Transcriptional Regulation in *Saccharomyces cerevisiae*: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators

Steven Hahn*† and Elton T. Young*

*Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, and †Department of Biochemistry, University of Washington, Seattle, Washington 98195

ABSTRACT Here we review recent advances in understanding the regulation of mRNA synthesis in *Saccharomyces cerevisiae*. Many fundamental gene regulatory mechanisms have been conserved in all eukaryotes, and budding yeast has been at the forefront in the discovery and dissection of these conserved mechanisms. Topics covered include upstream activation sequence and promoter structure, transcription factor classification, and examples of regulated transcription factor activity. We also examine advances in understanding the RNA polymerase II transcription machinery, conserved coactivator complexes, transcription activation domains, and the cooperation of these factors in gene regulatory mechanisms.

TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>705</td>
</tr>
<tr>
<td>Introduction</td>
<td>706</td>
</tr>
<tr>
<td>Upstream Activation Sequence Elements</td>
<td></td>
</tr>
<tr>
<td>Identification of UAS elements</td>
<td>706</td>
</tr>
<tr>
<td>UAS conservation and evolution</td>
<td>707</td>
</tr>
<tr>
<td>UAS function and combinatorial control</td>
<td>707</td>
</tr>
<tr>
<td>Transcription Factors</td>
<td></td>
</tr>
<tr>
<td>TF classification</td>
<td>708</td>
</tr>
<tr>
<td>Modularity</td>
<td>709</td>
</tr>
<tr>
<td>Functions of TF-DNA binding</td>
<td>709</td>
</tr>
<tr>
<td>Functions of TF-activating and -repressing transcription</td>
<td>710</td>
</tr>
<tr>
<td>Transcriptional Regulation</td>
<td></td>
</tr>
<tr>
<td>Adr1 Phosphorylation-dependent inhibition of Adr1 by Bmh</td>
<td>710</td>
</tr>
<tr>
<td>Gal4 regulation by intermolecular AD masking</td>
<td>711</td>
</tr>
<tr>
<td>Hap1 and Mal63 regulation by chaperones</td>
<td>712</td>
</tr>
<tr>
<td>Hsf1 regulation through conformational change</td>
<td>712</td>
</tr>
</tbody>
</table>

Copyright © 2011 by the Genetics Society of America
doi:10.1534/genetics.111.127019
Manuscript received January 19, 2011; accepted for publication February 8, 2011

*Corresponding author: Fred Hutchinson Cancer Research Center, PO Box 19024, 1100 Fairview Ave. N, Mailstop a1-162, Seattle, WA 98109. E-mail: shahn@fhcrc.org
I n concluding his 1995 review of yeast transcriptional regulation, Kevin Struhl (1995) proposed three major questions to direct future research: (1) How do activators and repressors affect the basic transcriptional machinery of the cell? (2) How are the activities of regulators themselves regulated? and (3) How are the various regulatory pathways integrated to coordinate cell growth and response to external signals? Studies in yeast have made seminal contributions in each of these areas in the subsequent 15 years. Enormous strides have been made to answer the first question. The structures of RNA polymerase (Pol) II and several general transcription factors have been determined, and the yeast system has been at the forefront in discoveries of fundamental transcription mechanisms. Transcription activators and repressors act, in part, by recruitment of the transcription machinery or repression complexes to gene regulatory regions. How the regulators are regulated (question 2) surprisingly seems to have almost as many answers as there are regulators in the cell. Question 3 is the subject of current “omics” research. Genome-wide methods are being used to analyze global expression and DNA binding by many transcription factors assayed under varied growth conditions and in multiple yeast species. Below, we provide an outline of recent advances in understanding the mechanisms of *Saccharomyces cerevisiae* transcription factors, the general transcription machinery, and the cooperation of these factors in transcriptional regulation.

**Upstream Activation Sequence Elements**

Transcriptional regulation begins with sequence-specific recognition of unique DNA elements by transcription factors (TFs), either transcription activators (TAs) binding to upstream activation sequences (UASs) or repressors (TRs) binding to upstream repression sequences (URSs). UASs and their associated TFs are probably required for expression of
all protein-coding genes transcribed by Pol II, which contain one or more UASs. Although there is a low level of non-activator-dependent “basal” transcription in vitro, basal transcription is repressed when a chromatin template, rather than free DNA, is used, suggesting that nucleosomes repress non-TF-dependent transcription initiation in vivo (Juan et al. 1993). Thus, the “ground-state” of a yeast promoter is inactive and transcription must be promoted by one or more TFs (Struhl 1999). Poly(dA:dT) elements (Iyer and Struhl 1995b) and altered chromatin (Han and Grunstein 1988; Han et al. 1988) may stimulate activator-independent transcription.

**Identification of UAS elements**

The classical method of UAS (or URS) identification utilizes fusions of regulatory region DNA or deletion derivatives to a core promoter and reporter gene such as *Escherichia coli* lacZ followed by expression analysis (Sentenac and Hall 1982; Struhl 1989, 1995). A technique for generalizing this experimental approach has recently been described (Doyon et al. 2010). However, sequencing of multiple yeast genomes has allowed a bioinformatics approach to largely supplant the classical approach for discovering and characterizing regulatory motifs. The development of algorithms that can find conserved or known motifs in promoter sequences has been instrumental, particularly when combined with global TF-binding data (Hu et al. 2010; Reid et al. 2010). Phylogenetic footprinting (Cliffen et al. 2003; Kellis et al. 2003; Borneman et al. 2007; Tuch et al. 2008; Lavoie et al. 2010) and nucleosome location data (Narlikar et al. 2007; Gordan et al. 2009) can enhance the probability that conserved motifs represent functional UASs (Wingender et al. 1996; Monteiro et al. 2008; Park et al. 2010). Genomewide mapping of 5’ mRNA ends (Miura et al. 2006; Nagar- lakshmi et al. 2008) and newly discovered or revised binding site preferences for >100 yeast transcription factors have also improved the ability to associate conserved motifs with known or suspected TFs. Proof of UAS/URS function still requires chromatin immunoprecipitation (ChIP) experiments to demonstrate TF occupancy and mutation of the binding site to demonstrate function.

**UAS conservation and evolution**

Evolution can be studied in yeast because yeast has a relatively short doubling time and can utilize a variety of nutrients, and its genome can be analyzed rapidly and completely (Dunham et al. 2002). The current paradigm in molecular evolution is that phenotypic diversity, and perhaps speciation, has primarily occurred by altering gene regulation rather than by altering the function of individual proteins. The results in yeast demonstrate that regulatory circuits are evolving rapidly using conserved TFs to regulate different sets of genes and are being driven by changes in both binding motifs and TFs (Tseng et al. 2003, 2006; Borneman et al. 2007; Chang et al. 2008; Tuch et al. 2008; Zheng et al. 2010).

**UAS function and combinatorial control**

UAS/URS function is generally orientation independent but dependent on being 5’ of the promoter with a few exceptions (Errede et al. 1984; Mellor et al. 1987; Company and Errede 1988; Fantino et al. 1992; Gray and Fassler 1996). Their number and position may influence the level of expression (Swamy et al. 2009). It is unknown why most yeast enhancers (UASs) cannot function 3’ of the promoter, but enhancers in higher eukaryotes can function both 5’ and 3’ of the promoter. The TF Gal4 can operate from both up- and downstream of +1 in higher eukaryotes (Webster et al. 1988) and can also function from a site >1 kb downstream of the start site of transcription when it binds near a telomere (de Bruin et al. 2001). Thus the position-dependent activity of UASs in yeast does not reside in the TF and may be due to unknown differences in either chromatin or the core promoter.

Although the exact distance of the UAS from the transcription start site may not be important, most functional UASs appear to be located in the nucleosome-depleted region of yeast promoters directly upstream of the transcription start site or are exposed on the surface of nucleosomes (Xue et al. 2004; Albert et al. 2007; Lee et al. 2007; reviewed in Li et al. 2007). One interpretation of this observation is that TF accessibility is adversely affected by the presence of...
a nucleosome, and thus the depleted region presents unimpeaded access to its binding site.

