Small molecule approaches to regulation of gene transcription

Regulation of gene expression is fundamental to the processes of cellular differentiation and organismal development. The human genome project has identified approximately 30,000 protein coding genes in the genome and the regulation of expression of these genes in time and space leads to the complex structure of the human form. How are these genes expressed at the appropriate time and in the appropriate place is a fundamentally important area of study which has yet to be fully defined. Regulation of gene transcription is a primary means of regulating gene expression. Transcription factors, both activators and repressors, are the molecular arbiters regulating gene expression in organisms as diverse as bacteria, yeast and mammals. Aberrant gene transcription plays a role in the onset of many diseases including cancer, neurodegenerative disease and others. In eukaryotes at least, transcriptional regulation involves a complex assembly of activators, repressors, coactivators/corepressors, general transcription factors and RNA polymerases to elicit proper expression of protein coding, tRNA and rRNA genes. In this review we focus on the regulation of protein coding genes by RNA polymerase II and small molecule approaches to regulating this complex process.

Transcriptional activators are a diverse set of molecules that activate gene expression in response to environmental and developmental cues. Activators generally bind to sequence elements upstream of protein coding genes and thus confer specificity of gene expression. Transcriptional activators are modular in nature containing discrete DNA binding domains and discrete activation domains (1). A wide variety of DNA binding domains have been defined both functionally and structurally and include domains such as the Zn-finger domain, the helix-turn-helix domain among others (2). Activation domains are less rigidly defined on the basis of structure and are generally sorted according to the preponderance of amino acids found in the activation domain. Thus activation domains have been grouped as acidic (1), proline-rich (3) and glutamine-rich (4).

Transcription factor modularity underlies current attempts at constructing artificial transcription factors (ATF) that will possess rationally designed activities. Domain swapping experiments used to define the modularity of DNA binding and activation domains were the first ATFs constructed and represent a proof of principle that such approaches were feasible. Thus far, rational design of ATFs has focussed on molecules that mimic or enhance DNA binding specificities as well as molecules that act as activation modules when fused to a DNA binding domain.
Triplex-forming oligonucleotides have been shown to possess DNA binding activity (5). Triplex-forming oligonucleotides fused to the VP16 activation domain have been shown to activate transcription in cell culture (6, 7). Peptide nucleic acids in which the phosphodiester backbone normally found in nucleic acids is replaced with amide bonds have also been used to bind DNA in a designated manner (8). DNA binding modules have also been designed based on the naturally occurring molecule distamycin. DNA binding molecules based on this structure bind to the minor groove of DNA, are cell permeable and have been shown to downregulate endogenous gene expression (9). Dervan and colleagues have gone further and designed a library of 27 molecules based on a hairpin polyamine scaffold which are currently available for testing in vivo (10). A related strategy has been to design zinc finger domains using a genetic selection approach (11). Designed zinc fingers when fused to the VP16 activation domain have been shown to upregulate endogenous erbB-2 and VEGF-A genes in cultured cell lines (12, 13).

Novel peptide activation domains were first identified in a screen of randomized E. coli sequences fused to a natural DNA binding domain (14). Randomized sequences containing hydrophobic sequences activated transcription in yeast as strongly as natural activation domains when fused to a natural DNA binding domain (15). Amphipathic isoxazolidine was one of the first chemical mimics of an activation domain to be discovered (16). Analogs of this structure have been synthesized and shown to possess transcriptional activation capability (17). A different approach involved identification of compounds that bind to coactivators. Based on their ability to interact with downstream proteins in the transcription complex these compounds when fused to DNA binding domains could presumably act as activator molecules in their own right. Two compounds adamanolol and its derivative wrenchnolol were identified as a micromolar ligand for Sur-2 a subunit of the human Mediator complex (18, 19). When fused to a DNA binding polyamide, the wrenchnolol conjugate could activate gene transcription in vitro (20). This hybrid molecule was shown to recruit both Sur2 and RNA polymerase II to the target DNA. Another screen involving the identification of molecules that bind the coactivator CBP/p300 led to the identification of two peptide ligands that bind a key domain of CBP/p300. When fused to a DNA binding domain these ligands could activate transcription in cell culture at levels comparable to a natural activation domain (21). In a library screen for compounds that bind the KIX domain of CBP/p300, Kodadek et al. identified a peptoid compound, KBP02, that binds the KIX domain in vitro and is able to strongly activate transcription in a cell culture based assay (22).

In addition RNA molecules have been found that when fused to a DNA binding domain can also activate transcription in yeast (23). While activation modules have been identified in several molecular forms, the levels of activation still do not approach those of natural activation domains and this remains an outstanding design problem. Transcriptional activation domains interact with a variety of factors to elicit regulation of gene expression. Knowledge of these key biochemical interactions has also been used to identify compounds that perturb these interactions and thus regulate gene expression. One of the most studied activation domain interactions is that of the p53 activation domain with MDM2. p53 is a transcriptional activator and tumor suppressor protein whose normal function is to activate genes involved in cell-cycle arrest, apoptosis and DNA repair (24, 25). MDM2 interacts with the p53 activation domain thereby blocking its transcriptional activation function. Screening and design based strategies have been used to identify a number of molecules that can inhibit the p53-MDM2 interaction (26). One of the first compounds to be identified that inhibits the p53-MDM2 interaction was nutlin. This molecule binds to MDM2 and can inhibit cancer cell growth and kill tumors in a mouse xenograft model after oral administration. Benzodiazepinone was also shown to be a human MDM2 (HDM2) antagonist (27). A distinct class of inhibitor was identified that interacts with p53 as opposed to MDM2. Bisthiophenylfluran 13 (RITA) was identified through a library screen and was shown to completely stop tumor growth in a mouse xenograft model system (28).

The histone acetyltransferase CBP/p300 is another coactivator that binds to many transcription factors to reorganize chromatin and activate transcription. CBP/p300 has several interaction domains that are able to interact with diverse transcription factors and each of these domains is a prime target for the discovery of molecules that inhibit transcription factor interaction. One of the transcription factors that target CBP/p300 is HIF-1alpha, the hypoxia responsive transcription factor. A fungal metabolite, chetomin, was isolated in a high throughput screen for inhibitors of the HIF-1alpha-p300 interaction (29). Chetomin inhibits the p53-MDM2 interaction in vitro as well as inhibiting hypoxia induced transcription in vivo by 50% (29). Because p300 binds many transcription factors, chetomin was also screened for inhibition of transcription by other transcriptional activators and it was shown that STAT2 activation was also inhibited by chetomin (29) indicating that specificity remains an issue in the design of CBP/p300 inhibitors. Using a NMR based screening approach for compounds that inhibit the transcriptional activator CREB and its interaction with CBP, Montminy et al. identified a molecule designated KG-501 as an inhibitor of this protein-protein interaction (30). In cell culture studies, KG-501 inhibited transcription of CREB dependent genes via its direct inhibition of the CREB-CBP interaction. Specificity was again shown to be an issue as KG-501 also inhibits CBP interaction with the transcriptional activator NF-kB.

In yet a further refinement of the construction of artificial transcription factors, several groups have developed ATPs whose activity is ligand dependent thereby introducing the capability of regulating the activation function of these factors in a controlled manner. In the first example, Bujard et al. developed a doxycyclin inducible...
system for regulating transcriptional activation (31). In a second, Schreiber et al. developed chemical inducers of transcription factor dimerization to regulate transcription factor activity (32). In yet another example, an allosteric ligand-responsive ATP was designed such that a zinc finger DNA binding domain was linked to the VP16 activation domain via the ligand binding domain of a steroid receptor. This fusion was responsive to steroid hormone and showed that regulatable transcription factors could be constructed from existing ligand binding domains (33).

The advent of whole genome expression profiling has added a new dimension to the search for small molecules that activate defined transcriptional pathways. Uretupamine was identified in a screen for small molecules that bind the yeast protein Ure2p (34). Subsequent comparison of the whole genome transcription profile of uretupamine treated cells with that of ure2 cells showed that uretupamine affects only a subset of genes affected by ure2 mutation, suggesting a previously unexpected level of specificity of the uretupamine molecule. In a similar manner splitomycin was isolated as a small molecule inhibitor of the histone deacetylase Sir2p and cells treated with splitomycin exhibited a transcriptional profile similar to that of sir2 deleted cells (35). Furthermore, the splitomycin gene expression profile also resembled to a lesser extent the profile of a hst1 deletion strain suggesting that splitomycin might inhibit not only Sir2 but an additional histone deacetylase Hst1. Derivatives of splitomycin were subsequently developed that showed selectivity of specificity of the uretupamine molecule. In a similar manner splitomycin was isolated as a small molecule inhibitor of the histone deacetylase Sir2p and cells treated with splitomycin exhibited a transcriptional profile similar to that of sir2 deleted cells (35). Furthermore, the splitomycin gene expression profile also resembled to a lesser extent the profile of a hst1 deletion strain suggesting that splitomycin might inhibit not only Sir2 but an additional histone deacetylase Hst1. Derivatives of splitomycin were subsequently developed that showed selectivity towards either Sir2 or Hst1 and subsequently it was shown that this selectivity could be corroborated by comparing the gene expression profiles of the derivatives with those of either sir2 or hst1 mutant cells (36). Whole genome expression profiling holds great promise for the identification of small molecules with specific gene expression targeting. By comparing the transcriptional profiles of new compounds with those of known therapeutic classes, predictions of a new compound’s potential therapeutic use and efficacy can be inferred. At present, the expense of microarray technology prohibits use of this technology as a standard screening tool, however an alternative that has been shown effective is the use of only a few marker genes from a complete profile as a sort of mini-array with which to look for compounds that can induce the mini-array. Stegmaier et al. used this approach to identify compounds that induce terminal differentiation of acute myelogenous leukemias cells (37). Compounds that were identified on the basis of their ability to induce a set of marker genes were subsequently shown to also induce the whole genome profile associated with differentiation. Clearly, if a set of marker genes can be inferred from a whole genome transcriptional profile this approach should be applicable to a more efficient screening for new compounds that affect a given transcriptional profile.

Small molecule approaches to the design of synthetic transcription factors as well as identification of compounds that operate through defined transcriptional tar-

**REFERENCES**

3. MERMOD N, O’NEILL E A, KELLY T J, TJIANT R 1989 The proline-rich transcriptional activator of CTP/NF-I is distinct from the replication and DNA binding domain. Cell 58: 741-753

Period Biol, Vol 111, No 1, 2009. 57
Small molecule approaches to regulation of gene transcription