ELSEVIER

Small organic molecules that modulate gene transcription

Dongju Jung[1], Yongmun Choi[1] and Motonari Uesugi[1,2]

[1] The Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA
[2] Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Regulation of gene expression by transcription factors touches all aspects of human biology and often induces extreme phenotypes. Its external, precise control by small organic molecules represents a challenge in chemistry and biology. Here, we summarize recent progress in the field, together with contributions from our laboratory. Small-molecule modulators of transcription, including small-molecule transcription factors could find their use in basic biological studies and therapeutic intervention.

Magical candy

Knowledge about bioactive small-molecules is a treasure to humankind. Small organic molecules with novel biological activities have often served as triggers for improving human health and for revealing many secrets of life. Recent progress in biology and chemistry suggests expansion of the treasure – we might be able to discover small molecules with amazing effects beyond our current imagination.

One illustrative, yet fictional, pharmacological effect was demonstrated in a cartoon movie that was very popular in Japan three decades ago. In the movie, the ten-year-old girl, called Merumo, who lost her parents in a traffic accident, is given a bottle of magical candies by angels. When the girl takes a blue candy, she appears to be ten years older; a red candy has the reverse effect. Through a range of clever uses of the candies, Merumo overcomes difficulties and raises her younger brother.

Merumo's candies are orally available, suggesting that the active ingredients are small organic molecules. The cellular regulation that most-often produces such extreme phenotypes is related to gene transcription, and we believe that its systematic, external control by organic molecules represents a challenge in chemistry and biology.

Small molecules and transcription

Gene transcription is usually regulated by a class of proteins called transcription factors, typically having separable DNA-binding and transcriptional activation domains. The DNA-binding domain provides gene specificity by sequence-specific recognition of appropriate promoters, whereas the activation domain stimulates gene transcription by binding to so-called co-activator proteins and thereby recruiting RNA polymerase II to the promoter. In principle, small molecules designed to modulate the binding of either one of the two domains would be capable of controlling selective gene transcription and might have unique biological or therapeutic effects.

It was originally thought that inhibition of the interaction of a DNA-binding domain with specific DNA sites by small organic molecules would be highly challenging. DNA structures are usually too uniform for small drug-like molecules to discriminate, and the interaction of the DNA-binding domain with DNA is typically spread out over a relatively wide surface area, which small organic molecules might not be able to span. Nevertheless, Dervan [1] and others succeeded in designing several hairpin polyamide molecules that bind to specific DNA sequences, by mimicking the DNA-recognition properties of DNA-binding natural products such as distamycin A. Some of these remarkable molecules are capable of penetrating cell membranes and selectively inhibiting transcription by binding to specific promoter regions of DNA [2–4].

It is also possible to modulate the activity of activation domains. It is well known that nuclear receptor transcription factors are modulated directly by small molecules, including steroid hormones or their mimics. Nuclear receptors are typically composed of a DNA-binding domain, a ligand-binding domain and an activation domain [5]. Binding of small molecules to the ligand-binding

Corresponding author: Uesugi, M. (uesugi@scl.kyoto-u.ac.jp)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2006.03.011

domain induces a conformational change of the activation domain, modulating its protein–protein interaction with co-activators or co-repressors [6]. As a consequence, small-molecule ligands can positively or negatively regulate specific gene expression. The small-molecule-mediated control of nuclear hormone receptors is an excellent naturally occurring switch, which has made nuclear receptors attractive targets for drug development. Drug development targeting nuclear receptor transcription factors has made considerable contributions to human health and is likely to continue to do so in the future. However, the human genome includes only 48 nuclear receptors. To control members of the much larger set of human transcription factors, one needs to develop a more generic strategy for the modulation of activation domain function.