When different types of UAS are present in the same promoter, they allow combinatorial control of transcription. The complex dynamics of gene expression during the cell cycle and after subjecting cells to stress have been assayed by ChIP analysis of the regulatory TFs and coactivators (McBride et al. 1997; Bhoite et al. 2001; Simon et al. 2001; Horak et al. 2002; Tan et al. 2008; Ni et al. 2009). Similarly, during sporulation (Kahana et al. 2010) and glucose starvation (Young et al. 2003; Tachibana et al. 2005; Ratnakumar and Young 2010) multiple TFs bind upstream of the promoter of coregulated genes. In the sulfur metabolic network, the non-DNA-binding Met4 activator interacts with multiple cofactors that stabilize DNA binding (Lee et al. 2010). Mcm1 heterodimerizes with different TFs to regulate cell-type and other processes (Tan and Richmond 1998; Hollenhorst et al. 2001; Tuch et al. 2008).

Transcription Factors

At present there are 169 genes designated as TFs in the Yeastact database (http://www.yeastact.com; Teixeira et al. 2006), making TFs one of the most abundant classes of proteins in the yeast genome. Analysis of other Saccharomyces genomes has allowed a comprehensive comparison of TFs between species (http://www.csmi.info/priloha/fm2008_drobna_tab2.pdf; Drobna et al. 2008). Many yeast TFs were discovered by genetic means because they influence expression of downstream target genes. Additional TFs were identified by homology. Whether all yeast TFs can be identified by these approaches is an open question. Recently, a metabolic enzyme involved in ornithine biosynthesis, Arg5,6 was identified in a screen for DNA-binding proteins (Hall et al. 2004). Arg5,6 represents just one of several examples in yeast of so-called “moonlighting” proteins, in this case a DNA-binding protein that also serves as an enzyme of intermediary metabolism (Gancedo and Flores 2008).

TF classification

TFs are grouped into three general classes on the basis of the type of DNA-binding domain (DBD) that they contain: zinc (Zn)\(^{2+}\) stabilized, helix-turn-helix, and zipper type; the classes were updated in 2000 using analysis of structural information in the Brookhaven Protein Database (http://www.biochem.ucl.ac.uk/bsm/prot_dna/prot_dna.html; Luscombe et al. 2000). Figure 1, A–E, illustrates five yeast TF DBDs.

The Zn\(^{2+}\)-stabilized DBD is the most abundant in all organisms and can be subdivided (Krishna et al. 2003) into C2H2 zinc fingers (Bohm et al. 1997), C6 (zinc knuckle or Zn2Cys6 binuclear zinc cluster) (MacPherson et al. 2006), and C4 or GATA fingers (Scanziochocchio 2000). Yeast also has at least one TF whose DNA-binding domain is stabilized by Cu (Zn)\(^{2+}\) (e.g., Ace1/Cup2). C2H2 zinc-finger proteins are ubiquitous as are C4 (GATA) factors. The C6 or zinc-knuckle type is unique to fungi.

C2H2 zinc fingers (53 members—e.g., Adr1, Mig1, Zap1) have a modular structure that is stabilized by tetrahedral coordination of a zinc ion by Cys and His ligands and by a hydrophobic core of conserved Phe and Leu residues (Bohm et al. 1997) (Figure 1A). The C terminus is an α-helix whose N-terminal amino acids confer DNA-binding specificity. The C2H2 TFs generally bind DNA as monomers with each finger recognizing consecutive triplets of bases; specificity and high affinity are achieved by multiple fingers.

The C6 proteins (55 members—e.g., Gal4, Mal63, Hap1, Leu3) have a DBD containing two zinc ions liganded to six Cys residues (MacPherson et al. 2006) (Figure 1B). The DBD is N-terminal in most C6 TFs with the DNA-binding residues C-terminal to the C6 region. The C6 TFs bind DNA sites containing CGG triplets flanking a region containing a variable number of residues. The orientation of the CGG triplets and their spacing are the primary determinants of DNA-binding specificity (Reece and Ptashne 1993). The C6 proteins bind as dimers to symmetric sites, utilizing a dimerization domain C-terminal to the DBD. Many C6 proteins bind DNA exclusively as homodimers, such as Gal4 and Leu3, whereas others, such as Oaf1 and Pip2 and Pdr1 and Pdr3, can form heterodimers as well. A few, such as Rgt1, are thought to bind DNA as monomers.

The C4 or GATA class of zinc-stabilized TFs consists of proteins that are primarily involved in nitrogen metabolism in S. cerevisiae (Cooper 2002) (5 members—Gln3, Gap1, Nil1, Dal80, Ash1). Ash1 represses HMothetical switching endonuclease (HO) transcription specifically in daughter cells through cytoskeletal-regulated partitioning of its mRNA in the nucleus (Cosma 2004), where it activates the filamentation pathway. In other fungi, the GATA factors are involved in multiple pathways that include mating-type switching (Scanziochocchio 2000).

The second most abundant class of TF is the zipper type (22 members). This class is characterized by a DBD consisting of a dimerization motif and a basic region. The first subclass, bZIP proteins (14 members—e.g., Gcn4, Yap1-8, Sko1), also called leucine zippers, have a dimerization domain consisting of multiple Leu residues (Fernandes et al. 1997; Moye-Rowley 2003) (Figure 1C). The second subclass, bHLH proteins (9 members—e.g., Ino2, Ino4, Pho2, Pho4), has paired amphipathic α-helices separated by a loop of variable length followed by a basic region (Robinson and Lopes 2000) (Figure 1D). These proteins generally bind DNA as heterodimers, giving rise to a multitude of different complexes (Fernandes et al. 1997; Robinson and Lopes 2000; Moye-Rowley 2003; Chen and Lopes 2007, 2010; Tan et al. 2008).

The third most abundant class, helix-turn-helix (HTH) TFs (Mat1, Mat2; Mata1; eight members), are most closely related to homeodomain-containing proteins in higher eukaryotes and to prokaryotic repressors and activators. They form both homo- and heterodimers. The classical HTH protein in yeast is Mata2, which, together with Mcm1, represses a specific gene in Matα haploids (Figure 1E) and
forms a heterodimer with Mata1 to repress haploid-specific genes in MATα/MATα diploid cells.

The forkhead (Fkh) or MADS-box transcription factors Mcm1, Fkh1, Fkh2, and Hcm1 and the heat-shock factor Hsf1 are related to the HTH proteins (Tan and Richmond 1998; Kaestner et al. 2000). Three helices and two large loops or “wings,” which led to the name “winged helix,” form the DBD. Mcm1 associates with several proteins and acts as both an activator and a repressor to control cell-specific gene expression (Elble and Tye 1991). Fkh proteins are involved in numerous processes, including cell cycle regulation, where they appear to endow other transcription factors with promoter specificity (Hollenhorst et al. 2001; Voth et al. 2007).

**Modularity**

TFs have a modular structure that consists of multiple, independently functioning domains. One consequence of this modularity is that, unlike globular proteins, they have been resistant to structural analysis as a single entity. Some proteins classified as TFs, such as Met4 (Lee et al. 2010) and Swi6 (Sidorova and Breeden 1993), lack a DBD motif and interact with DNA through a binding partner. Other TFs, such as Gcr1 and Dal81, have a DBD that is dispensable, apparently because it forms a heterodimer and the other subunit is sufficient for DNA binding (Bricmont et al. 1991; Scott et al. 2000; Tornow et al. 1993). The effector domains of a TF may include an activation/repression domain (AD/RD), a nuclear localization sequence (NLS), and a regulatory domain. The first evidence suggesting the modular nature of eukaryotic TFs accompanied the discovery of the Gen4 and Gal4 ADs (Hope and Struhl 1986; Ma and Ptashne 1987b). The ability of Gal4 to function universally attests to the conserved nature of the activation process in fungi, plants, insects, and metazoans (Fischer et al. 1988; Kakidani and Ptashne 1988; Ma et al. 1988; Webster et al. 1988).

