Generic binding sites, generic DNA-binding domains: where does specific promoter recognition come from?

Adrien B. Georges,*,†,‡ Bérénice A. Benayoun,*,† Sandrine Caburet,*,† and Reiner A. Veitia*,†,1

*Unité Mixte de Recherche 7592-Centre National de la Recherche Scientifique, Institut Jacques Monod, Paris, France; †Université Paris Diderot/Paris 7, Paris, France; and ‡Ecole Normale Supérieure de Paris, Paris, France

ABSTRACT **A transcription reaction relies on the specific recognition of *cis*-regulatory regions containing short DNA motifs. Such sequences are bound by transcription factors (TFs) involved in the recruitment, direct or not, of the transcriptional machinery. A eukaryotic genome can contain tens of genes encoding TFs that recognize very similar consensus DNA target sequences. In this review, we explore in a simple way how TFs coexpressed in the same cells and recognizing generic consensus sites with generic DNA-binding domains can achieve a specific modulation of target gene expression. We dissect the strategy followed by eukaryotes, which involves the formation of complex nucleoprotein structures involving many TFs and their cognate binding sites. This multiplicity of actors increases the effective length of the target DNA recognized by the TFs and might help paralogous TFs establish specific interactions. From this perspective, eukaryotic gene regulation implies the cooperation of several TFs, which is also the basis of information integration. Such cooperative TFs are likely to form a combinatorial partner code whose ultimate molecular hallmark is the assembly of enhanceosome-like structures ensuring the formation of an activation surface that is complementary to other coactivators and to the transcriptional machinery itself.—Georges, A. B., Benayoun, B. A., Caburet, S., Veitia, R. A. Generic binding sites, generic DNA-binding domains: where does specific promoter recognition come from? *FASEB J.* 24, 346–356 (2010). www.fasebj.org**

Key Words: cooperativity · promoter · specificity · synergy · transcription

Gene expression and its regulation are dependent, to a great extent, on transcription. An elementary transcription reaction is initiated by the recognition of *cis*-regulatory regions and especially promoters. These regions contain a series of short DNA sequence motifs that are bound by general and/or specific transcription factors (TFs) (1). To achieve a well-modulated gene expression (*i.e.*, the right gene expressed at the right place, at the right moment, at the right level), it is mandatory to achieve a rather high degree of specificity in the recognition events that will allow the recruitment, direct or not, of the transcriptional machinery (2). A eukaryotic genome typically contains hundreds to thousands of genes encoding transcription factors, often grouped in families according to the structure of their DNA-binding domains. For instance, more than 80% of TFs in higher eukaryotes contain a helix-turn-helix domain, a basic helix-loop-helix, a zinc finger domain, or a leucine zipper. Helix-loop-helix and leucine zipper domains allow dimerization of factors that contain an N-terminal α helix that ensures interactions with DNA (1). This is linked to the fact that eukaryotic genome evolution has involved a series of whole-genome/gene duplication events, which has favored the emergence of large families of paralogous genes encoding TFs sharing structural (and functional) properties. As a consequence, it is common that many TFs within a family recognize either the same, or very similar, consensus DNA target sequences. Interestingly, whereas prokaryotic TFs recognize quite long target sequences (as long as 25 bp in *Escherichia coli*) in eukaryotes, TFs recognize shorter sequences (for instance, the average TF binding site (TFBS) in *Drosophila melanogaster* is ~13 bp long) (3). The emergence of these TFs binding short target sequences, coupled with a need for tight gene expression in multicellular organisms, demands the existence of mechanisms generating specificity, especially in the case of related TFs. For instance, the many members of the Forkhead box (Fox) TF family, share a highly conserved DNA-binding domain, known as the "forkhead box" or "winged helix" domain and a well-conserved target consensus sequence (refs. 4, 5 and references therein). High-affinity binding sites for 14 Forkhead proteins have been identified (**Table 1**), and their consensus sequence is 5′-(G/A)(T/C)(A/C)AA(C/T)A-3′ (ref. 5 and references therein). Fox factors are involved in processes as diverse as eye organogenesis [FoxC1–2 (6)], language acquisition [FoxP2 (6)], stress response, aging regulation, and tumor suppression [FoxO (7)], and liver development [FoxA1–2 (8)]. Thus, the high conservation of the Forkhead domains and of their binding sites on the one hand and the diversity of the

[1] Correspondence: Equipe Génétique et Génomique du Développement Gonadique, Institut Jacques Monod, Bâtiment Buffon, 15 rue Hélène Brion, 75205 Paris cedex 13, France. E-mail: reiner.veitia@inserm.fr

doi: 10.1096/fj.09-142117

0892-6638/10/0024-0346 © FASEB

TABLE 1. *Described high-affinity binding sites of various Forkhead transcription factors*

FOX factor name	Name in original description	High-affinity binding site described								Original description (ref.)
FOXF2	FREAC-2	G/A	T	A	A	A	C/T	A	A	54
FOXC1	FREAC-3	G	T	A	A	A	C/T	A	A	54
FOXD1	FREAC-4	G/A	T	A/C	A	A	C	A	N	54
FOXL1	FREAC-7	G/A	T/C	A/C	A	A	C/T	A	N	54
FOXQ1	HFH-1	A/C	T	A	A	A	C	A	A/T	55
FOXD3	HFH-2	A/T	T	A	A	A	C	A	A/T	55
FOXA3	HNF-3	G/A	T/C	A/C	A	A	C/T	A	A/T	55
FOXI1	N/A	G/A	C	C	A	A	T	C/G	A	56
FOXK2	ILF-1	G	T	A	A	A	C	A	A	57
FOXO1	FKHR	G	T	A	A	A	C	A	A	58
FOXO3A	FKHRL1	G	T	A	A	A	C	A	A	58
FOXO4	AFX	G	T	A	A	A	C	A	A	58
FOXP1	N/A	A	T/C	A	A	A	C	A	A	35
FOXL2	N/A	G	T	C/G	A	A	G	G	T	5
Forkhead	General consensus	G/A	T/C	A/C	A	A	C/T	A	N	

N/A, not available.

biological processes the FOX genes are involved in on the other hand imply that a high degree of specificity must exist to ensure that different family members can successfully play their different (paralog-specific) roles. Another well-known family of TFs with close DNA-binding specificity involves the Homeobox (Hox) factors. Indeed, the *in vitro* DNA-binding specificity of Hox TFs is poor, because they recognize highly similar sequences centered around a 5′-TAAT-3′ core (9) (**Table 2**). However, *in vivo*, Hox proteins demonstrate a high degree of target specificity, and ectopic expression of any Hox protein has severe consequences on the developmental program (9).

To complicate matters, similar high-affinity TFBSs may also be shared by TFs coming from families with seemingly independent evolutionary origins. This is the case with the Forkhead TF FOXL2, the nuclear estrogen receptors (ERs) ERα and ERβ, and other nuclear receptors, including steroidogenic factor SF1 and NR4A1 (5, 10–15). Moreover, recall that any TF and its competitors, from the same family or not, have to locate *cis*-regulatory regions in a sea of DNA sequence. If we consider a 6-mer binding site and assume a random sequence composition, it will be represented >700,000 times in the human genome just by chance. However, this would be an underestimation because a TF can recognize degenerate sequences. Obviously, only a small subset of such sites is occupied, and chromatin modification may play an important role, because it can partition the genome into open/expressed and closed/nonexpressed compartments, which reduces the chromatin space to be explored by any TF (16).

How can TFs recognizing similar binding sites achieve a specific modulation of target gene expression? Several nonexclusive mechanisms can help in providing specificity: *1*) differences in expression patterns (as demonstrated for Fox and Hox TFs), *2*) limited, but existing, specificity in TFBS recognition, and *3*) existence of different cofactors (either because they are tissue-specific, because the relevant TFs contain different protein/protein interaction domains, or because distinct signals will modify their molecular behavior, through post-translational modifications). Here, we further explore these issues. However, we do not attempt to provide a comprehensive review on the mechanisms of promoter recognition and transcription. We focus on the particular question of how specific sequence recognition and hence specific gene expression can be achieved by paralogous TFs coexpressed in the same cells. We

TABLE 2. *Described high-affinity binding sites of various Homeobox transcription factors*

Homeobox factor name	Homeobox factor species	High-affinity binding site described								Original description (ref.)
vnd/NK-2	*Drosophila melanogaster*	T	T/C	A	A	G	T	G	G/C	59
Engrailed	*Drosophila melanogaster*	G	T	A	A	T	G	A	C	60
OTX5	*Drosophila rerio*	N	T	A	A	G	A	C	T	61
Dlx3	*Xenopus tropicalis*	A	T	A	A	T	T	G/A	C/G	62
Xom	*Xenopus tropicalis*	C	T	A	A	T	T	A/G	G/C	63
CdxA (CTO)	*Gallus gallus*	A	T	A	A	A	T/G	N	N	64
CdxA (CTS)	*Gallus gallus*	A	T	T/A	A	T/A	T	N	N	64
CDX3	*Gallus gallus*	C	T	A	A	T	T	N	N	64
CAD	*Gallus gallus*	A	T	A	A	A	N	N	N	64
HOXA13	*Homo sapiens*	A	T	A	A	A	C/G	C/G	N	65
TGIF	*Homo sapiens*	T	C	A	A/T	T/A	A/T	C	N	66
SHOX	*Homo sapiens*	N	T	A	A	T	G	N	N	67
Homeobox	General consensus	A	T	A	A	T	T	N	N	

illustrate our points with theoretical and experimental examples of increasing complexity.

RECOGNITION OF TARGET DNA BY HOMODIMERS: TWO IS BETTER THAN ONE

Recognition of DNA by a preassembled homodimer

The ability of TFs to form homodimers is a fairly common property (1). This is the basis of the existence of a first layer of recognition specificity, even when a TF is in competition with other proteins for similar sequence elements on DNA. How can the formation of homodimers increase recognition specificity compared with that of a monomer? This strategy basically implies an increase in the total effective length of the recognized sequence. Effective length should be understood in terms of informational content to foster specific interactions (which is, of course, linked to physical length). The informational content of TFBSs is low because of their short lengths, which leads to spurious occurrences in the genome. Thus, use of two (or more) close binding sites increases the reliability of a recognition event.

To understand how this strategy works, let us focus on the simple case of a preassembled homodimer (**Fig. 1*A***). The dimer TF-TF will recognize its target with an affinity constant K_{dimer}. As a first approximation, the energy of the interaction between the dimer and DNA involves the additive free energies of the interactions of each monomeric TF with their target sites. Basic physicochemical rules indicate that K_{dimer} is on the order of $(K_{monomer})^2$. This huge difference between K_{dimer} and $K_{monomer}$ is the elementary basis of discrimination between a random monomeric DNA sequence resembling a TFBS (genomic noise), recognized with an affinity K, and a true bipartite TFBS (signal) bound with a much higher affinity, $\sim K^2$. A particularly striking, naturally occurring, example of this is provided by the atypical E2F proteins. Several typical E2F proteins involved in the regulation of the cell cycle, apoptosis, and differentiation have been described. They possess an N-terminal DNA-binding domain followed by a dimerization domain, which allows interactions with a dimerization partner (DP) protein (that brings a similar DNA-binding domain). Dimerization is required for high-affinity, sequence-specific DNA binding. Interestingly, atypical E2F members have a duplicated DNA-binding domain. For the reasons outlined above, the presence of two DNA-binding domains allows DNA recognition in a DP-independent fashion and DNA binding is abolished by a mutation in either of the two DNA-binding domains (17). Thus, the ability of TFs in competition for the target sequences to form dimers, or not, can make a difference in the strength of protein-DNA interactions. For instance, if TF1 forms dimers and TF2 does not, dimers of TF1 will be active even at very low concentrations, and TF2 would not be an important source of interference. However, all of these arguments are valid for any homodimeric TF and cannot explain by themselves how different homodimers will lead to specific target regulation. The simplest ingredient ensuring target discrimination is the fine-tuning of the spacing between similar monomeric DNA-binding sites, which should be mirrored by the quaternary structure of the relevant dimers.

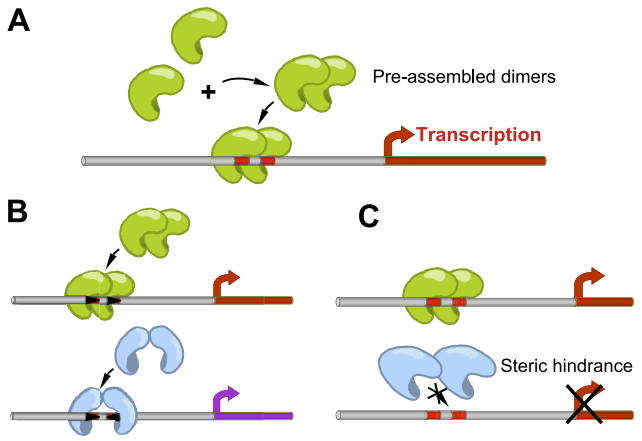

Figure 1. Promoter specificity in dimeric TFs. Effects of binding site orientation and spacing. *A*) Binding of a preassembled TF dimer (green proteins) on the red binding sites in the promoter is required to induce a transcriptional response of the corresponding gene (red arrow). *B*) Promoter specificity provided by TFBS orientation. Tandem orientation of two identical binding sites in one promoter (red gene) will enable binding of one type of TF, here the green TF1-TF1 dimer, but not of the symmetric blue TFs. The palindromic orientation of the binding site in the promoter of another gene (purple) will allow binding of the blue dimer. *C*) Promoter specificity provided by TF binding-site spacing. The proximity of two identical binding sites will enable the binding of one type of TF, here the green TF1-TF1 dimer, but not of the too voluminous blue TF2-TF2 dimer recognizing the same half site (*i.e.*, belonging to the same TF family), because of steric hindrance.

There are two main ways of forming homodimers. Most frequently, each monomer has a domain that interacts with the same domain of the other monomer (*e.g.*, a leucine zipper). In these cases, the resulting dimers are symmetric, and the ideal arrangement of the two binding sites should then be a palindrome, a geometry that also displays a central symmetry (Fig. 1*B*). This is the case, for example, of the ER and its bipartite TFBS (18, 19). The other way occurs when one domain of one monomer recognizes a different domain of the other monomer. In such cases some allosteric change should occur to prevent further association of new monomers and stabilize the dimeric form (avoiding polymerization). A TF complex formed according to such a pattern would preferably recognize half-binding sites arranged in a head-to-tail configuration, as shown in Fig. 1*B*. Examples of such a configuration have been shown to occur *in vivo* and can be found in the literature (15, 20). As displayed in Fig. 1*C*, the structure of dimers TF1-TF1 (*i.e.*, green TF) and TF2-TF2 (*i.e.*, blue TF) can be such that they cannot cross-recognize their bipartite binding sites. This strategy is easily conceivable because even though TF1 and TF2 share the same DNA-binding domain, the rest of these proteins can be (and usually are) very divergent. As a result, the dimers TF1-TF1 and TF2-TF2 can have

different quaternary structures. Thus, differences in the distance between the two DNA-binding domains in the dimer correlating with different spacings between the TFBSs in a promoter are a key parameter of DNA-binding specificity. This has been shown in the case of the thyroid hormone, vitamin D_3, and retinoic acid response elements, which share similar half-sites (15). Indeed, the spacing between the half-sites is instrumental for the generation of specific responses: a thyroid hormone response element can be converted into a retinoic acid response element or a vitamin D_3 response element by varying half-site spacing (15).

Cooperativity in DNA recognition, specificity, and promoter activity

In some cases, homodimers are not preassembled but are formed during the recognition event. This implies that monomers are virtually unable to interact in solution owing to their low concentration and to their relatively low dimerization constant. Homodimer formation directly onto the promoter is linked to the idea of cooperativity, which postulates that the binding of one monomer can facilitate binding of a second one to a nearby TFBS (**Fig. 2**) (21). Where does cooperativity come from? One possibility is that on DNA recognition, a newcomer TF molecule is attracted synergistically by both the bound TF and the neighboring free TFBS (*i.e.*, DNA would make possible a protein-protein interaction not favored in solution). Another explanation is that the interaction of the first monomer with DNA induces an allosteric change in the protein, which in turn increases its affinity for an incoming one. In any case, the interaction between two monomers is enhanced and stabilized by DNA. By favoring efficient binding of the right dimer (matching a particular promoter geometry), even at low monomer concentration, cooperativity is a key component of binding specificity. It is also possible that, in the absence of direct protein-protein interactions, binding of one monomer to DNA may lead to a conformational change in the neighboring binding site, thereby increasing its affinity for a newcomer monomer. Such a mechanism does not imply

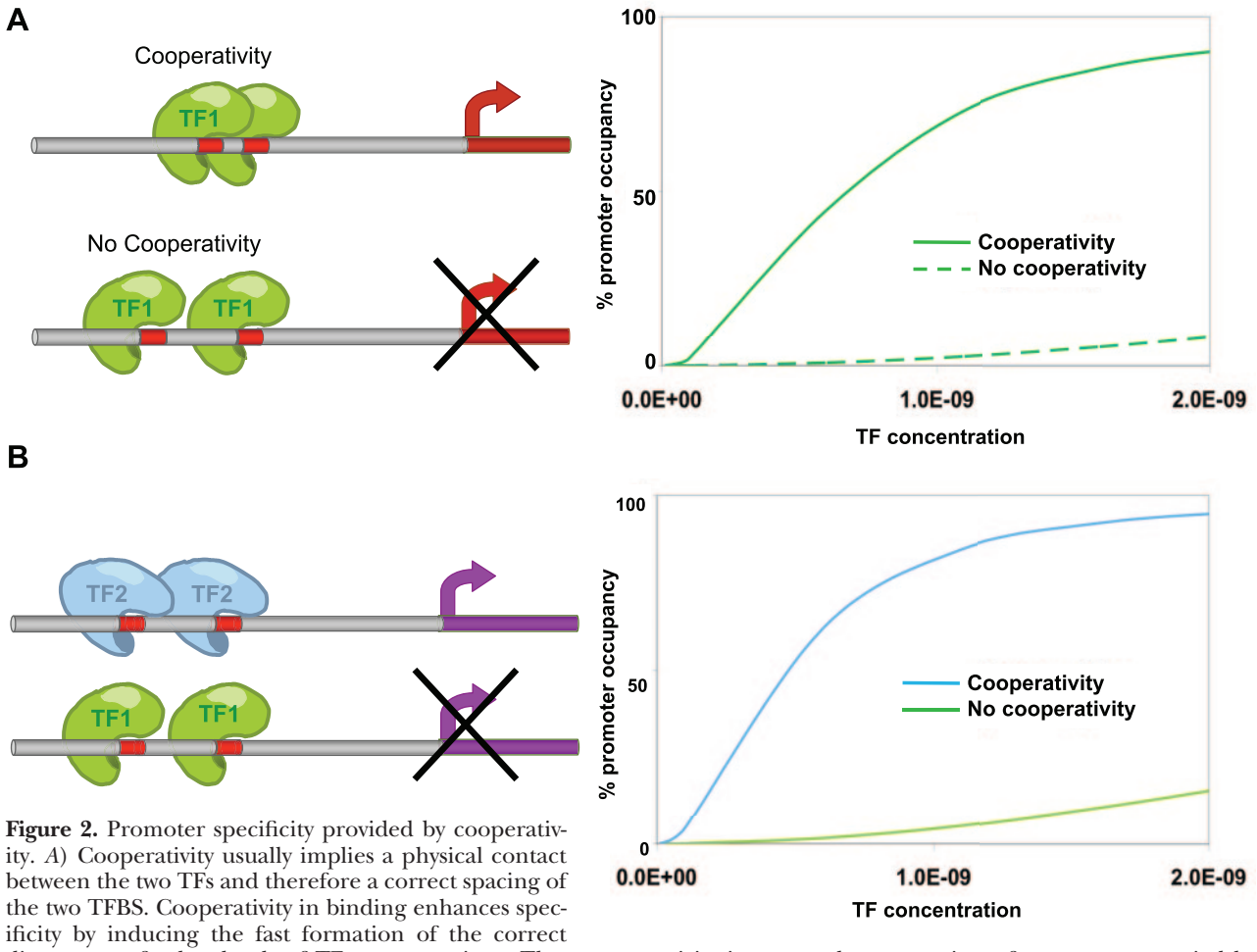

Figure 2. Promoter specificity provided by cooperativity. *A*) Cooperativity usually implies a physical contact between the two TFs and therefore a correct spacing of the two TFBS. Cooperativity in binding enhances specificity by inducing the fast formation of the correct dimer, even for low levels of TF concentrations. Thus, cooperativity increases the proportion of promoter occupied by functional dimers as a function of the monomeric TF concentration. This mechanism leads to a sharp transcriptional response for a slight increase of monomeric TF concentration (see absence of cooperativity for comparison). *B*) Another spacing of the same TFBS in another promoter (purple) allows cooperative binding of TF2 monomers, while imposing a noncooperative binding to TF1. The sigmoidal curves were derived using Eq. 1 from Veitia and Nijhout (33). Quantitative details are available from the authors.

formation of a dimer and is apparently less likely to account for discrimination between two TFs (*i.e.*, TF1 and TF2). However, if binding of TF1 and TF2 to DNA does not lead to the same conformational change, the distance and orientation of the TFBSs (and of the bound proteins) may be such that some kind of long-range discrimination can emerge. This will become clearer if we include another ingredient—transcriptional synergy—that we discuss below.

As shown by the graphs in Fig. 2, cooperativity leads to higher-than-additive responses so that a small change in activator concentration may elicit dramatic changes in binding and most likely in promoter activity (and gene expression). Such systems generating an S-shaped (sigmoidal) response can be used as molecular switches, because TF concentrations just below the threshold are unable to substantially stimulate transcription, whereas concentrations above the threshold can induce strong responses.

Affinity, availability, and specificity: closer to reality

Up to now, we have considered that all members of a family have similar affinities for the available TFBSs. The known consensus site for a TF is often derived from *in vitro* experiments and is an average sequence representing many different sites that are recognized with strong, yet different, affinities. For instance, although the members of the Fox family are able to bind to the general forkhead consensus sequence, it does not mean that they all have the same absolute affinity for it and that variants of the site are not recognized with even higher affinities. Consider, for instance, the case of FOXL2, whose high-affinity binding site (FLRE) diverges from the general binding consensus. However, it is also known to transactivate through a TFBS closer to the general consensus (mFLRE in ref. 5). Moreover, it has been shown that FOXL2 can transactivate *in vivo* using binding sites close to the general forkhead consensus present in the target GnRHR and FST promoters (22, 23).

Intrinsic affinity variations may elicit different responses after cell- or signal-dependent changes in the relative concentrations of the members of the TF family. Indeed, the relative availability of the TFs in potential competition for the same binding sites is a very important parameter. To illustrate this point, let us consider two dimeric TFs (TF1-TF1 and TF2-TF2) using the same sites in a target promoter and binding DNA with global affinities K_1 and K_2, respectively, being $K_1 < K_2$. Because binding efficiency is a function of both affinity and TF concentration, variations of TF1 and TF2 concentrations can switch promoter preference from TF2 to TF1 (*i.e.*, after a strong increase in TF1 concentration in response to a signal, for instance). Moreover, a TF can activate expression of a specific gene in a given cell lineage, but not in another, because of variations of its availability compared with that of its competitors. In a cell lineage in which TF2 is poorly expressed, TF1 will activate a target promoter specifically, which will not be the case in another cell lineage in which TF2 is highly expressed. Although neither TF1 nor TF2 is, in this case, specifically required to activate this target, experiments will show that gene expression is TF1-dependent in one lineage and TF2-dependent in the other. An example is provided by the FOXO factors (FOXO1, 3a, and 4), as many of their targets have been shown to be regulated by any of them, through a common regulatory sequence (the DBE) (24). Although FOXO factors are fairly ubiquitous, according to UniGene EST profiles (Hs.370666, Hs.220950, and Hs.584654) their abundances differ among tissues. For instance, FOXO3a is the only one expressed in adipose tissue, FOXO4 prevails in placenta, and FOXO1 dominates in lymph nodes, which suggests that FOXO target regulation could be driven by one member or another according to the tissue of interest. This finding is consistent with the differences in phenotypes of the three separate knockout-mouse models, which show that Foxo TFs, although often working through regulation of identical targets, have a degree of specificity (owing, at least in part, to preferential tissue expression) (25–27). In the case of Hox proteins, Goff and Gabin (28) proposed that they might modulate the expression of the same targets involved in growth. However, some Hox proteins would be more effective on some promoters than others, depending on their relative concentrations. The overall rate of growth in a given embryonic region would be the result of a competition of the Hox genes expressed in that region for the same targets (28).