Protein–protein interactions of activation domains

Small-molecule modulation of the activation domains of transcription factors would be facilitated by a molecular understanding of the protein–protein interactions between their activation domains and co-activators. Activation domains have little sequence homology beyond a preponderance of particular amino acids (e.g. acidic residues, glutamine and proline) and usually only have limited folded structure in the absence of their target proteins [7,8]. The absence of conserved primary sequences and structural information for activators had made it difficult to design small-molecule modulators of transcription. To gather structural information, we have studied the activation domain of viral protein 16 (VP16), a prototypic acidic activator, as a model. A combination of NMR spectroscopy and biochemical techniques revealed that a short peptide segment of the VP16 activation domain folds into an α-helix upon binding to its target and that this short α-helical module is essential for the ability of the activation domain to stimulate transcription [9]. A systematic analysis of the activation domains of 65 human transcription factors showed that similar α-helical modules are present in the activation domains of disease-linked human transcription factors, including p53, nuclear factor-κB (NF-κB), p65, acute lymphoblastic leukemia-1 (ALL1), CCAAT/enhancer binding protein β (C/EBPβ), nuclear factor of activated T cells 1 (NFAT1), heat shock factor-1 (HSF-1) and epithelial-specific ETS transcription factor (ESX) [10,11]. These short motifs are essential for transcriptional activation by their corresponding transcription factors.

The amphipathic α-helical nature of activation domains had been proposed before our studies. For example, the successful design of synthetic activation peptides by Ptashne and co-workers [12] implied that acidic activation domains are α-helical, and the crystal structure of the helical p53 activation domain bound to mouse double minute 2 (MDM2), the cellular attenuator of p53, proposed by Pavletich and co-workers [13] also suggested that the p53 activation domain is α-helical even when bound to its co-activators. Our studies lend credence to these early suggestions.

The importance of α-helical motifs in transcriptional activation domains has now been described in several other human transcription factors. Well-known examples include the kinase-inducible activation domain of cAMP-response-element binding protein (CREB), which folds to form a pair of helices upon binding to the transcriptional co-activator CREB-binding protein (CBP) [14]; the activation domain of hypoxia-inducible factor-1α (HIF-1α), which forms several helices upon binding to CBP [15,16]; and the activation domain of hepatocyte nuclear factor-1α (HNF-1α), which forms a unique four-helix bundle upon binding to the transcriptional co-activator dimerizing cofactor for HNF-1 (DcoH) [17]. The conservation of the amphipathic α-helical structural motif in transcriptional activation is remarkable and raises the possibility of modulating gene transcription by small-molecule α-helix mimics.

Discovery of adamanolol and wrenchnolol

Of the α-helical motifs that we have characterized, there was particular interest in the activation domain of ESX, because of its strong activating potency and high-relevance to human cancer. ESX is an epithelial-specific transcription factor that activates human epidermal growth factor receptor-2 (Her2), an oncogene whose overexpression occurs in ~30% of breast cancer patients [18,19]. ESX binds and activates the Her2 promoter [20] and the ESX-binding site in the Her2 promoter is essential for the high-level expression of Her2 in breast cancer cells [21]. Our group biochemically isolated human suppressor of Ras (Sur-2), which is a nuclear protein that binds selectively to the α-helical motif of the ESX activation domain [22]. The 130 kDa nuclear protein is a Ras-linked subunit of the human mediator complexes known as: vitamin D receptor co-activator complex (DRIP); co-activator complex for Sp1 (CRSP); and activator-recruited cofactor (ARC) [23–26] (Figure 1a). The human mediator complexes are required for high levels of gene transcription in human cells, which could explain the powerful activation of Her2 expression by ESX.

The interaction of ESX with Sur-2 is mediated by one face of an eight-amino-acid α-helical region in the ESX activation domain, and the tryptophan residue in the hydrophobic face of the helix makes a unique contribution to the specificity of the interaction [22] (Figure 1b). Adamanolol, a small-molecule inhibitor of this interaction, has been identified by screening a chemical library enriched in indole-mimicking π-electron-rich pharmacophores (Figure 2) [27]. This unique molecule targets Sur-2, inhibiting the ability of the ESX activation domain to stimulate transcription in cells, blocks the interaction of ESX with Sur-2 *in vitro*, represses the expression of the *her2* gene in cultured cells and impairs the viability of Her2-positive breast-cancer cell lines.

Structure–activity relationship studies based around adamanolol led to the design of the second-generation compound named wrenchnolol [28] (Figure 3). The wrench-shaped molecule inhibited the ESX–Sur-2 interaction *in vitro* more potently than adamanolol (K_d of wrenchnolol with Sur-2 is 1.8 μM). This stable and water-soluble derivative of adamanolol is amenable to large-scale synthesis and preliminary studies suggest that wrenchnolol is effective in animal models of Her2-positive breast tumors (unpublished).

Other small molecules that block the protein–protein interactions of activation domains

In addition to adamanolol and wrenchnolol, several other molecules that modulate the protein–protein interaction of an activation domain have been described. Montminy and co-workers [29] reported the discovery of a small molecule that blocks the interaction between the KIX domain of CBP and the activation domain of CREB. An NMR-based screen of a focused small-molecule library identified several compounds that bind to different surfaces on

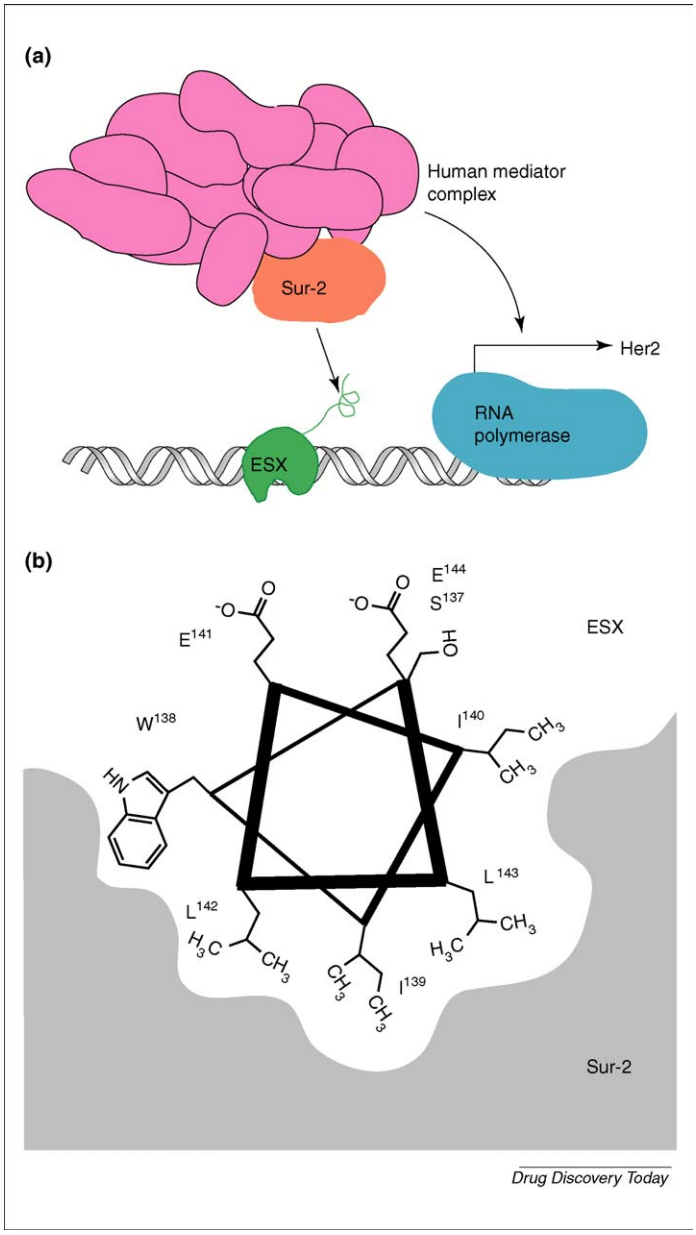

FIGURE 1

The epithelial-specific ETS transcription factor (ESX) activation domain. (a) The α-helical module in the ESX activation domain binds selectively to suppressor of Ras (Sur-2), a Ras-linked subunit of the human mediator complex, to activate the transcription of Her2. **(b)** A helical-wheel illustration of the activation domain of ESX bound to Sur-2.

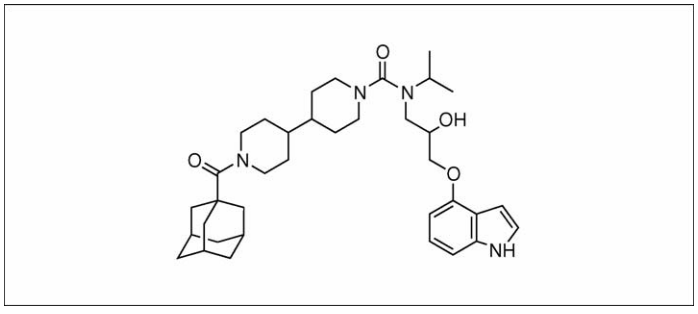

FIGURE 2

Chemical structure of adamanolol.

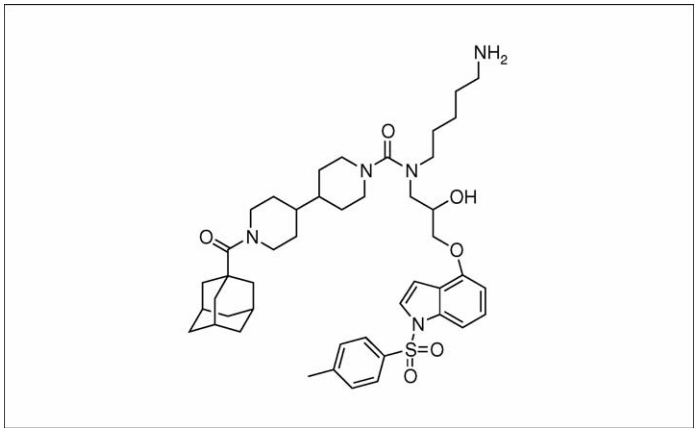

FIGURE 3

Chemical structure of wrenchnolol.

KIX. One of these, KG-501 (Figure 4), seemed to target a KIX surface distal to the CREB-binding site but still disrupted the CREB–CBP interaction and attenuated target-gene induction produced by cAMP in cultured cells.

As found in the examples of ESX and CREB, activation domains usually bind to co-activators for gene activation. However, activation domains also interact occasionally with negative cofactors or co-repressors. In such a case, specific inhibition of co-repressor interaction would restore the ability of the transcription factor to activate transcription. One prominent example is the small-molecule-induced restoration of p53, a transcription factor that is most frequently inactivated in human cancers [30]. In addition to inactivating mutations of the p53 DNA-binding domain, p53 is inactivated through the binding of its activation domain to MDM2, the cellular attenuator of p53 that is often overexpressed in human cancers. A team at Roche discovered the small-molecule antagonists of MDM2 (called nutlins) from screening a chemical library [31] (Figure 5). The *cis*-imidazoline compounds inhibited the binding of p53 to MDM2 (with IC_{50} values of 100–300 nM) and activated the p53 pathway in cancer cells, leading to cell-cycle arrest and apoptosis. Remarkably, oral administration of nutlin-3 reduced tumor growth in mouse xenografts.

It has generally been thought that protein–protein interactions are particularly difficult to target by small organic molecules because binding typically occurs over a relatively large surface area and because the binding surfaces tend to be flat and often lack pockets that might provide binding sites suited for small organic molecules. Nevertheless, the case studies of wrenchnolol, KG-501, nutlin and others highlight the promise of targeting

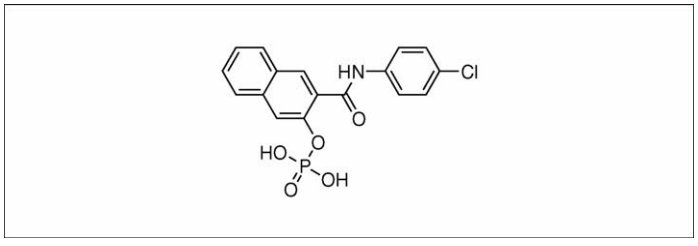

FIGURE 4

Chemical structure of KG-501 (naphthol-AS-E-phosphate).

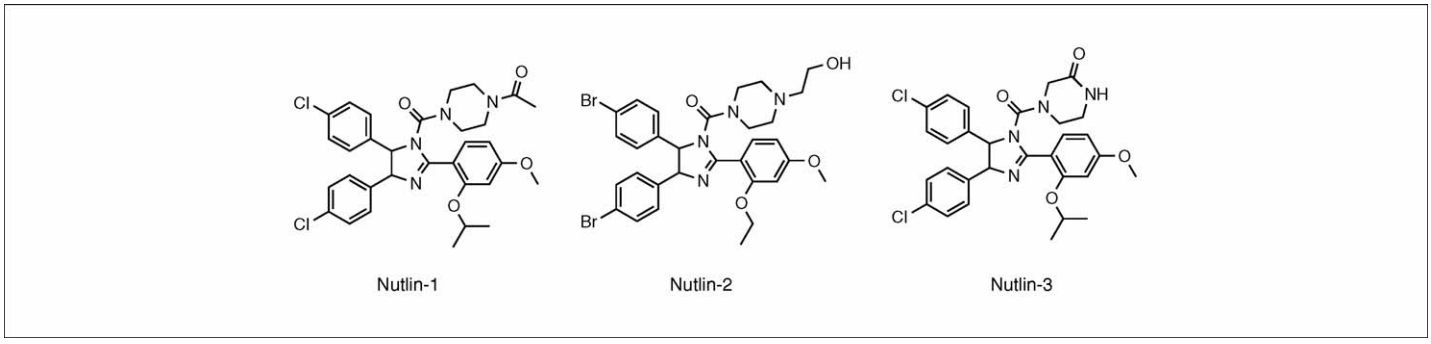

FIGURE 5

Chemical structure of nutlins. Nutlins were identified from a chemical library as inhibitors of p53–mouse double minute 2 (MDM2) binding. One of them, nutlin-3, suppressed the growth of tumor xenografts in nude mice.

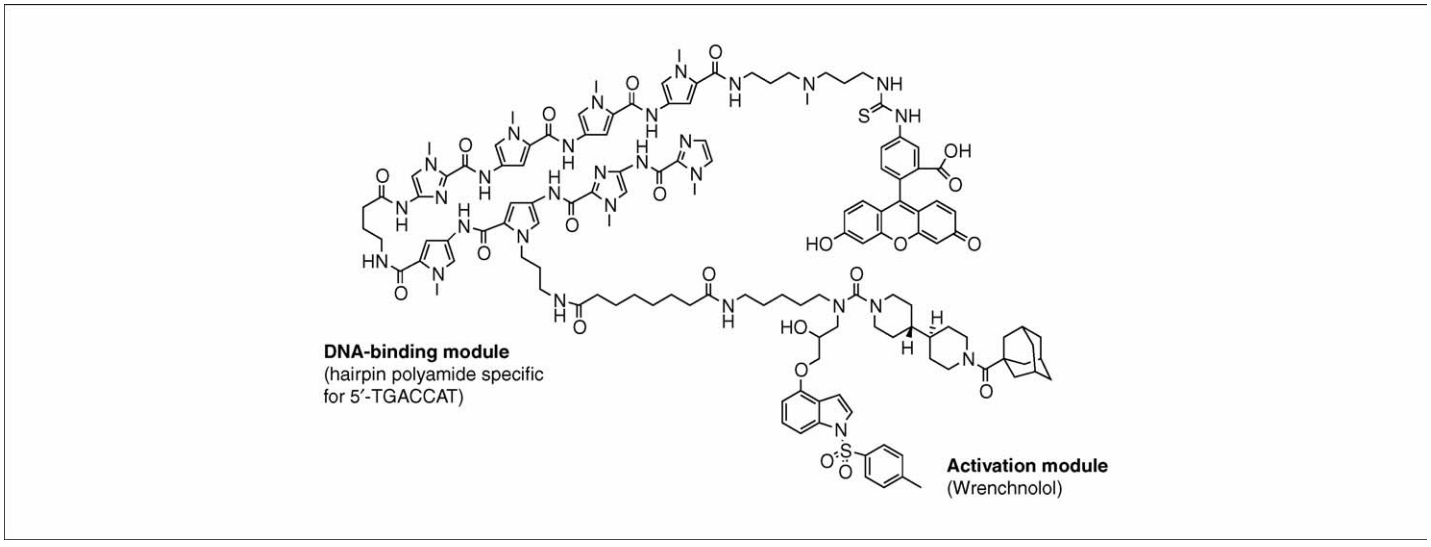

FIGURE 6

Chemical structure of synthetic transcription factor 1 (STF1). The hybrid molecule is composed of two modules: a hairpin-polyamide that recognizes and binds the TGACCAT sequence in DNA and wrenchnolol that binds to suppressor of Ras (Sur-2) co-activator.

activation-domain co-activator interactions by small organic molecules.

Small-molecule transcription factor

Structural analysis of wrenchnolol indicated that the wrench-shaped molecule mimics the α-helical interface of the ESX activation domain to block the ESX–Sur-2 interaction. We, thus, envisioned that the molecule could function as a transcriptional activation domain if it were covalently attached to a small-molecule DNA-binding domain such as a hairpin polyamide.

In collaboration with Dervan's group, our group developed the entirely organic transcription factor synthetic transcription factor 1 (STF1) by coupling the hairpin polyamide and wrenchnolol [32] (Figure 6). As expected, STF1 activated transcription of a reporter containing hairpin-polyamide-binding sites *in vitro*, whereas the control molecule lacking the wrenchnolol moiety had no detectable activity. A reporter construct with point mutations in the hairpin-polyamide-binding sites was unresponsive to STF1, suggesting a high-degree of selectivity of STF1. Biochemical investigation showed that STF1 stimulates transcription by recruiting the human mediator complex to the promoter through simultaneous contacts with Sur-2 and DNA, just as a naturally occurring transcription factor does. These results directly demonstrate that it is possible to generate a transcription factor out of completely organic components.

Notwithstanding the success of STF1 *in vitro*, STF1 had limited ability to penetrate the cell membrane, even though the hairpin polyamide and the wrenchnolol molecule are completely cell

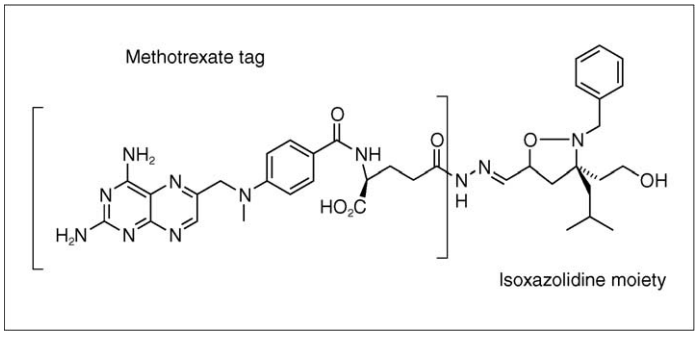

FIGURE 7

Chemical structure of an isoxazolidine activator. The isoxazolidine molecule (shown here) activated transcription *in vitro* in the presence of a dihydrofolate-reductase (DHFR)–LexA fusion protein when conjugated with methotrexate.

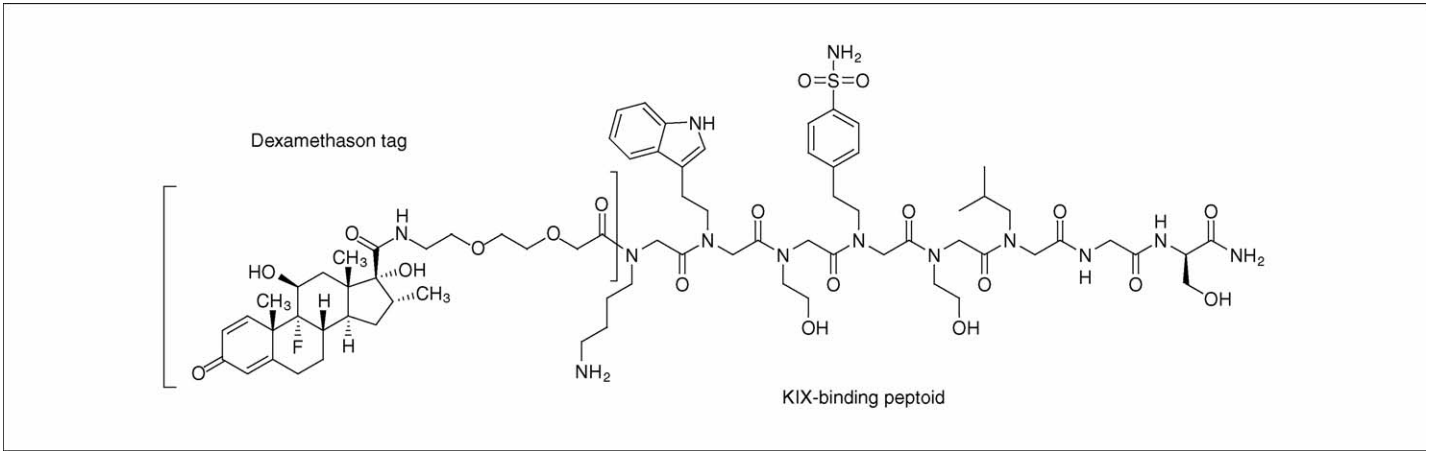

FIGURE 8

Chemical structure of a KIX-binding peptoid activator. The dexamethason conjugate of the peptoid showed potent gene activation in cultured mammalian cells in the presence of a Gal4–glucocorticoid receptor fusion protein.

permeable as separate compounds. A biotinylated version of wrenchnolol appears to be cell permeable and is capable of activating transcription in cultured mammalian cells when an artificial streptavidin–Gal4 fusion protein is expressed in cells to localize the wrenchnolol molecule to DNA, indicating that wrenchnolol functions as an activation module in cells (unpublished). Reduction of the size of the molecule and optimization of physical properties could increase the cell permeability of wrenchnolol-derived transcription factors.

Other small-molecule activation domains

Discovery of several other small-molecule activation domains has been described in recent literature. Mapp and co-workers [33] discovered small-molecule activation domains containing isoxazolidine structures. An isoxazolidine ring was selected as a scaffold because its conformationally constrained ring facilitates the stereo-controlled introduction of diverse functional groups. Each isoxazolidine derivative was coupled to methotrexate, a tag molecule that facilitated localization of the isoxazolidines through the interaction with artificial dihydrofolate-reductase–LexA (DHFR–LexA) fusion protein (Figure 7). Remarkably, two of the five molecules they synthesized exhibited activity in an *in vitro* transcription assay, although their targets remain unknown.

By contrast, Kodadek and co-workers [34] used a large combinatorial library of ~100,000 peptoids to discover small-molecule activation domains. Two peptoids were found to bind to the co-activator CBP, a target of several naturally occurring activation domains. When conjugated with dexamethasone, a high affinity ligand of the glucocorticoid receptor used to localize the peptoids to DNA, one of the two peptoids exhibited a dose-dependent and highly potent activation of a reporter gene in cultured cells (Figure 8).

Although these molecules have not been tested as a component of an entirely organic small-molecule transcription factor, their distinct chemical properties and high potencies certainly serve as a basis for designing a potent, compact activation module for small-molecule transcription factors.

None of the small molecules described in this review are yet as good as Merumo's magical candies. However, the demonstration of the completely organic small-molecule transcription factor suggests that it might eventually be possible to design a more drug-like transcription factor with a clear biological effect. Regulation of gene expression by transcription factors touches all aspects of human biology: development, aging, cell death, oncogenesis, metabolism, and many others. It is an exciting prospect that drug-like small-molecule transcription factors could precisely control all of these biological processes.

Acknowledgements

We are grateful to several collaborators, colleagues, funding agencies and pharmaceutical companies for their support and their kind understanding of our vision. We would also like to thank David D. Moore for his critical review of the manuscript and Shinji Kamisuki for assistance in figure preparation.

References

1 Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* 9, 2215–2235

2 Gottesfeld, J.M. *et al.* (1997) Regulation of gene expression by small molecules. *Nature* 387, 202–205

3 Dickinson, L.A. *et al.* (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl. Acad. Sci. U. S. A.* 95, 12890–12895

4 Olenyuk, B.Z. *et al.* (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16768–16773

5 Mangelsdorf, D.J. *et al.* (1995) The nuclear receptor superfamily: the second decade. *Cell* 83, 835–839

6 Feng, W. *et al.* (1998) Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. *Science* 280, 1747–1749

7 Triezenberg, S.J. (1995) Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* 5, 190–196

8 Tjian, R. and Maniatis, T. (1994) Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 77, 5–8

9 Uesugi, M. *et al.* (1997) Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science* 277, 1310–1313

10 Uesugi, M. and Verdine, G.L. (1999) The alpha-helical FXXPhiPhi motif in p53: TAF interaction and discrimination by MDM2. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14801–14806
11 Choi, Y. *et al.* (2000) Divergent hTAFII31-binding motifs hidden in activation domains. *J. Biol. Chem.* 275, 15912–15916
12 Giniger, E. and Ptashne, M. (1987) Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. *Nature* 330, 670–672
13 Kussie, P.H. *et al.* (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* 274, 948–953
14 Radhakrishnan, I. *et al.* (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell* 91, 741–752
15 Freedman, S.J. *et al.* (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5367–5372
16 Dames, S.A. *et al.* (2002) Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5271–5276
17 Rose, R.B. *et al.* (2000) Structural basis of dimerization, coactivator recognition and MODY3 mutations in HNF-1alpha. *Nat. Struct. Biol.* 7, 744–748
18 Slamon, D.J. *et al.* (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235, 177–182
19 Slamon, D.J. *et al.* (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244, 707–712
20 Chang, C.H. *et al.* (1997) ESX: a structurally unique Ets overexpressed early during human breast tumorigenesis. *Oncogene* 14, 1617–1622
21 Scott, G.K. *et al.* (1994) Binding of an ETS-related protein within the DNase I hypersensitive site of the HER2/neu promoter in human breast cancer cells. *J. Biol. Chem.* 269, 19848–19858
22 Asada, S. *et al.* (2002) External control of Her2 expression and cancer cell growth by targeting a Ras-linked coactivator. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12747–12752
23 Naar, A.M. *et al.* (1999) Composite co-activator ARC mediates chromatin-directed transcriptional activation. *Nature* 398, 828–832
24 Rachez, C. *et al.* (1999) Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature* 398, 824–828
25 Boyer, T.G. *et al.* (1999) Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature* 399, 276–279
26 Ryu, S. *et al.* (1999) The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* 397, 446–450
27 Asada, S. *et al.* (2003) A gene-expression inhibitor that targets an alpha-helix-mediated protein interaction. *J. Am. Chem. Soc.* 125, 4992–4993
28 Shimogawa, H. *et al.* (2004) A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction. *J. Am. Chem. Soc.* 126, 3461–3471
29 Best, J.L. *et al.* (2004) Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17622–17627
30 Hollstein, M. *et al.* (1991) p53 mutations in human cancers. *Science* 253, 49–53
31 Vassilev, L.T. *et al.* (2004) *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* 303, 844–848
32 Kwon, Y. *et al.* (2004) Small molecule transcription factor mimic. *J. Am. Chem. Soc.* 126, 15940–15941
33 Minter, A.R. *et al.* (2004) A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* 126, 10504–10505
34 Liu, B. *et al.* (2005) A potent transactivation domain mimic with activity in living cells. *J. Am. Chem. Soc.* 127, 8254–8255