ADs, sometimes more than one, have been documented in numerous TAs (Table 1, “Activation domains”). Although most yeast TAs possess both a DBD and an AD within the same polypeptide, in some heterodimeric TFs only one subunit has an AD. The ReTroGrade (RTG) response, a mechanism by which the mitochondria and nuclear compartments communicate (Butow and Avadhani 2004), is orchestrated by two TFs, Rtg1 and Rtg3. These bHLH proteins form a heterodimer in which only Rtg3 contains an AD that responds to the RTG signal (Rothermel et al. 1997). RDs have been found in many TFs that have a negative role in gene expression (Table 1, “Repression domains”). Some of these proteins, including Ash1, Rgt1, and Rap1, have the ability to both activate and repress transcription, depending on the promoter, chromatin context, and growth conditions.

NLSs have been identified by a similar approach by making deletion mutations in the native protein and/or by creating chimeras, usually by fusion to E. coli β-galactosidase, and assessing its intracellular location (Silver et al. 1984, 1986). This approach has identified NLSs on numerous yeast TFs (Hahn et al. 2008). The receptors for classical NLSs are nuclear importins functioning as α/β heterodimers where the α-subunit recognizes the NLS and the β-subunit recognizes the nuclear pore complex (Silver et al. 1989; Brodsky and Silver 1999).

**Functions of TF-DNA binding**

The DNA-binding specificity of TFs largely determines where it acts in the genome. Recent high-throughput biochemical screens identified new and revised binding sequences for a large number of TFs (Badis et al. 2008; Zhu et al. 2009). Gene expression analysis in strains deleted for or overexpressing specific TFs have identified both direct and indirect TF targets. Short-term overexpression appears to identify direct targets more reliably than analysis of deletion strains (Chua et al. 2006; Sopko et al. 2006). In some cases, direct and indirect effects can be distinguished by integrating binding-site information with data on nucleosome positions and gene expression (Beyer et al. 2006; Gordan et al. 2009).

ChIP-chip and ChIP-Seq (chromatin IP followed by hybridization to microarrays or used in high-throughput

<table>
<thead>
<tr>
<th>Table 1 Yeast transcription factors with demonstrated activation or repression domains</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transcription factors</strong></td>
</tr>
<tr>
<td>Abf1</td>
</tr>
<tr>
<td>ADR1</td>
</tr>
<tr>
<td>Ash1</td>
</tr>
<tr>
<td>Gal4</td>
</tr>
<tr>
<td>Gcn4</td>
</tr>
<tr>
<td>Gcr1</td>
</tr>
<tr>
<td>Hap4</td>
</tr>
<tr>
<td>Hsf1</td>
</tr>
<tr>
<td>Ime1</td>
</tr>
<tr>
<td>Ino2/Ino4</td>
</tr>
<tr>
<td>Lys14</td>
</tr>
<tr>
<td>Mga2</td>
</tr>
<tr>
<td>Met4</td>
</tr>
<tr>
<td>Rap1</td>
</tr>
<tr>
<td>Rim1, Smp1</td>
</tr>
<tr>
<td>Rtg3</td>
</tr>
<tr>
<td>Spt23</td>
</tr>
<tr>
<td>Tea1</td>
</tr>
<tr>
<td>Ume6</td>
</tr>
<tr>
<td>Zap1</td>
</tr>
<tr>
<td>Ash1</td>
</tr>
<tr>
<td>Mata2</td>
</tr>
<tr>
<td>Mig1</td>
</tr>
<tr>
<td>Rap1</td>
</tr>
<tr>
<td>Rgt1</td>
</tr>
<tr>
<td>Rme1</td>
</tr>
<tr>
<td>Rox1</td>
</tr>
<tr>
<td>Ume6</td>
</tr>
</tbody>
</table>
DNA sequencing, respectively) are the gold standard for determining TF-binding sites in vivo. ChIP-Seq is particularly valuable because it can provide nucleotide-level resolution of TF-binding sites (Guo et al. 2010). Condition-dependent binding of a TF is one potential complication of ChIP analysis as is the possibility that a bound TF may be inactive (Gao et al. 2004). Genome-wide localization analysis has been reported for TFs involved in numerous pathways [Saccharomyces Genome Database, (http://www.yeastgenome.org/); Yeastract (http://www.yeasttract.com)]. Determining how TFs achieve their remarkable promoter-binding specificity is an important goal (Georges et al. 2010).

**Functions of TF-activating and -repressing transcription**

As described more fully below, much data support the recruitment model of transcription activation proposed by Ptashne (Ptashne 1988; Ptashne and Gann 1997). It has also been proposed that gene localization within the nucleus is sometimes regulated in response to regulatory signals (Menon et al. 2005; Sarma et al. 2007). According to this model, genes and their associated activators move to the nuclear periphery where they encounter the transcriptional machinery at nuclear pore complexes. It has been suggested that retention at the nuclear periphery could explain the phenomenon of transcriptional memory, the ability of an induced but subsequently repressed gene to be rapidly reactivated (Brickner 2009).

The recruitment model has also been invoked to explain the function of RDs. The TRs Matα2, Mig1, Rgt1, Rox1, Rfx1, and Sko1 repress transcription by recruiting the Tup1–Ssn6/Cyc8 complex to promoters (Smith and Johnson 2000; Courey and Jia 2001). Repression occurs by a histone-dependent pathway (Wu et al. 2001; Davie et al. 2002) and through interactions with RNA Pol II and the coactivator termed Mediator (Gromoller and Lehning 2000; M. Lee et al. 2000; Papamichos-Chronakis et al. 2000). The extent of the repressed chromatin domain can be quite large, as at the silent copies of the MAT locus and at telomeres. Repression at these loci is due in part to the Rap1-dependent recruitment of the Sir2/3/4 complex (Wyrick et al. 1999). The extent of the repressed chromatin domain, characterized by hypoacetylated histones, is controversial (Ducker and Simpson 2000; Courey and Jia 2001; Wu et al. 2001).

Ume6, OpI, and Ash1 modify chromatin and repress transcription by recruiting the Sin3–Rpd3 histone deacetylase complex to create a localized domain of hypoacetylated histones H3 and H4 (Kadosh and Struhl 1998; Carrozza et al. 2005). Ume6 represses meiotic genes during vegetative growth but activates them during sporulation (Rubin-Bejerano et al. 1996; Kassir et al. 2003). Ash1 similarly has both repressive and activating functions (Chandralapaty and Errede 1998; Maxon and Herskowitz 2001). The Rme1 repressor is unusual in that it has overlapping repression and activation domains and no apparent requirement for specific DNA binding (Blumental-Perry et al. 2002). All known TRs recruit complexes that maintain chromatin in a generally repressive state that may inhibit the binding of an activator and the recruitment of coactivators and prevent chromatin remodeling or block a subsequent step in transcription.

**Transcriptional Regulation**

Some pathways in yeast are transcriptionally controlled by altering the expression level rather than the activity of the TFs involved. The prototypical yeast activator Gcn4 is one such factor whose level is controlled by translational readthrough of short upstream open reading frames and by ubiquitination-dependent turnover (Hinnebusch 2005). Cell-type determination is another example of complex regulation that occurs at multiple levels (Galgoczy et al. 2004).

The nuclear localization of numerous TFs is altered by regulated interaction with nuclear importins and karyopherins (Komeili and O’Shea 2000). In contrast, the TFs described below represent TAs whose activation function is regulated by inter- or intramolecular interactions triggered by ligand binding, phosphorylation, or stress.

**Adr1 Phosphorylation-dependent inhibition of Adr1 by Bmh**

Adr1 activates catabolic pathways that are essential for growth in the absence of a fermentable sugar (derepression) (Young et al. 2003). The AMP-activated, protein kinase homolog Snf1 is essential for promoter binding of Adr1, Adr1-dependent chromatin remodeling, and transcription when glucose is exhausted (Di Mauro et al. 2002; Young et al. 2002; Agricola et al. 2004, 2006; Biddick et al. 2008a,b). Glucose repression inhibits Adr1 activity by multiple mechanisms, including ADR1 expression (Blumberg et al. 1988), DNA binding (Sloan et al. 1999; Kacherovsky et al. 2008), and transcription activation (Cherry et al. 1989; Cook et al. 1994; Tachibana et al. 2007; Ratnakumar et al. 2009). PKA phosphorylates S230 in the regulatory domain in vitro (Denis et al. 1992), but its activity is not required for
phosphorylation (Ratnakumar et al. 2009) or glucose inhibition of Adr1 activity in vivo (Denis et al. 1992; Dombek et al. 1993).

Promoter binding of Adr1 is also inhibited by hypoacetylation of the histone tails (Verdone et al. 2002). Loss of histone deacetylase activity and consequent hyperacetylation of the histone tails overcomes glucose inhibition of promoter binding, chromatin remodeling, and preinitiation complex recruitment, but not transcription (Verdone et al. 2002; Tachibana et al. 2007). Activated Adr1 and activation of Snf1 stimulate the poised, inactive preinitiation complex and together allow complete escape from glucose repression (Tachibana et al. 2007).

Recent evidence suggests that yeast 14-3-3 (Bmh) proteins inhibit Adr1 activation function by binding to the S230-phosphorylated regulatory domain (Parua et al. 2010). 14-3-3 proteins are components of multiple signaling pathways and affect protein localization or activity by binding to phosphorylated domains (Yaffe 2002; Van Heusden and Steensma 2006). ADR1S alleles, such as ADR1-S230A, disrupt this inhibition as would be expected if phosphorylation were important for Bmh binding (Parua et al. 2010). As illustrated in Figure 2, Bmh has direct and indirect roles in Adr1 regulation, and both roles involve Snf1. Identifying the S230 kinase and the mechanism whereby Snf1 promotes its dephosphorylation (Ratnakumar et al. 2009) might link Adr1 activity to other nutrient-signaling pathways.

**Gal4 regulation by intermolecular AD masking**

The regulation of GAL gene expression and the description of the regulators involved constitute an important paradigm for eukaryotic gene regulation (Johnston and Carlson 1992). One focus of current research is the nature of promoter-bound but transcriptionally inactive Gal4 (Wightman et al. 2008; Jiang et al. 2009). In the absence of galactose, the C-terminal AD cannot recruit coactivators because it is occluded by Gal80 (Johnston et al. 1987; Ma and Ptashne 1987a). The sensor is Gal3, which binds and inactivates Gal80 when galactose is present. Genetic and structural analyses indicate that Gal4 and Gal3 bind to Gal80 at distinct but overlapping sites (Pilauri et al. 2005; Thoden et al. 2007). Thus, Gal3 holds the key to understanding how galactose is sensed by Gal80 and ultimately by Gal4.

Two opposing models have been proposed to explain Gal3 inhibition of Gal80 (Figure 3). A nondissociation model suggests that Gal3 binds Gal80 but does not remove it from Gal4 at the promoter (Johnston et al. 1987; Ma and Ptashne 1988; Chasman and Kornberg 1990; Parthun and Jaehning 1992; Platt and Reece 1998). Although this tripartite complex has been observed in vitro in the presence of galactose and ATP (Platt and Reece 1998), it has eluded detection in vivo. Recent fluorescence resonance energy transfer (FRET) experiments showed that a Gal80–Gal3 complex is present in both the cytoplasm and the nucleus after galactose induction; however, association of Hsc82 (Hsp90, green crescent), apparently without dissociation of the repressive chaperones (Lee and Zhang 2009).
model. Cytoplasmic Gal3 can sequester Gal80 away from the nuclear compartment in galactose-inducing conditions, suggesting that Gal80 is not bound to Gal4 at the promoter (Peng and Hopper 2000; Peng and Hopper 2002; Pilauri et al. 2005). Using live-cell imaging and a novel array of Gal4-binding sites, Gal80 was shown to rapidly dissociate from Gal4 in a Gal3, galactose-dependent manner (Jiang et al. 2009), suggesting that there is not a stable Gal3–Gal80–Gal4 complex at the promoter.

The GAL system has provided extraordinary insights into eukaryotic gene regulation and the evolution of a regulatory system. The galactose sensor Gal3 is a homolog of the Gal1 galactokinase (Bajwa et al. 1988; Suzuki-Fujimoto et al. 1996), and Gal80 is structurally related to a glucose–fructose oxidoreductase (Thoden et al. 2007). Thus, these two regulatory genes appear to have evolved from genes involved in intermediary metabolism. The GAL system also demonstrated how yeast genetics and molecular biology can be combined with biochemistry and structural biology to reveal fundamental mechanisms of eukaryotic transcription.

**Hap1 and Mal63 regulation by chaperones**

Hap1 and Mal63 are regulated by molecular chaperones as are steroid hormone receptors in mammalian cells (Pearl and Prodrumou 2006; Wandinger et al. 2008). Hap1, the major oxygen sensor in the yeast cell, is regulated by the Hsp70/Hsp90 chaperone machine as illustrated in Figure 4. In the absence of heme, repressive regions in the central part of Hap1 keep it in an inactive state (Hach et al. 1999). In response to heme, a biosynthetic precursor of active cytochromes, repression is relieved and genes in the Hap1 regulon are activated (Lee and Zhang 2009). Heme binds in the central regulatory region of Hap1 (Zhang et al. 1998; Hon et al. 2000), allowing transcription activation by the C-terminal AD.

Hap1 forms subcomplexes with Hsp70 and the co-chaperones Ydj1 and Sro9 that facilitate inhibition of Hap1 (Lan et al. 2004). One possibility is that heme binds Hap1 in the inhibited complex, alters its conformation such that it can bind Hsp90, which converts Hap1 to an activation-competent state (Hon et al. 2001). The nature of the functional activator, whether free or bound to Hsp90 and the other chaperones, is unresolved, as is the location of the inhibited complex when expressed at the endogenous level.

Mal63, the activator of the MAL genes, is also a client protein of the Hsp70/Hsp90 chaperone machine (Bali et al. 2003), but two different co-chaperones, Sti1 and Aha1, facilitate inhibition (Ran et al. 2010). In the presence of maltose, Hsp90 apparently displaces the inhibitors and converts Mal63 to an active form (Ran et al. 2008). The step(s) in transcription activation that are inhibited by the chaperone complex is unclear.

**Hsf1 regulation through conformational change**

Yeast Hsf1 is constitutively bound to some target genes but is inactive in the absence of stress (Jakobsen and Pelham 1988). Stress, such as increased temperature or oxidative damage, induces binding to new targets (Hahn et al. 2004) and activation of Hsf1 through two differentially acting ADs, one N- and one C-terminal to the central DBD (Nieto-Sotelo et al. 1990; Sorger 1990). The essential C-terminal AD modulates the response to high temperature (Nieto-Sotelo et al. 1990; Sorger 1990) and is indispensable for response to heat stress for genes containing a single heat shock element (Hashikawa et al. 2006). The N- and C-terminal ADs activate different sets of genes in the Hsf1 regulon during heat shock (Eastmond and Nelson 2006). Stress-induced activation is accompanied by Hsf1 hyperphosphorylation, but whether it is important for activation or for return to an inactive state is unclear (Sorger 1990; Jakobsen and Pelham 1991; Hoj and Jakobsen 1994; Hashikawa et al. 2006).

The sensor for both transient and sustained response appears to be the DNA-binding and oligomerization domains of Hsf1 on the basis of heterologous fusion proteins and other studies (Bonner et al. 1994, 2000a,b). Temperature-resistant mutations near the DBD enable Hsf1 lacking its C-terminal AD to activate transcription (Hashikawa et al. 2006). These suppressors implicate the DBD in mediating the response, as did single point mutations that constitutively enhanced transcription (Bulman et al. 2001). A conformational change accompanying activation can be detected by electrophoretic mobility shift in vitro and requires two trimers of Hsf1 bound to DNA (Bonner et al. 2000a; S. Lee et al. 2000), but determining whether a similar alteration occurs in vivo is challenging. Genetic evidence suggests that repression of DNA-bound Hsf1 may be mediated by chaperones (Duina et al. 1998; Bonner et al. 2000a). Thus, mutations in Hsf1 that cause constitutive activity could alter its interactions with inhibitory chaperones.

**Leu3 binding to a metabolite activates transcription**

Leu3 is a zinc-knuckle TF that acts as both repressor and activator of genes encoding enzymes involved in branched-
Chain amino acid biosynthesis (Kohlhaw 2003) The immediate signal for up-regulation of transcription is an intermediate in the pathway, α-isopropylmalate (α-IPM), that accumulates during leucine starvation. Leu3 is apparently DNA bound in the absence of α-IPM because it represses the low level of constitutive gene expression that occurs in the presence of leucine by an unknown mechanism. ChiP experiments are needed to confirm its direct action.

As illustrated in Figure 5, internal masking of the Leu3 AD, as opposed to recruitment of a corepressor, is favored by the Kohlhaw group primarily because Leu3 is regulated by α-IPM in both mammalian cells and mammalian cell-free extracts (Wang et al. 1997, 1999; Kohlhaw 2003). This interpretation assumes that corepressors are able to bind Leu3 are absent in mammalian cells. Genetic studies of mutant Leu3 activators are consistent with intramolecular as opposed to intermolecular interactions regulating the Leu3 AD (Friden et al. 1989; Wang et al. 1997, 1999). Other TFs that activate specific metabolic pathways (Arg81, Bas1, Lys14, Ppr1, Put3, War1) are also regulated by metabolites (Sellick and Reece 2005). In the case of Put3, direct binding of proline has been demonstrated (Sellick and Reece 2003).

**Zap1: AD regulation by Zn+2 binding**

Zap1 is the TF for a regulon that responds to limiting amounts of the essential but potentially toxic metal ion Zn+2 (Eide 2009). As shown in Figure 6, Zap1 has two ADs embedded in zinc regulatory regions that function independently of one another in response to the level of Zn+2 (Bird et al. 2000). Both ADs have been highly conserved in the Hemiascomycetes, indicating that both are important for Zap1 function. Mutation of potential Zn+2 ligands in AD1 blocks the inhibition of Zap1 activity that normally occurs in response to high Zn+2 levels (Herbig et al. 2005). Purified AD1 binds multiple Zn+2 atoms, consistent with a direct role for metal binding. AD2 consists of finger 2 of the seven canonical C2H2 zinc fingers, the last five of which constitute the DBD. When Cys and His residues in zinc finger 1 and zinc finger 2 are mutated to abolish Zn+2 binding, AD2 is constitutively active, suggesting that Zn+2 binding stabilizes the interaction of zinc fingers 1 and 2 and that this interaction is essential to inhibit Zap1 AD function (Wang et al. 2006). Genetic and biochemical studies suggest a model for the structure of this pair of fingers that implicates hydrophobic residues in Zap1 regulation (Eide 2009). The coactivator targets of the Zap1 ADs and the Zap1 residues involved in the interaction have not been determined.

In most of the examples described above, transcriptional activity is inhibited or activated by direct ligand binding. In some cases, the ligand is a protein (Gal4, Adr1, Hap1); in other cases, it is a small molecule, either a metabolite (Leu3, Put3) or a metal ion (Zap1). In the case of Hsf1, a conformational change may be ligand independent although this has yet to be directly demonstrated. It is assumed that the step in activation that is inhibited is recruitment of coactivators although this remains to be directly demonstrated. Structural studies showing how ligand binding influences the interaction of coactivators with ADs should lead to new insights into the mechanism of transcription activation.

**Core Promoter Architecture**

Regulatory signals from UAS and URS elements converge at the core promoter, the site where RNA Pol II and the general transcription factors assemble in the transcription preinitiation complex (PIC) before transcription initiation begins. The core promoter was first identified in mammalian gene regulatory regions and is defined as “the minimal DNA element required for basal transcription” (Smale and Kadonaga 2003). At least 60 bp of promoter DNA is occupied in the PIC, where nearly every base pair is in contact with Pol II and/or the general transcription factors (Douziech et al. 2000; Kim et al. 2000; Miller and Hahn 2006). Work with yeast has been especially important in defining different types of core promoters, the pathways of activator-stimulated PIC assembly, the structure of the PIC, and the role of chromatin structure in different classes of promoters.

Functional sequence elements in core promoters include the TATA element, initiator (INR), downstream promoter element (DPE), motif 10 element (MTE), and TFIIB recognition element (BRE) (Smale and Kadonaga 2003; Juven-Gershon and Kadonaga 2010). TATA is the recognition site for the general transcription factor TATA-binding protein (TBP), while INR, DPE, and MTE are recognition sites for the TBP-associated factors (TAF) subunits of the coactivator TFIID and BRE is a recognition site for the general factor TFIIB. All of these core promoter elements are short, degenerate, low-specificity elements. The combination of these elements varies among promoters and can determine activator and enhancer specificity.

Of these metazoan motifs, TATA is the only one clearly conserved in yeast (Basehoar et al. 2004; Sugihara et al. 2011). Since ~90% of yeast genes are TFIID dependent (Shen et al. 2003; Huisings and Pugh 2004), it is likely that yeast-specific TFIID recognition elements exist, but have not yet been recognized—perhaps because they are degenerate or have significantly diverged in sequence from their metazoan counterparts. Additional conserved yeast core promoter elements may be identified in future work. For example, a recent study found functionally redundant A- and T-rich sequences within the TATA-less RPS5 gene that
may be recognition sites for a component of the transcription machinery (Sugihara et al. 2011).

**Two promoter assembly pathways**

There are two pathways for assembly of the PIC that use the coactivators TFIID or SAGA (Kuras et al. 2000; Li et al. 2000; Bryant and Ptashne 2003; Qiu et al. 2004). These coactivators contact activators at UAS elements and are responsible for recruitment of TBP to promoters. Genome-wide analysis showed that TATA-containing promoters are primarily SAGA dependent, highly regulated, and generally stress responsive (Huisingsa and Pugh 2004). Only ~19% of yeast promoters contain TATA elements, and many of these (~10% of all yeast promoters) are dependent on SAGA coactivator function (Basehoar et al. 2004). In contrast, ~90% of yeast promoters are primarily TFIID dependent, are usually more constitutively active, and generally lack TATA elements. Efficient transcription requires coupling of compatible UAS and core promoters that favor nuclearosome-depleted regions, often found at the 5’ and 3’ end of genes, is noncoding transcripts that include both divergent and antisense RNAs (Seila et al. 2009). Several recent genome-wide studies have found that nearly 75% of the yeast genome is transcribed and that most noncoding transcripts seem to be the primary determinant of initiation (Runish et al. 1999; Herbig et al. 2010). A consequence of nucleosome-depleted regions, often found at the 5’ and 3’ end of genes, is noncoding transcripts that include both divergent and antisense RNAs (Seila et al. 2009). Several recent genome-wide studies have found that nearly 75% of the yeast genome is transcribed and that most noncoding transcripts initiate from nucleosome-depleted regions associated with the 5’ or 3’ end of genes (Nagalakshmi et al. 2008; Xu et al. 2009). Recent work has shown that several of these noncoding RNAs play important regulatory roles, and more examples of this are certain to emerge in future studies (Berretta and Morillon 2009).

An exception to the uniformly nucleosome-depleted regions is found at some regulated promoters where modulation of chromatin structure is part of the gene regulatory mechanism. For example, a positioned nucleosome at PHO5 must be removed prior to gene activation because it covers a binding site for the Pho4 activator (Almer et al. 1986; Fascher et al. 1990). Expression of genes containing nucleosomes positioned within the promoter are highly dependent on recruitment of chromatin-remodeling factors by transcription activators (Cairns 2009).

**Pol II Transcription Machinery and the Mechanism of Initiation**

Specific transcription initiation by eukaryotic and archaeal RNA polymerases requires the general transcription factors
(Hahn 2004; Thomas and Chiang 2006). These factors recognize core promoters, recruit RNA Pol into an active transcription initiation complex, and interact with coactivators and repressors to modulate transcription. Each of the three nuclear Pols (Pol I, II, and III) has its own set of general factors, sharing only TBP, which is required for transcription by all three enzymes. Many key advances in understanding the mechanism and regulation of Pol II and the general factors have been made in the yeast system. These include structure determination of Pol II in many different forms, the first isolation of genes encoding general transcription factors (TBP and TFIIA), models for assembly of the PIC, and using a combination of genetics and biochemistry to determine conserved mechanisms of gene regulation.

**RNA Pol II**

All multi-subunit RNA Pols are related in sequence and structure (Lane and Darst 2010a,b). Pol II is composed of 12 subunits, termed Rpb1–12, and numbered from largest to smallest (Cramer et al. 2008). All of these subunits except Rpb4 and Rpb9 are essential. Much of our detailed understanding about the mechanism of eukaryotic Pol II is derived from the groundbreaking structural work on yeast Pol II from the Kornberg and Cramer laboratories (Cramer et al. 2001, 2008; Gnatt et al. 2001). The structures of Pol II, either free or bound to other transcription factors or in various elongation complexes, form an invaluable framework for understanding mechanisms of gene regulation by all nuclear Pols (Figure 7A). The two largest Pol II subunits, Rpb1 and Rpb2, correspond to the β′- and β-subunits of bacterial Pol; together, these two subunits form the active site, the pore for entering nucleotide triphosphates, and the binding sites for DNA and the DNA–RNA hybrid in the transcription elongation complex. Rpb3 and Rpb11 correspond to the dimer of bacterial α-subunits, and Rpb6 corresponds to the bacterial ω-subunit, important for assembly and stability of bacterial Pol. The remainder of the Pol II subunits have no homology to bacterial subunits and are distributed around the surface of the enzyme where they perform roles in interaction with general factors, nucleic acids, and/or coactivators (Werner and Grohmann 2011). Eukaryotic Pols share 5 subunits (Rpb5, -6, -8, -10, -12) with 7 other Pol II subunits having sequence similarity with their Pol I and III counterparts. Pol I and III also have subunits that are similar to two of the Pol II general factors (Carter and Drouin 2010; Geiger et al. 2010), representing general factors that were stably incorporated into an ancestral form of Pol I and III (Werner and Grohmann 2011).

Pol II is unique among the multi-subunit Pols in containing a repeated seven-residue motif (YSPTSTS) at the C

**Figure 7** (A) RNA Pol II structure. Selected subunits are labeled. (B) Model for arrangement of TBP-TFIIA-TFIIA and TATA-DNA in the PIC. Red sphere, Zn. (C) Model for structure of the yeast Pol II PIC. The DNA template and non-template strands are blue and pink, and the base pair where DNA melting initiates is in dark blue and red; the TFIIF large and small subunits (Tfg1 and Tfg2) are orange and red; the TBP conserved domain (TBP) is green; TFIIA is yellow; and the large and small TFIIA subunits (Toa1 and Toa2) are brown and magenta. TFII dim, TFII dimerization domain; Tfg2 WH, Tfg2 winged helix domain; Zn ribbon, TFII N-terminal Zn ribbon domain. Dashed line represents a Tfg2 loop connecting the dimerization and winged helix domains. The 5′ and 3′ ends of the noncoding strand of promoter DNA are indicated.
terminus of Rpb1, termed the C-terminal domain (CTD) (Buratowski 2009). The unstructured CTD plays a role in assembly of the PIC, functionally interacting with the coactivator Mediator, and it is a target of several kinases that phosphorylate Ser at positions 2, 5, and 7 of the repeat. Pol II with a nonphosphorylated CTD is preferentially assembled in the PIC and then phosphorylated at Ser5 and -7 during initiation, principally by Kin28/Cdk7, a subunit of the general factor TFIIH (Feaver et al. 1994; Akhtar et al. 2009), although Srb10/Cdk9 can functionally substitute upon inhibition of Kin28 (Liu et al. 2004; Kanin et al. 2007). Upon transition of Pol II to the elongating mode, Ctk1 and Bur1, both related to the mammalian kinase Cdk9, phosphorylate the CTD at Ser2 (Buratowski 2009; Liu et al. 2009; Qiu et al. 2009; Zhou et al. 2009). The levels of CTD phosphorylation are precisely modulated during initiation, elongation, and termination, and these modifications function in regulating the association of many important factors with elongating Pol, including mRNA capping factors, chromatin modifiers, mRNA export, and transcription termination factors.

**General transcription factors and PIC assembly**

The human Pol II general factors were discovered as factors essential for specific transcription initiation from a TATA-containing core promoter (Matsui et al. 1980). A subset of the general factors recognize promoter DNA and form an asymmetric platform for the binding of Pol II and the incorporation of the remaining general factors (Buratowski et al. 1989). These factors include TBP, TFIIA, and TFIIIB, all of which form a complex with promoter DNA (Figure 7B). As described below, TFII B directly contacts Pol II and, along with TFII F, is critical for Pol II recruitment, initiation activity, and start site recognition. The remaining factors, TFII E and TFII H, play key roles in separation and stabilization of promoter DNA strands during transition of the transcription machinery into the active open complex state. Several major advances have been made in understanding the function of the general factors since recent extensive reviews (Hahn 2004; Thomas and Chiang 2006), and these are summarized below.

TFII B is a key component in assembly of the PIC, contacting TBP, DNA, and Pol II. TFII B consists of an N-terminal Zn ribbon domain that contacts the Pol II dock domain, an unstructured segment termed the B-reader and linker regions, and two cyclin folds that form the TFII B core domain, binding TBP and DNA on either side of the TATA (Hahn 2004; Thomas and Chiang 2006). Positioning of TFII B on Pol II has been visualized by site-directed crosslinking and targeted hydroxyl radical cleavage assays and by several X-ray structures of the Pol II–TFII B complex (Chen and Hahn 2003, 2004; Bushnell et al. 2004; Kostrewa et al. 2009; Liu et al. 2010). Two recent structures of Pol II–TFII B show the B-reader and linker regions near two critical regions in the Pol II active site. In the complex, a helical portion of the B-reader is in position to interact with single-stranded DNA in the open complex state and is proposed to “read” the DNA sequence, contributing to start site recognition (Kostrewa et al. 2009). A helical portion of the B-linker region in the complex is positioned near the presumed site of DNA strand unwinding and proposed to contribute to DNA melting (Kostrewa et al. 2009; Liu et al. 2010). As first shown by site-specific biochemical probes (Chen and Hahn 2004; Chen et al. 2007), the N-terminal cyclin fold (core) of TFII B binds the Pol II wall domain where it functions to position the TBP–DNA complex over the Pol II active site cleft (Figure 7C). This positioning is the key to setting the architecture of the PIC and to positioning promoter DNA directly over the Pol II cleft (Miller and Hahn 2006).

Yeast TFII F is composed of two conserved subunits, Tfg1 and Tfg2, as well as a third nonconserved subunit that is a component of several other complexes involved in gene regulation (Henry et al. 1994). TFII F is involved in stabilizing Pol II in the PIC, setting the transcription start site, stabilizing the RNA–DNA hybrid in early elongation complexes, and stimulating transcription elongation in vitro (Hahn 2004; Thomas and Chiang 2006). Two structured domains of TFII F, the dimerization domain and the Tfg2 winged helix domain, are located on two separate sites on the Rpb2 surface above the active site cleft (Figure 7C) (Eichner et al. 2010; A. Chen et al. 2010). The winged helix domain is in position to contact DNA upstream of the TATA and bend it over the top of Pol II, possibly contributing to stabilization of the PIC. These two TFII F domains are connected by an essential nonstructured linker that may play a direct role in initiation (Eichner et al. 2010).

**Open complex formation and transcription initiation**

Transition of the PIC into the open complex involves a dramatic conformational change (Murakami and Darst 2003) requiring insertion of double-stranded promoter DNA into the jaw and downstream cleft of Pol II, the TFII H helicase-dependent separation of DNA strands surrounding the transcription start site from ~ −9 to +1 (with respect to the transcription start) (Wang et al. 1992; Holstege et al. 1997; Reyvak et al. 2004), and insertion of the single-stranded DNA template strand into the active site of Pol II. This step can be highly regulated in bacteria and the mechanism of open complex formation is one of the major unanswered questions for all multi-subunit RNA polymerases (Figure 8).

Recently, the yeast PIC structure model was merged with the structure of yeast-elongating Pol II to generate the first
structural model of the open complex (Kostrewa et al. 2009; Liu et al. 2010). In this model, there are 14 unwound bases with the TFIIB B-reader segment interacting directly with single-stranded DNA upstream of the transcription start site (Kostrewa et al. 2009). Future advances in understanding the open complex will emerge from structural and biochemical studies examining other general factors situated within the enzyme active site and determining how they interact with each other and DNA to promote start site selection, initiation, and the initial steps of elongation.

Transcription start site scanning

Yeast Pol II scans downstream sequences for a suitable transcription start (Giardina and Lis 1993; Kuehner and Brow 2006; Steinmetz et al. 2006). Consistent with this, DNA in vivo at the GAL1 and GAL10 promoters is unwound from ~20 bp downstream of TATA (the approximate site of initial strand unwinding for metazoan Pol II) through the transcription start ~90 bp distant from the TATA (Giardina and Lis 1993). This scanning mechanism does not involve transcription of the DNA between the TATA and initiation site (Khaperskyy et al. 2008). Since the cost of disrupting a DNA base pair is ~2 kcal/mol, there is a significant energetic cost of unwinding 10–70 bp. A reasonable model to explain these findings (Miller and Hahn 2006) is that the yeast-scanning mechanism involves strand unwinding and DNA translocation promoted by the TFIIF Rad25/XPB helicase using the energy of ATP hydrolysis (Figure 8). There is no evidence yet to determine whether the DNA between the initial melting site and the transcription start site is unwound all together or whether an ~10-bp bubble translocates downstream. Answering this question will likely require single molecule studies of the initiation reaction.

Pioneering genetic experiments by Hampsey, as well as later work, suggested that TFIIB, TFIIF, and Pol II are all involved in start site selection since mutations in any of these factors can alter the normal start site distribution (Hampsey 1998; Faitar et al. 2001; Ghazy et al. 2004; Chen et al. 2007). Mutations in both the TFIIB B-reader and the switch 2 segment of Rpb1 cause transcription to start farther away from TATA but still at sequences matching the initiator consensus (Faitar et al. 2001; Kostrewa et al. 2009). These mutations act as though they decrease the efficiency of initiator recognition. In contrast, mutations in the TFIIF dimerization domain or in two subunits of Pol II that interact with TFIIF, Rpb2, and Rpb9, start transcription closer to TATA than in wild-type cells and behave as though they have relaxed specificity for initiator recognition (Khaperskyy et al. 2008). The study of this mechanistic step in S. cerevisiae will undoubtedly reveal important details about start site selection in other eukaryotes.

Transcription Coactivators

Transcription activation is one of the most important mechanisms for control of gene regulation and is a common endpoint for many signaling pathways, including those controlling cell growth and the response to environmental or metabolic stress. The principal targets of activators are coactivators, large protein complexes that enhance activated transcription by direct contact with the basal transcription machinery and/or by chromatin modification. These two activities cooperate to stimulate PIC formation, leading to increased transcription. Although much progress has been made in defining coactivators and their mechanism of action, there is much to be learned about the architecture of coactivator complexes and how they interact with the transcription machinery and integrate signals to modulate transcription. The coactivators discussed below are conserved in eukaryotes, and several of the coactivators (Mediator, SAGA, NuA4) were first discovered in yeast. The yeast system has often led the way in understanding the nature and mechanism of coactivators and how these complexes function in other eukaryotes.

Mediator

Mediator is a 25-subunit complex that functions as an intermediate between transcription regulators and the general transcription factors (Biddick and Young 2005; Bjorklund and Gustafsson 2005). All yeast Mediator subunits have homologs in insects and mammals, and a common nomenclature has been developed for Mediator subunits (Bourbon et al. 2004; Bourbon 2008). A set of 17 mediator subunits is conserved in all eukaryotes and forms a core for assembly of other organism-specific subunits (Bourbon 2008). Human Mediator is more complex than its yeast counterpart, containing additional subunits, and exists in multiple forms with variable subunit composition (Conaway et al. 2005).

Mediator binds transcription activation domains and Pol II, allowing activator-dependent Pol II recruitment (Bjorklund and Gustafsson 2005; Malik and Roeder 2005), but its role in gene regulation is much more complex than simply linking activators and polymerase. Mediator also stimulates basal transcription, at least in part, by stabilization of the PIC and by stimulation of TFIIF-dependent Pol II CTD phosphorylation. Mediator can positively or negatively affect transcription and seems to cooperatively interact with other coactivators (Bryant and Ptashne 2003; X. Liu et al. 2008). Finally, it has been proposed that Mediator may be a direct target of signaling pathways, although this remains to be firmly established (Malik and Roeder 2005; Taatjes 2010). Because of these diverse roles, Mediator is thought to be a major target of transcriptional regulatory signals that are integrated and transmitted to the transcription machinery in a promoter and gene-specific manner.

Mediator was discovered in yeast in a classic example of biochemistry and genetics converging to identify the same factor, with the sum of functional insights greater than could be revealed by either approach alone. In a biochemical tour de force, Roger Kornberg’s laboratory developed a yeast
basal transcription system using purified general factors and Pol II that did not respond to activators. A factor that they termed Mediator was purified that stimulated basal transcription and allowed stimulation by the activators Gal4-VP16 and Gcn4 (Flanagan et al. 1991). At the same time, in an elegant series of genetic experiments, Rick Young’s lab isolated suppressors of cold-sensitive yeast mutants containing a shortened Pol II CTD. Young’s group found that these Srb proteins (Suppressor of RNA polymerase B) copurified with Pol II and were identical to some of the Mediator subunits (Thompson et al. 1993). Further comparison of the Mediator subunits with other regulatory genes showed that many of the Mediator subunits had been identified in genetic screens for defects in specific regulatory pathways [Table 1 in Biddick and Young 2005; (Bjorklund and Gustafsson 2005)].

**Structural organization of Mediator**

Mediator can be subdivided into four distinct modules: head, middle, tail, and kinase (Figure 9) (Davis et al. 2002; Cai et al. 2009). Extensive biochemical, genetic, and two-hybrid studies have mapped protein–protein interactions and the four modules within Mediator (Kang et al. 2001; Guglielmi et al. 2004). Electron microscopy revealed that the modules form separate domains that are highly flexible with respect to each other (Davis et al. 2002; Cai et al. 2009). Well-diffracting crystals from Mediator modules have not yet been obtained because of the flexibility of Mediator (Toth-Petroczy et al. 2008). Sequence analysis predicts that many Mediator subunits are composed of protein–protein interaction motifs connected by large inherently unstructured segments. To date, X-ray and NMR structures have been determined for five relatively small and well-ordered Mediator domains (Baumli et al. 2005; Hoeppner et al. 2005; Lariviere et al. 2006; Koschubs et al. 2009; Thakur et al. 2009).

