellular mechanisms. By recognizing the prominent role of natural dimerizers in cell biology, the properties of organic equivalents can now also be readily and reliably predicted. Formulating the structures of CIDs and synthesizing them will be a new, and worthwhile, challenge. Nature has provided clear illustrations of the feasibility of this approach in the form of the natural products CsA, FK506, and rapamycin. These low-molecular-weight, organic heterodimerizers are monomeric molecules equipped with two distinct protein-binding surfaces. One approach to constructing a designed CID involved converting the monomeric heterodimerizer FK506 into the dimeric homodimerizer FK1012 by eliminating one of FK506's protein-binding surfaces and replacing it with an element that crosslinked the modified natural product [18,19].

It is easy to imagine the mixing and matching of different protein-binding surfaces using synthetic organic chemistry, to create new dimerizers with tailor-made properties. Since protein dimerizers simply create a high local concentration of a particular protein at a particular cellular location, their actions will not require the geometric precision associated with allosteric agents. This fact, combined with our increasing ability to generate specific protein-binding surfaces either by screening or by design, suggests that synthetic dimerizers might be even more readily accessible than the classical allosteric agents.

References

1. Linder, M.E. & Gilman, A.G. (1992). G proteins. *Sci. Am.* **267**, 36–43.
2. Cunningham, B.C., Ultsch, M., de Vos, A.M., Mulkerrin, M.G., Clauser, K.R. & Wells, J.A. (1991). Dimerization of the extracellular domain of the human growth hormone receptor by a single hormone molecule. *Science* **254**, 821–825.
3. de Vos, A., Ultsch, M. & Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306–312.
4. Chou, K.-C., Carlacci, L., Maggiora, G.M., Parodi, L.A. & Schulz, M. W. (1992). An energy-based approach to packing the 7-helix bundle of bacteriorhodopsin. *Prot. Sci.* **1**, 810–827.
5. Zhang, R.-G., Joachimiak, A., Lawson, C.L., Schevitz, R.W., Otwinowski, Z. & Sigler, P.B. (1987). The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* **327**, 591–597.
6. Otwinowski, Z., *et al.*, & Sigler, P.B. (1988). Crystal structure of the *TRP* repressor/operator complex at atomic resolution. *Nature* **335**, 321–329.
7. Merluzzi, VJ., *et al.*, & Sullivan, J.L. (1990). Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science* **250**, 1411–1413.
8. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783–1790.
9. Pawson, T. & Schlessinger, J. (1993). SH2 and SH3 domains. *Curr. Biol.* **3**, 434–442.
10. Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. & Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790.
11. Aronheim, A., Engleberg, D., Li, N., Al-Alawi, N., Schlessinger, J. & Karin, M. (1994). Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway. *Cell* **78**, 949–961.
12. Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I. & Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin–cyclosporin A and FKBP-FK506 Complexes. *Cell* **66**, 807–815.
13. Brown, E. J., *et al.*, & Schreiber, S. L. (1994). A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* **369**, 756–758.
14. Sabatini, D.M., Erdjument-Bromage, H., Lui, M., Tempst, P. & Snyder, S.H. (1994) RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* **78**, 35–43.
15. Stern, L.J. & Wiley, D.C. (1994). Antigenic peptide binding by Class I and Class II histocompatibility proteins. *Structure* **2**, 245–251.
16. Tijan, R. & Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77**, 5–8.
17. Scheffner, M., Huibregtse, J.M., Vierstra, R.D. & Howley, P.M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**, 495–505.
18. Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. (1993). Controlling signal transduction with synthetic ligands. *Science* **262**, 1019–1024.
19. Pruschy, M., Spencer, D.M., Kapoor, T.M., Miyake, H., Crabtree, G. R. & Schreiber, S.L. (1994). Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chemistry & Biology* **1**, 163-172.
20. Nicholls, A., Sharp, K.A. & Honig, B. (1991). *Prot. Struct. Funct. Genet.* **11**, 282–293.

David J Austin and Stuart L Schreiber, Howard Hughes Medical Institute, Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA and Gerald R Crabtree, Howard Hughes Medical Institute, Departments of Developmental Biology & Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA.