PERIODICUM BIOLOGORUM
VOL. 111, No 1, 55–58, 2009

UDC 57:61
CODEN PDBIAD
ISSN 0031-5362

Short review

Small molecule approaches to regulation of gene transcription

JOSIPA PEŠUT
BRANKA JELIČIĆ
MARY SOPTA

Department of Molecular Biology
Ruđer Bošković Institute
Bijenička 54, 10000 Zagreb, Croatia

Correspondence:
Mary Sopta
Department of Molecular Biology
Ruđer Bošković Institute
Bijenička 54, 10000 Zagreb, Croatia
E-mail: msopta@irb.hr

Received October 30, 2008.

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 occuring 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. Bisthiophenylfuran 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 HIF-1alpha-p300 interaction in vitro as well as inhibiting hypoxia induced transcription in vivo by 50% *(29)*. Because p300 binds many transcription factors, chetomin was also screed 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 ATFs 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 ATF 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 yeast 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 targeting represent a promising avenue of research for the future. In this review we have highlighted a few examples which typify the approaches used thus far to isolate a variety of molecules that directly act on transcription factors and the assembly of a functional transcription complex at the promoters of protein coding genes. While initial progress in the design of synthetic activator molecules is promising thus far these molecules still lack many characteristics of natural activator molecules. Synthetic activators still do not approach the levels of transcriptional activation observed with natural activators, and as well do not show cell or tissue specificity. In addition cell permeability remains an issue for both synthetic activator molecules and small compound regulators. Clearly a convergence of chemical and biological information will be required to advance the field to a point where these approaches will ultimately give rise to new therapeutic agents.

REFERENCES

1. PTASHNE M, GANN A 2002 Genes&Signals. Laboratory Press, Cold Spring Harbor.
2. HARRISON S C 1991 A structural taxonomy of DNA-binding domains. *Nature 353*: 715-719
3. MERMOD N, O'NEILL E A, KELLY T J, TJIAN R 1989 The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell 58*: 741-753
4. GERBER H P, SEIPEL K, GEORGIEV O, HOFFERER M, HUG M, RUSCONI S, SCHAFFNER W 1994 Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science 263*: 808-811
5. FARIA M, GIOVANNANGELI C 2008 Triplex-forming molecules: from concepts to applications. *J Gene Med 3*: 299-310
6. STANOJEVIC D, YOUNG R A 2002 A Highly Potent Artificial Transcription Factor. *Biochemistry 41*: 7209-7216
7. KUZNETSOVA S, AIT-SI-ALI S, NAGIBNEVA I, TROALEN F, LE VILLAIN J P, BELLAN A, SVINARCHUK F 1999 Gene activation by triplex-forming oligonucleotide coupled to the activating domain of protein VP16. *Nucl Acids Res 27*: 3995-4000
8. NIELSEN P E 2001 Peptide Nucleic Acid Targeting of Double-Stranded DNA. *In:* Jonathan B C (*ed*) Methods in Enzymology: Drug-Nucleic Acid Interactions. Academic Press, London, p 329
9. DERVAN P B 2001 Molecular recognition of DNA by small molecules. *Bioorg Med Chem 9*: 2215-2235
10. HSU C, PHILLIPS J, TRAUGER J *et al*. 2007 Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron 63*: 6146-6151
11. BEERLI R R, BARBAS C F 2002 Engineering polydactyl zinc-finger transcription factors. *Nat Biotech 20*: 135-141
12. BEERLI R R, DREIER B, BARBAS C F 2000 Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci U S A 97*: 1495-1500
13. LIU P Q, REBAR E J, ZHANG L *et al.* 2001 Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. *J Biol Chem 276*: 11323-11334
14. MA J, PTASHNE M 1987 A new class of yeast transcriptional activators. *Cell 51*: 113-119
15. LU X, ANSARI A Z, PTASHNE M 2000 An artificial transcriptional activating region with unusual properties. *Proc Natl Acad Sci U S A 97*: 1988-1992
16. MINTER A R, BRENNAN B B, MAPP A K 2004 A Small Molecule Transcriptional Activation Domain. *J Am Chem Soc 126*: 10504-10505
17. ROWE S P, CASEY R J, BRENNAN B B, BUHRLAGE,S J, MAPP A K 2007 Transcriptional Up-regulation in Cells Mediated by a Small Molecule. *J Am Chem Soc 129*: 10654-10655