TRANSCRIPTIONAL SYNERGY: AN UNDERESTIMATED INGREDIENT OF SPECIFIC *CIS*-REGULATORY SEQUENCE RECOGNITION

The previous discussion was based on the idea that two occupied sites will promote a much stronger gene transactivation than a single site. This idea is supported by a model in which transcription activation is exponentially dependent on the number of TF molecules bound to the promoter, which are able to attract in a concerted fashion the transcriptional machinery (or an intermediary mediator complex) (29–32). In other words, if we suppose that the transcriptional machinery binds a TF molecule in contact with DNA with an affinity K, then its interaction with a complex involving two DNA-bound TFs is supposed to imply an affinity K^2, and so on. This is the molecular explanation of transcriptional synergy (*i.e.*, homosynergy if only one type of TF is involved) and means, in practical terms, that the species contributing the most to transcription is the promoter with all its TFBSs occupied (33).

The existence of homosynergy allows, in principle, sequence-specific recognition even in the absence of classic cooperativity. Where would specificity come from in this case? As shown in **Fig. 3**, the transcriptional machinery (or the mediator complex (ref. 34 and references therein) might interact, through generic domains, with TF molecules bound to DNA on sites that have a characteristic spacing. Synergy will ensure discrimination between TF1 and TF2 molecules bound to

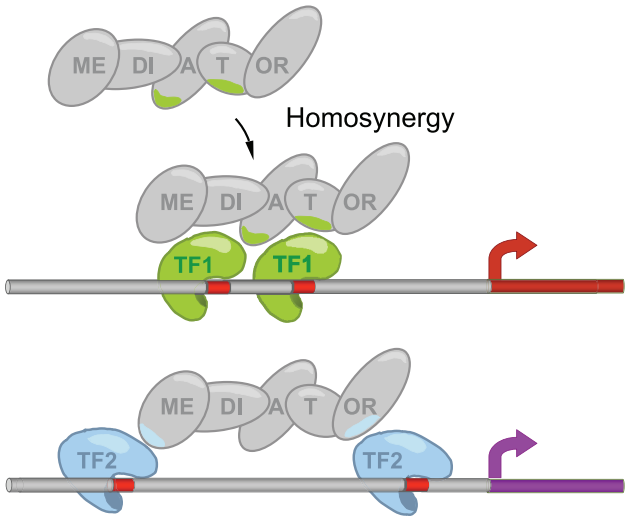

Figure 3. Promoter specificity provided by transcriptional synergy. Synergy stems from the concerted attraction of the transcriptional machinery, or the mediator complex, by TFs bound to DNA. In this simplified example, a multimeric complex (mediator? the RNA polymerase itself?) is supposed to bind differently to various couples of TFs, therefore ensuring specific activation of the transcription when the relevant TFs are present on the promoter at the right spacing. Colored regions in the mediator complex are supposed to be generic protein-protein interaction domains able to recognize activation domains of the TFs.

DNA irrespective of how the TF-DNA interaction took place (*i.e.*, cooperatively or not). Of course, the existence of cooperativity in the binding of the monomers would provide a further layer of specificity.

Thus far we have focused on only two DNA-binding sites, but this paradigm can be extended to *cis*-regulatory regions containing multiple TFBSs for the same TF. Indeed, it is known that increasing the number of sites available for binding dramatically increases the responsiveness of the promoter to even very low concentrations of the relevant TF (see refs. 31, 32 for mathematical models).

BIPARTITE INTERACTIONS: HETERODIMERIZATION AND HETEROSYNERGY

Heterodimers and specificity

Although TFs of the same family have a well-conserved DNA-binding domain, the remaining domains of these proteins are often highly divergent. For instance, forkhead TFs only share the DNA-binding domain (4), and no good sequence alignment is possible outside the forkhead, although subsets of FOX proteins can share other domains. Namely, FOXP subfamily members have zing finger domains and FOXK subfamily members have FHA domains (35, 36). Similar observations can be made for the Hox factors. Thus, different TFs can establish specific and unambiguous interactions with other TFs, through their more variable domains. The differential formation of heterodimers, in solution or upon DNA binding, with other TFs can turn a generalist TFBS into a specific one (**Fig. 4**). As in the case of cooperativity described previously, there is no need for very strong protein-protein interactions to ensure a spectacular gain of specificity. An example of this heterotypic cooperation is provided by the interaction between NFAT and FOXP3 in the control of regula-

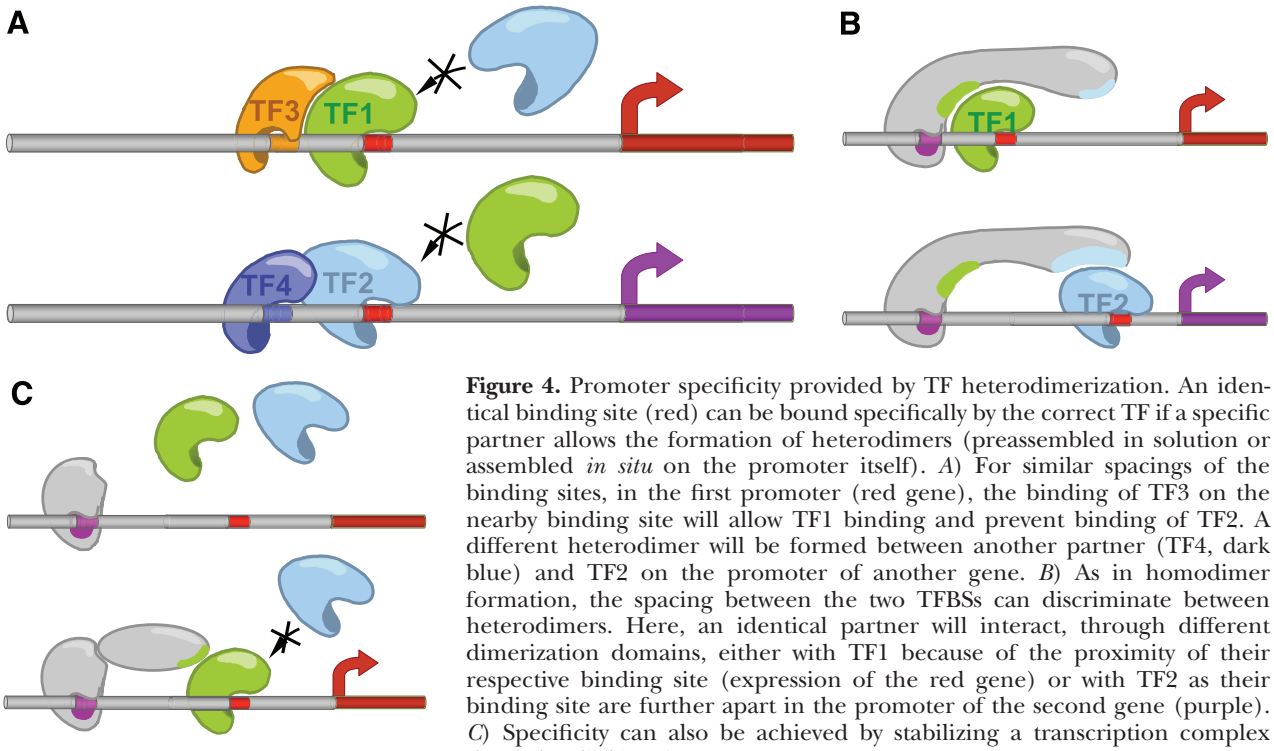

Figure 4. Promoter specificity provided by TF heterodimerization. An identical binding site (red) can be bound specifically by the correct TF if a specific partner allows the formation of heterodimers (preassembled in solution or assembled *in situ* on the promoter itself). *A*) For similar spacings of the binding sites, in the first promoter (red gene), the binding of TF3 on the nearby binding site will allow TF1 binding and prevent binding of TF2. A different heterodimer will be formed between another partner (TF4, dark blue) and TF2 on the promoter of another gene. *B*) As in homodimer formation, the spacing between the two TFBSs can discriminate between heterodimers. Here, an identical partner will interact, through different dimerization domains, either with TF1 because of the proximity of their respective binding site (expression of the red gene) or with TF2 as their binding site are further apart in the promoter of the second gene (purple). *C*) Specificity can also be achieved by stabilizing a transcription complex through a linking factor.

tory T cell function (37). Another interesting example involves members of the Hox family that bind DNA by making nearly identical contacts with the major groove. However, *in vivo*, their specificity often depends on extended regions flanking the homeodomain that link the latter to a cofactor [*i.e.*, Extradenticle (Exd)]. Indeed, a HOX-binding site in the promoter of the *forkhead* gene is recognized *in vivo* and *in vitro* by the protein Scr (Sex combs reduced) in combination with Exd. However, other Hox-Exd heterodimers fail to properly recognize this sequence. Interestingly, when a mutation is introduced in this target element, it will be bound by a different Hox-Exd heterodimer. It has been shown that amino acid residues located outside the homeodomain help Hox proteins recognize specific Hox-Exd binding sites *via* interactions with the minor groove and that different target sequences elicit different conformational changes of these extrahomeodomain protein segments. Thus, these residues, which are conserved in a paralog-specific manner, might confer specificity (ref. 38 and references therein). This observation can be rationalized as follows. Let us consider that the constants of the interaction between factors TF1, TF2, and TF3 and a target DNA are K_1, K_2, and K_3. The constants of the interaction between the stable heterodimers TF1-TF3 and TF2-TF3 with DNA will be aK_1K_3 and bK_2K_3, where a and b are connected to the free energy released by protein and DNA conformational changes on binding. The gain in global affinity provided by the interaction with the common partner TF3 (*i.e.*, Exd) is the same for both heterodimers and cannot explain any increased discrimination power or specificity. Thus, the most plausible source of discrimination should come from differences in the intrinsic affinities between TF1 and TF2 and their targets but above all from differences in the free energies spent in conformational changes undergone by the interacting partners (*i.e.*, aK_1 and bK_2 should be very different).

The partnership strategy outlined in this section (and discussed further below) makes sense in the light of evolution. The number of protein/DNA-binding pairs that can emerge by molecular evolution is limited. However, binary (and higher-order) combinations of subsets of protein/DNA-binding pairs decrease the number of innovations necessary to increase the complexity of regulatory patterns. It should be noted that, once again, the TFs in question do not have to interact directly but may form a higher-order complex through an intermediary linking (not DNA-binding) factor (Fig. 4*C*).

Finally, note that the effect of a DNA-binding event on gene transcription might not necessarily be gene activation through direct recruitment of the transcriptional machinery. A TF may also recruit histone modifiers or induce local chromatin conformation changes. This is, for instance, the case of FOXO1, which was shown to be able to induce local chromatin decondensation at the insulin growth factor-binding protein 1 (IGFBP1) promoter consequent to binding (39) and that of the glucocorticoid receptor, whose binding has been shown to also induce local chromatin remodeling (40). Despite the varying underlying mechanisms, the basic principles of sequence recognition outlined above remain valid for these cases.

Heterosynergy: heterotypic interactions with the transcriptional machinery

We can now consider the case of transactivation by recruitment of the transcriptional machinery by two different TFs (directly interacting or not) sitting close to each other on a promoter. Heterosynergy (*i.e.*, the simultaneous recruitment of the transcriptional machinery by both bound TFs) is a way to obtain a strong transcriptional response (**Fig. 5**). Indeed, heterosynergy implies that if K_1 and K_2 are the affinities of the transcriptional machinery (or the mediator complex) for DNA-bound TF1 and TF2, the overall affinity for the target protein-DNA complex will be on the order of $K_1 * K_2$. The ability of the transcriptional machinery (or an intermediate complex) to bind TF1 and TF2 simultaneously must depend on the geometry of the promoter (intervening distance and relative orientation of the two TFBS), on the shapes of the transactivation domains of both TFs, and on the ability of each of them to recruit the transcriptional machinery in concert. Alterations in the structure of the transactivation domain of several TFs, including the glucocorticoid receptor, have been found to specifically modulate synergy (41).

HIGHER-ORDER INTERACTIONS: A COMBINATORIAL PARTNER CODE?

We have thus far considered simple cases involving only two generic TFs. However, when one is considering TF families such as those of FOX or HOX TFs, dozens of members are generally available simultaneously in the

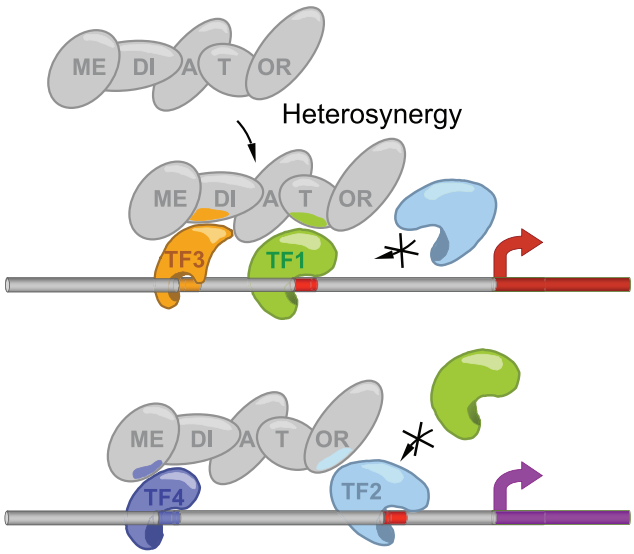

Figure 5. Promoter specificity provided by heterosynergy. As in homosynergy, heterosynergy occurs when the mediator complex is more efficiently recruited by two or more (here, different) factors, instead of only one. Specificity would be ensured by the binding of the multimeric mediator to the relevant DNA-bound TFs (with an appropriate spacing).

cell. This scenario is rendered even more complex if we consider the existence of different post-translational modification (PTM) variants. Indeed, PTM may induce functional modulation by allowing differential recruitment of partners, interfering with TF availability (*i.e.*, nuclear exclusion or shuttling) (42), inducing shifts in DNA-binding affinity and site preference (43, 44), and/or changing TF behavior with respect to transactivation/transrepression (ref. 45 and references therein). For instance, while studying the transcriptional activity of FOXL2, we have noticed that its transactivation capacity is modulated in a promoter-dependent fashion by its acetylation status (ref. 45 and references therein). FOXL2 regulates the activity of the promoters of *SIRT1* and *MnSOD* (mitochondrial superoxide dismutase) genes. This regulation, which relies on FOXL2 acetylation, is enhanced on oxidative stress. However, SIRT1-deacetylated FOXL2 is virtually inactive on the *MnSOD* promoter, but is more efficient than normal FOXL2 on the *SIRT1* promoter. This finding shows that highly acetylated or deacetylated FOXL2 forms are not functionally interchangeable. Moreover, this correlates with the fact that the high-affinity DNA-binding sites that we have derived *in vitro* for a highly modified FOXL2 (from AT29C cells) and for recombinant FOXL2, expressed in bacteria, are rather different downstream of the AAGG core region [*i.e.*, GTC-AAGG-TCA *vs.* GT(C/G)-AAGG-GTG, respectively] (46). One is tempted to assume that, somehow, the sequences flanking the core modulate the type of FOXL2 that is recruited to a particular site (along with as yet unknown partners?).

As already stated, sequences similar to true binding sites are present thousands of times in the genome, and intracellular concentrations of TFs, promoters, and the transcriptional machinery are low. Therefore, regulation, even in the simplest eukaryotes, is expected to imply the cooperation of several TFs (ensuring discrimination of signal from noise). Cooperation is also the basis of information integration flowing from different signal transduction pathways (45). This general feature can be exploited to ensure specificity *via* a combinatorial TF partner code that should mirror particular combinations of TFBS, which is an extension of what we have outlined earlier. The idea of code has to be understood as a system of elements (sequence motifs and the corresponding binding proteins) involving a rule for converting a piece of information (occupancy of DNA features) into another form of representation or response (regulated gene expression). With the use of a code, combinations of a limited set of motifs/proteins lead to a wide variety of different meanings or responses. Namely, combinations of a few DNA motifs (*e.g.*, FKH-binding sites, HOX-binding sites, GC boxes, and CAAT-boxes), separated by specific intervening distances, give rise to a wild variety of different promoters.