18. ASADA S, CHOI Y, UESUGI M 2003 A gene-expression inhibitor that targets an alpha-helix-mediated protein interaction. *J Am Chem Soc 125*: 4992-4993
19. SHIMOGAWA H, KWON Y, MAO Q, KAWAZOE *et al.* 2004 A Wrench-Shaped Synthetic Molecule that Modulates a Transcription Factor-Coactivator Interaction. *J Am Chem Soc 126*: 3461-3471
20. KWON Y, ARNDT H D, MAO Q, CHOI Y, KAWAZOE Y, DERVAN P B, UESUGI M 2004 Small Molecule Transcription Factor Mimic. *J Am Chem Soc 126*: 15940-15941
21. RUTLEDGE S E, VOLKMAN H M, SCHEPARTZ A 2003 Molecular Recognition of Protein Surfaces: High Affinity Ligands for the CBP KIX Domain. *J Am Chem Soc 125*: 14336-14347
22. LIU B, ALLURI P G, YU P, KODADEK T 2005 A Potent Transactivation Domain Mimic with Activity in Living Cells. *J Am Chem Soc 127*: 8254-8255
23. SAHA S, ANSARI A Z, JARRELL K A, PTASHNE M 2003 RNA sequences that work as transcriptional activating regions. *Nucl Acids Res 31*: 1565-1570
24. KO L J, PRIVES C 1996 p53: puzzle and paradigm. *Genes Dev 10*: 1054-1072
25. VOGELSTEIN B, LANE D, LEVINE A J 2000 Surfing the p53 network. *Nature 408*: 307-310
26. MAJMUDAR C Y, MAPP A K 2005 Chemical approaches to transcriptional regulation. *Curr Opin Chem Biol 9*: 467-474
27. GRASBERGER B L, LU T, SCHUBERT C *et al.* 2005 Discovery and Cocrystal Structure of Benzodiazepinedione HDM2 Antagonists That Activate p53 in Cells. *J Med Chem 48*: 909-912
28. ISSAEVA N, BOZKO P, ENGE M *et al*. 2004 Small molecule RITA binds to p53, blocks p53-HDM-2 interaction and activates p53 function in tumors. *Nat Med 10*: 1321-1328
29. KUNG A L, ZABLUDOFF S D, FRANCE D S *at al.* 2004 Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell 6*: 33-43
30. BEST J L, AMEZCUA C A, MAYR B *et al*. 2004 Identification of small-molecule antagonists that inhibit an activator:coactivator interaction *Proc Natl Acad Sci U S A 101*: 17622-17627
31. BARON U, BUJARD H 2000 Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol 327*: 401-421
32. BELSHAW P J, HO S N, CRABTREE G R, SCHREIBER S L 1996 Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc Natl Acad Sci U S A 93*: 4604-4607
33. BEERLI R R, SCHOPFER U, DREIER B, BARBAS C F, III 2000 Chemically Regulated Zinc Finger Transcription Factors. *J Biol Chem 275*: 32617-32627
34. KURUVILLA F G, SHAMJI A F, STERNSON S M, HERGENROTHER P J, SCHREIBER S L 2002 Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. *Nature 416*: 653-657
35. BEDALOV A, GATBONTON T, IRVINE W P, GOTTSCHLING D E, SIMON J A 2001 Identification of a small molecule inhibitor of Sir2p. *Proc Natl Acad Sci U S A 98*: 15113-15118
36. HIRAO M, POSAKONY J, NELSON M, HRUBY H, JUNG M, SIMON J A, BEDALOV A 2003 Identification of Selective Inhibitors of NAD+-dependent Deacetylases Using Phenotypic Screens in Yeast. *J Biol Chem 278*: 52773-52782
37. STEGMAIER K, ROSS K N, COLAVITO S A, O'MALLEY S, STOCKWELL B R, GOLUB T R 2004 Gene expression-based high-throughput screening(GE-HTS) and application to leukemia differentiation. *Nat Genet 36*: 257-263