When a TF in competition for a site with other TFs is able to establish specific interactions with its neighbors, either directly or through an intermediary bridge, its effective affinity for the tri/multipartite site will be much higher than that of its competitors. For instance, as shown in **Fig. 6**, if TF1, but not TF2, forms a complex with TF3 and TF5, only TF1 (along with its partners) will be able to drive transcription efficiently. This can also be a way to

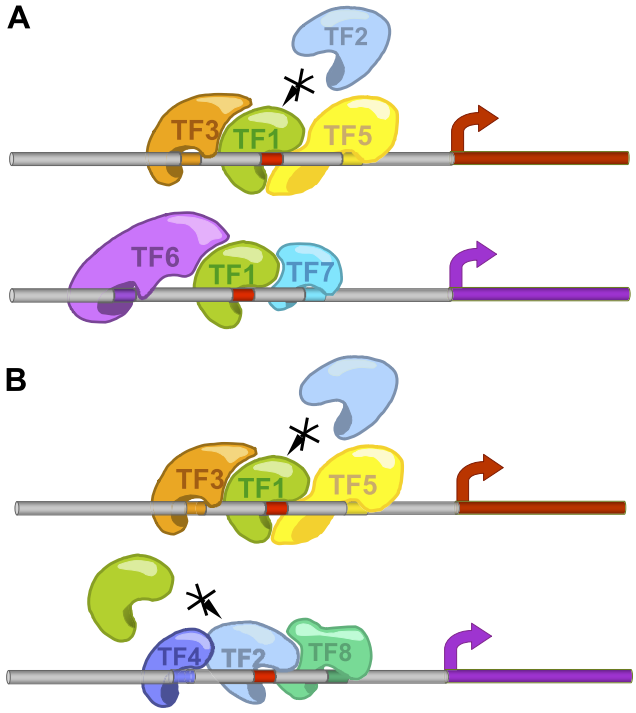

Figure 6. Promoter specificity provided by combinatorial TF binding. *A*) Formation of heterotrimeric protein complexes (in solution or directly on the promoter through heterotypic cooperation) can dictate selection of target promoters. Binding of TF1 to TF3 and TF5, on their respective binding sites, dictates the activation of the red gene, whereas the same TF1, when interacting with TF6 and TF7, is directed to another promoter (purple). *B*) Combinatorial binding of heterotrimeric TF can also lead to the selection of the relevant TF on a promoter, by its binding to proper helper cofactors. Here, the specificity of the binding of TF1 on its binding site is ensured by the binding of both TF3 and TF5. The same binding site in another promoter is specifically bound by TF2, because TF2 interacts with its helper TF4 and TF8. Here, the change of interacting cofactors results in the selection of another TF bound to the common site.

activate different target subsets in the same or different cells. For instance, if TF1 can also interact with TF6 and TF7 in an exclusive way, the complexes TF1.3.5 and TF1.6.7 will be assembled on (and will activate) different target promoters (Fig. 6*A*). From this perspective, promoter recognition is a trial-and-error process in which recognition of the same promoter in different cells might be started by different factors of the same family even with low affinities. The first-landing factor will try to recruit a partner. When two true partners interact with the promoter, the complex becomes much more stable and interaction with other partners will further increase its stability, locking on the complex. This increases the fraction of time the transcription complex exists as such and therefore the probability of a transcription initiation reaction.

Can such a combinatorial mechanism fully explain specific binding for a whole family of related factors? For the sake of simplicity, let us speculate that a TF will establish at most two direct interactions with neighboring TFs (hence, the following reasoning is valid for the central TF). According to this assumption, interactions with only

6 different helper partners will allow the existence of 36 theoretically different tripartite sites, sharing a common central consensus site for the TF of the family in question. This number might be enough to ensure specific interactions for a whole TF family. Figure 6*B* shows the hypothetical example of TF1 and TF2, belonging to the same family, whose specificity in promoter recognition depends on the interaction with different helper cofactors. We have provided an operational definition of helper TFs as those surrounding the central one. Of course, the central TF in a set of promoters can become, in turn, helper for other promoters. This combinatorial strategy is also applicable to cases in which the three interacting TFs belong to families of paralogs. Such proteins can interact with each other in different ways. Thus, the combination of binding sites and intervening distances can lead to an important number of specific tripartite binding sites. This discussion assumes that the idea of classic/homotypic cooperative binding can be extended to include interactions of several different factors, each of them facilitating (and stabilizing) the binding of others (heterotypic cooperativity). This combinatorial mechanism can be further extended to convey specific cell- or signal-dependent transcriptional outputs elicited by the binding of different complexes to the same promoter, as shown in **Fig. 7**.

From an experimental point of view, a myriad of TF-TF interactions have been described. For instance, >24 binding partners have been uncovered for FOXO TFs, including nuclear hormone receptors, Hox factors, and other Fox factors, allowing specific transcriptional modulations (47). For instance, formation of a FOXO/Smad3–4 complex leads to specific up-regulation of the $p21^{cip1}$ gene in response to transforming growth factor-β cytostatic signals. This action is counteracted by an increase in FOXG1 levels, which competes with Smads for interaction with FOXO on the $p21^{cip1}$ gene promoter (47). Members of another family of heterodimeric TFs, the Fos-Jun family, have been shown to interact with >55 different TFs (48). For example, AP-1, a complex involving a cFos-cJun heterodimer, interacts with the functional heterodimer Foxl2-Smad3 to promote transcription of the gonadotropin-releasing hormone receptor in pituitary cells (23).

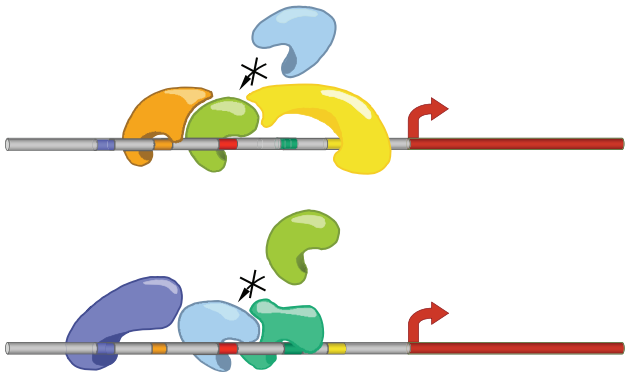

Figure 7. Signal- or cell-specific response provided by a combinatorial TF code on a single promoter. Both TF1 (green) and TF2 (blue) are able to bind to their common binding site (in red), but specificity is achieved by the binding of specific partners to their respective binding site and the formation of distinct complexes. Therefore, a single promoter can be bound alternatively by different transcriptional complexes in response to different signals.

Joint regulation of targets by directly interacting HOX and FOX factors has been suggested previously as a potential conserved mechanism of gene regulation. For instance, Foxa2 and Engrailed antagonize each other to regulate the MAP1B promoter, which contains overlapping homeoprotein and FOX binding sites. In Foxa2 the interacting domain is the Forkhead, whereas in Engrailed, one of the interacting domains is the homeobox itself. Interestingly, Foxa2 binds other homeoproteins such as Lim1, Gsc, and Hoxa5 by, at least, the homeodomain. This finding suggests that interactions between Fox and homeoproteins are a general phenomenon (49). However, from our perspective other cofactors are required to improve specificity of promoter selection (otherwise any promoter having HOX-FOX TFBS would respond). Indeed, the fact that identical HOX-FOX functional heterodimers can lead to different regulatory responses in different contexts might result from the existence of yet unknown interactors and/or potential effects owing to cell- or signal-dependent modifications. This could, for instance, explain why the HOXA5-FOXO complex induces repression of *IGFBP1* in hepatocytes but cooperative activation in fibroblasts (47, 50).

An example of more complex (multipartite) TF-DNA interactions involving FOX and homeobox proteins is provided by FoxJ1 and Pitx2, which exhibit overlapping expression patterns in the dental and oral epithelium. The FoxJ1 promoter is regulated by Pitx2, Lef-1, and β-catenin. Interestingly, FoxJ1 physically interacts with the PITX2 homeodomain to synergistically regulate its own expression in concert with PITX2, Lef-1, and β-catenin, thus generating positive feedback (51).

All in all, heterotypic cooperative binding to DNA and heterosynergy ensuring the joint recruitment of the transcriptional machinery, but also of chromatin modifiers, is another important potential way to specifically regulate transcription.

BACK TO REALITY: THE INTERPLAY OF DIFFERENT STRATEGIES

Multiple strategies can help differentiate TFs with closely related binding domains and ensure specific responses, and all have gained more or less experimental support. There is no obvious reason that evolution would have favored one of these mechanisms over the others. The effects of geometry (*e.g.*, spacing of TFBS or their orientation), of homo- and heterodimerization, and of interaction with many protein partners are complementary. Indeed, they rely on the recognition of multiple TFBS on DNA, which increases the surface of protein/DNA interaction, hence increasing both global affinity and specificity. Each one of these effects taken separately may not be able to fully explain specificity, and their interplay is most probably required to achieve a highly specific gene expression regulation in eukaryotes. This strategy has led to the emergence of enhanceosomes, which are defined as higher-order nucleoprotein complexes, cooperatively assembled, that work as transcription preinitiation/stimulatory complexes. Enhanceosomes contain multiple

TFBS for different effectors involved in combinatorial interactions. An enhanceosome is believed to ensure the formation of a specific activation surface that is complementary to other coactivators (*e.g.*, the mediator complex) or to the transcriptional machinery itself. Interestingly, assembly of an enhanceosome is enhanced (often in a nonsequence-specific way) by architectural proteins that bend DNA, such as the high-mobility group proteins, which allow otherwise energetically forbidden interactions (28, 52).

Before we close this section, it is interesting to mention that seemingly nonspecific (promiscuous) responses may also be important in defined contexts, because different, although closely related, TFs may, by this means, have redundant functions. This would allow these TFs to act as fail-safe mechanisms and ensure robustness of a response when a key component is damaged. This kind of functional redundancy is illustrated quite well by the FoxO TFs, which have largely been characterized as tumor suppressor genes. Surprisingly, single knockout mice are not more tumor-prone than wild-type mice (25–27). However, the combined triple Foxo knockout leads to an increase in tumor development (53). This result suggests that the absence of tumor development in single knockouts results from functional overlap and redundancy, ensuring robustness in tumor suppression by the FoxO subfamiliy of TFs. Therefore, a partial lack of specificity can sometimes foster an evolutionary advantage. FJ

We gratefully thank two anonymous referees for their most helpful comments on a previous version of this manuscript. Work in R.A.V.'s laboratory is supported by the Association pour la Recherche contre le Cancer, the University Paris Diderot-Paris7, the Centre National de la Recherche Scientifique, and the Institut National de la Santé et de la Recherche Médicale. R.A.V. is member of and supported by the Institut Universitaire de France. A.B.G. is funded by Ecole Normale Supérieure de Paris. B.A.B. is funded by an AMN/Université Paris Diderot Paris VII fellowship.

REFERENCES

1. Lewin, B. (2000) Initiation of transcription—promoters for RNA polymerase II have short sequence elements. In *Genes VII*, pp. 634–637, Oxford University Press/Cell Press, New York
2. Ptashne, M., and Gann, A. (1997) Transcriptional activation by recruitment. *Nature* **386,** 569–577
3. Bilu, Y., and Barkai, N. (2005) The design of transcription-factor binding sites is affected by combinatorial regulation. *Genome Biol.* **6,** R103
4. Kaestner, K. H., Knochel, W., and Martinez, D. E. (2000) Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **14,** 142–146
5. Benayoun, B. A., Caburet, S., Dipietromaria, A., Bailly-Bechet, M., Fellous, M., Vaiman, D., and Veitia, R. A. (2008) The identification and characterization of a FOXL2 response element provides insights into the pathogenesis of mutant alleles. *Hum. Mol. Genet.* **17,** 3118–3127
6. Lehmann, O. J., Sowden, J. C., Carlsson, P., Jordan, T., and Bhattacharya, S. S. (2003) Fox's in development and disease. *Trends Genet.* **19,** 339–344
7. Greer, E. L., and Brunet, A. (2005) FOXO transcription factors at the interface between longevity and tumor suppression. *Oncogene* **24,** 7410–7425
8. Lee, C. S., Friedman, J. R., Fulmer, J. T., and Kaestner, K. H. (2005) The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435,** 944–947
9. Hueber, S. D., and Lohmann, I. (2008) Shaping segments: Hox gene function in the genomic age. *Bioessays* **30,** 965–979
10. Klein-Hitpass, L., Ryffel, G. U., Heitlinger, E., and Cato, A. C. (1988) A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. *Nucleic Acids Res.* **16,** 647–663
11. Little, T. H., Zhang, Y., Matulis, C. K., Weck, J., Zhang, Z., Ramachandran, A., Mayo, K. E., and Radhakrishnan, I. (2006) Sequence-specific deoxyribonucleic acid (DNA) recognition by steroidogenic factor 1: a helix at the carboxy terminus of the DNA binding domain is necessary for complex stability. *Mol. Endocrinol.* **20,** 831–843
12. Wilson, T. E., Fahrner, T. J., and Milbrandt, J. (1993) The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. *Mol. Cell. Biol.* **13,** 5794–5804
13. Sladek, R., Bader, J. A., and Giguere, V. (1997) The orphan nuclear receptor estrogen-related receptor α is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. *Mol. Cell. Biol.* **17,** 5400–5409
14. Cooney, A. J., Tsai, S. Y., O'Malley, B. W., and Tsai, M. J. (1992) Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. *Mol. Cell. Biol.* **12,** 4153–4163
15. Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* **65,** 1255–1266
16. Mohn, F., and Schubeler, D. (2009) Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends Genet.* **25,** 129–136
17. Lammens, T., Li, J., Leone, G., and De Veylder, L. (2009) Atypical E2Fs: new players in the E2F transcription factor family. *Trends Cell Biol.* **19,** 111–118
18. Welboren, W. J., van Driel, M. A., Janssen-Megens, E. M., van Heeringen, S. J., Sweep, F. C., Span, P. N., and Stunnenberg, H. G. (2009) ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* **28,** 1418–1428
19. Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell* **75,** 567–578
20. Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* **375,** 203–211
21. Courey, A. J. (2001) Cooperativity in transcriptional control. *Curr. Biol.* **11,** R250–R252
22. Blount, A. L., Schmidt, K., Justice, N. J., Vale, W. W., Fischer, W. H., and Bilezikjian, L. M. (2009) FoxL2 and Smad3 coordinately regulate follistatin gene transcription. *J. Biol. Chem.* **284,** 7631–7645
23. Ellsworth, B. S., Burns, A. T., Escudero, K. W., Duval, D. L., Nelson, S. E., and Clay, C. M. (2003) The gonadotropin releasing hormone (GnRH) receptor activating sequence (GRAS) is a composite regulatory element that interacts with multiple classes of transcription factors including Smads, AP-1 and a forkhead DNA binding protein. *Mol. Cell. Endocrinol.* **206,** 93–111
24. Salih, D. A., and Brunet, A. (2008) FoxO transcription factors in the maintenance of cellular homeostasis during aging. *Curr. Opin. Cell Biol.* **20,** 126–136
25. Nakae, J., Kitamura, T., Kitamura, Y., Biggs, W. H., 3rd, Arden, K. C., and Accili, D. (2003) The forkhead transcription factor Foxo1 regulates adipocyte differentiation. *Dev. Cell* **4,** 119–129
26. Castrillon, D. H., Miao, L., Kollipara, R., Horner, J. W., and DePinho, R. A. (2003) Suppression of ovarian follicle activation in mice by the transcription factor Foxo3a. *Science* **301,** 215–218
27. Hosaka, T., Biggs, W. H., 3rd, Tieu, D., Boyer, A. D., Varki, N. M., Cavenee, W. K., and Arden, K. C. (2004) Disruption of forkhead transcription factor (FOXO) family members in mice reveals their functional diversification. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 2975–2980
28. Goff, D. J., and Tabin, C. J. (1997) Analysis of Hoxd-13 and Hoxd-11 misexpression in chick limb buds reveals that Hox

genes affect both bone condensation and growth. *Development* **124,** 627–636
29. Carey, M. (1998) The enhanceosome and transcriptional synergy. *Cell* **92,** 5–8
30. Herbomel, P. (1990) Synergistic activation of eukaryotic transcription: the multiacceptor target hypothesis. *New Biol.* **2,** 1063–1070
31. Veitia, R. A. (2003) A sigmoidal transcriptional response: cooperativity, synergy and dosage effects. *Biol. Rev. Camb. Philos. Soc.* **78,** 149–170
32. Gibson, G. (1996) Epistasis and pleiotropy as natural properties of transcriptional regulation. *Theor. Popul. Biol.* **49,** 58–89
33. Veitia, R. A., and Nijhout, H. F. (2006) The robustness of the transcriptional response to alterations in morphogenetic gradients. *Bioessays* **28,** 282–289
34. Casamassimi, A., and Napoli, C. (2007) Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie* **89,** 1439–1446
35. Wang, B., Lin, D., Li, C., and Tucker, P. (2003) Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors. *J. Biol. Chem.* **278,** 24259–24268
36. Pohl, B. S., and Knochel, W. (2004) Isolation and developmental expression of *Xenopus* FoxJ1 and FoxK1. *Dev. Genes Evol.* **214,** 200–205
37. Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A. D., Stroud, J. C., Bates, D. L., Guo, L., Han, A., Ziegler, S. F., Mathis, D., Benoist, C., Chen, L., and Rao, A. (2006) FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* **126,** 375–387
38. Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., Jacob, V., Aggarwal, A. K., Honig, B., and Mann, R. S. (2007) Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* **131,** 530–543
39. Hatta, M., and Cirillo, L. A. (2007) Chromatin opening and stable perturbation of core histone:DNA contacts by FoxO1. *J. Biol. Chem.* **282,** 35583–35593
40. Hager, G. L., Fletcher, T. M., Xiao, N., Baumann, C. T., Muller, W. G., and McNally, J. G. (2000) Dynamics of gene targeting and chromatin remodelling by nuclear receptors. *Biochem. Soc. Trans.* **28,** 405–410
41. Iniguez-Lluhi, J. A., and Pearce, D. (2000) A common motif within the negative regulatory regions of multiple factors inhibits their transcriptional synergy. *Mol. Cell. Biol.* **20,** 6040–6050
42. Brunet, A., Bonni, A., Zigmond, M. J., Lin, M. Z., Juo, P., Hu, L. S., Anderson, M. J., Arden, K. C., Blenis, J., and Greenberg, M. E. (1999) Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96,** 857–868
43. Brent, M. M., Anand, R., and Marmorstein, R. (2008) Structural basis for DNA recognition by FoxO1 and its regulation by posttranslational modification. *Structure* **16,** 1407–1416
44. Tsai, K. L., Sun, Y. J., Huang, C. Y., Yang, J. Y., Hung, M. C., and Hsiao, C. D. (2007) Crystal structure of the human FOXO3a-DBD/DNA complex suggests the effects of post-translational modification. *Nucleic Acids Res.* **35,** 6984–6994
45. Benayoun, B. A., and Veitia, R. A. (2009) A post-translational modification code for transcription factors: sorting through a sea of signals. *Trends Cell Biol.* **19,** 189–197
46. Benayoun, B. A., Caburet, S., Dipietromaria, A., Bailly-Bechet, M., Batista, F., Fellous, M., Vaiman, D., and Veitia, R. A. (2008) The identification and characterization of a FOXL2 response element provides insights into the pathogenesis of mutant alleles. *Hum. Mol. Genet.* **17,** 3118–3127
47. Van der Vos, K. E., and Coffer, P. J. (2008) FOXO-binding partners: it takes two to tango. *Oncogene* **27,** 2289–2299
48. Chinenov, Y., and Kerppola, T. K. (2001) Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene* **20,** 2438–2452
49. Foucher, I., Montesinos, M. L., Volovitch, M., Prochiantz, A., and Trembleau, A. (2003) Joint regulation of the MAP1B promoter by HNF3β/Foxa2 and Engrailed is the result of a highly conserved mechanism for direct interaction of homeoproteins and Fox transcription factors. *Development* **130,** 1867–1876
50. Foucher, I., Volovitch, M., Frain, M., Kim, J. J., Souberbielle, J. C., Gan, L., Unterman, T. G., Prochiantz, A., and Trembleau, A. (2002) Hoxa5 overexpression correlates with IGFBP1 upregulation and postnatal dwarfism: evidence for an interaction between Hoxa5 and Forkhead box transcription factors. *Development* **129,** 4065–4074
51. Venugopalan, S.R., Amen, M.A., Wang, J., Wong, L., Cavender, A.C., D'Souza, R.N., Akerlund, M., Brody, S.L., Hjalt, T.A., and Amendt, B.A. (2008) Novel expression and transcriptional regulation of FoxJ1 during oro-facial morphogenesis. *Hum. Mol. Genet.* **17,** 3643–3654
52. Merika, M., and Thanos, D. (2001) Enhanceosomes. *Curr. Opin. Genet. Dev.* **11,** 205–208
53. Paik, J. H., Kollipara, R., Chu, G., Ji, H., Xiao, Y., Ding, Z., Miao, L., Tothova, Z., Horner, J. W., Carrasco, D. R., Jiang, S., Gilliland, D. G., Chin, L., Wong, W. H., Castrillon, D. H., and DePinho, R. A. (2007) FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* **128,** 309–323
54. Pierrou, S., Hellqvist, M., Samuelsson, L., Enerback, S., and Carlsson, P. (1994) Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J.* **13,** 5002–5012
55. Overdier, D. G., Porcella, A., and Costa, R. H. (1994) The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. *Mol. Cell. Biol.* **14,** 2755–2766
56. Zeng, J., Yan, J., Wang, T., Mosbrook-Davis, D., Dolan, K. T., Christensen, R., Stormo, G. D., Haussler, D., Lathrop, R. H., Brachmann, R. K., and Burgess, S. M. (2008) Genome wide screens in yeast to identify potential binding sites and target genes of DNA-binding proteins. *Nucleic Acids Res.* **36,** e8
57. Nirula, A., Moore, D. J., and Gaynor, R. B. (1997) Constitutive binding of the transcription factor interleukin-2 (IL-2) enhancer binding factor to the IL-2 promoter. *J. Biol. Chem.* **272,** 7736–7745
58. Furuyama, T., Nakazawa, T., Nakano, I., and Mori, N. (2000) Identification of the differential distribution patterns of mRNAs and consensus binding sequences for mouse DAF-16 homologues. *Biochem. J.* **349,** 629–634
59. Wang, L. H., Chmelik, R., and Nirenberg, M. (2002) Sequence-specific DNA binding by the vnd/NK-2 homeodomain of *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12721–12726
60. Draganescu, A., and Tullius, T. D. (1998) The DNA binding specificity of engrailed homeodomain. *J. Mol. Biol.* **276,** 529–536
61. Nishio, S., Kakizawa, T., Chatelain, G., Triqueneaux, G., Brunet, F., Rambaud, J., Lamonerie, T., and Laudet, V. (2008) OTX5 regulates pineal expression of the zebrafish REV-ERBα through a new DNA binding site. *Mol. Endocrinol.* **22,** 23–32
62. Feledy, J. A., Morasso, M. I., Jang, S. I., and Sargent, T. D. (1999) Transcriptional activation by the homeodomain protein distalless 3. *Nucleic Acids Res.* **27,** 764–770
63. Trindade, M., Tada, M., and Smith, J. C. (1999) DNA-binding specificity and embryological function of Xom (Xvent-2). *Dev. Biol.* **216,** 442–456
64. Margalit, Y., Yarus, S., Shapira, E., Gruenbaum, Y., and Fainsod, A. (1993) Isolation and characterization of target sequences of the chicken CdxA homeobox gene. *Nucleic Acids Res.* **21,** 4915–4922
65. Knosp, W. M., Saneyoshi, C., Shou, S., Bachinger, H. P., and Stadler, H. S. (2007) Elucidation, quantitative refinement, and in vivo utilization of the HOXA13 DNA binding site. *J. Biol. Chem.* **282,** 6843–6853
66. Bertolino, E., Reimund, B., Wildt-Perinic, D., and Clerc, R. G. (1995) A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *J. Biol. Chem.* **270,** 31178–31188
67. Rao, E., Blaschke, R. J., Marchini, A., Niesler, B., Burnett, M., and Rappold, G. A. (2001) The Leri-Weill and Turner syndrome homeobox gene SHOX encodes a cell-type specific transcriptional activator. *Hum. Mol. Genet.* **10,** 3083–3091

Received for publication July 18, 2009.
Accepted for publication August 27, 2009.