The head module can be reconstituted upon coexpression of its seven individual subunits, Med6, -8, -11, -17, -18, -20, and -22 (Takagi et al. 2006; Lariviere et al. 2008; Cai et al. 2010). The N terminus of Med8 is reported to bind TBP (Lariviere et al. 2006), but it has not yet been shown if this interaction occurs in functional transcription complexes. From electron microscopy (EM) studies, the head module seems to closely interact with the back surface of Pol II, closely approaching the subunits Rpb3/11 and the protruding Rpb4/7 subunits that modulate closing of the Pol II cleft (Davis et al. 2002; Cai et al. 2010). Recent studies suggest that the head module binds weakly to a minimal PIC consisting of Pol II, TFII, TFIIb, TBP, and promoter DNA (Takagi et al. 2006). Mediator does not appear to bind directly to the Pol II CTD, so the molecular basis by which Srb mutations were originally isolated is still unclear; most of the Srb subunits are located within the head module.

The middle module is composed of eight to nine subunits with Med14/Rgr1 connecting the middle and tail modules. A seven-subunit recombinant complex lacking Med19 and Med14 was analyzed using protein biochemistry (Koschubs et al. 2010), revealing that the middle module is elongated and highly flexible. Part of this is likely due to the Med7/21 interface composed of a four-helix bundle and a flexible protrusion connected by a flexible linker (Baumli et al. 2005). This structure is very elongated, nearly one-third the length of Mediator, and the linker may contribute to Mediator conformational changes upon binding Pol II.

The tail module is composed of four to five subunits and is a target of at least several transcriptional activators (Kang et al. 2001). The subunits Med2, -3, and -15 (Gal11) form a submodule that can be recruited in vivo to a DNA-bound activator in a strain containing a MED16/SIN4 deletion (Zhang et al. 2004). Mutation of any one of these subunits disrupts this submodule. The best-characterized tail subunit is Med15/Gal11, containing four N-terminal activator-binding domains separated by glutamine-asparagine-rich flexible linkers (Herbig et al. 2010; Jedidi et al. 2010).

The kinase module is composed of four subunits including Cdk8 and cyclinC (Liao et al. 1995). This module has both positive and negative effects on expression (Bjorklund and Gustafsson 2005; van de Peppel et al. 2005; Taatjes 2010). Early work showed that Cdk8 levels are reduced as cells reach stationary phase, leading to expression of genes induced by nutrient deprivation (Holstege et al. 1998). Studies with the human Mediator suggest that the kinase module inhibits a Mediator conformational change that opens up a pocket containing the Pol II-binding surface (Taatjes 2010). Conversely, it was demonstrated that Cdk8 can act positively to promote initiation and CTD phosphorylation in the absence of
Cdk7 kinase activity (Liu et al. 2004), and studies in human cells showed that Cdk8 activity stimulates transcription of genes in the serum response network (Donner et al. 2010).

**Mediator targets**

Mediator gene targets are still somewhat controversial. Pioneering genome-wide studies using a temperature-sensitive MED17 allele showed that transcription of nearly all Pol II-transcribed genes was rapidly shut down upon heat shock (Holstege et al. 1998), a powerful argument for direct action of Mediator at all Pol II genes. Genome-wide ChIP studies later reported that Mediator is located in control regions of nearly all Pol II-transcribed genes (Andrau et al. 2006). This finding was challenged by results suggesting that Mediator nearly always crosslinks to the UAS element rather than to the promoter, where it is expected to interact with Pol II. These results may be explained in part by inefficient crosslinking of Mediator to promoter DNA, a result expected if Mediator is located several interactions away from proteins that directly bind promoter DNA.

In contrast to the model that Mediator is a simple link between activators and Pol II, mutations in the nonessential Mediator subunits result in both positive and negative effects on mRNA expression. This was first observed when a mutation in MED16/RGR1 was found to increase expression of HO in the absence of its activator Swi5 (Stillman et al. 1994). These effects were examined systematically using microarray analysis to study genome-wide changes in mRNA levels upon deletion of the nonessential Mediator subunits (Holstege et al. 1998; van de Peppel et al. 2005). Although mutation of each of these subunits has both positive and negative effects, elimination of some subunits predominantly increases gene expression while elimination of others generally decreases expression. For example, deletion of MED19 or kinase module subunits results in increased expression from a large set of genes. Conversely, mutation of tail or head subunits results in predominantly decreased gene expression. Mutation of middle module subunits falls somewhere in between. Although not yet understood, these complex phenotypes may be due to changes in Mediator that preclude or enhance interaction with specific transcription factors or the general transcription machinery in a promoter-specific fashion. Understanding these mechanisms in more detail will greatly help in understanding how Mediator integrates inputs from various signaling pathways and why Mediator is so complex.

**TFIID**

TFIID is composed of TBP and 14 Tafs (Matangkasombut et al. 2004; Cler et al. 2009). Thirteen of the yeast TAFs are conserved in eukaryotes, with one yeast-specific Taf (Taf14) a subunit of other multi-subunit complexes involved in transcriptional regulation such as TFIIF, INO80, and SWI/SNF (Hahn 2004). The strong sequence conservation of Tafs has allowed a unified Taf nomenclature for all eukaryotes (Tafs 1–14) (Tora 2002). Although TBP is sufficient to promote basal transcription from a TATA-containing promoter when combined with purified Pol II and the other general transcription factors, transcription from TATA-less promoters and, in many cases, response to activators requires the TFIID complex.

The subunit composition of TFIID was first revealed in human and Drosophila cells, where TBP is tightly associated with Tafs (Dynlacht et al. 1991; Kokubo et al. 1993). Yeast TBP is not as stably bound to the Tafs compared to other eukaryotes, explaining why yeast TBP was originally purified as a single polypeptide (Buratowski et al. 1988). Yeast Tafs were isolated later using streamlined and gentler purification methods such as TBP affinity columns or immune purification of TBP (Reese et al. 1994; Poon et al. 1995). Depletion of yeast Tafs from whole-cell extracts abolished activation by the strong heterologous activator Gal4-VP16, and purified yeast TFIID allowed modest transcription stimulation by activators using purified general factors and Pol II (Reese et al. 1994; Poon et al. 1995). Yeast Tafs are encoded by single-copy genes, while several human and Drosophila Tafs are encoded by multiple genes. In these more complex systems, one Taf allele is typically expressed only in specific cell types, contributing to tissue and developmental-specific gene expression (D’Alessio et al. 2009).

**TFIID structure and TBP-DNA binding**

Nine of the 13 conserved Taf subunits contain histone fold domains (HFDs), and biochemical and structural analysis showed that these domains are used for dimerization of specific Taf pairs: Tafs 6–9, 11–13, 8–10, 3–10, and 4–12 (Cler et al. 2009). These Tafs are all thought to be present in at least two copies per TFIID complex. The coactivator SAGA
contains a related subset of these HFD Tafs except that Taf4–12 has been replaced by Ada1-Taf12; Taf13–11 has been replaced by Spt3, which contains two HFDs; and Taf3–10 has been replaced by Spt7-Taf10. Whether these HFD Tafs form related structural modules in TFIID and SAGA is not yet known.

EM studies showed that TFIID is a large flexible complex and that human and yeast TFIIDs have very similar structures (Andel et al. 1999; Leurent et al. 2002). TFIID is composed of three linked lobes termed A, B, and C, which form a "horseshoe"-shaped structure (Figure 10) (W. L. Liu et al. 2008; Papai et al. 2009). TFIID is assembled around two molecules of the WD-40 repeat-containing Taf5, which binds the other Taf HFD pairs to form a core crescent-shaped structure consisting of TFIID lobes B, C, and the lower part of lobe A (Cler et al. 2009). Although this subunit arrangement predicts a symmetric core domain built around the Taf5 dimer, localization studies and biochemical analysis suggest that the core is only pseudosymmetric.

Tafs 1, -2, and -7, along with TBP, bind to one lobe of the core domain, forming the complete TFIID complex. In this arrangement, TBP is positioned within the center of the horseshoe-shaped cleft. It is not yet clear how or if TBP interacts with DNA when TFIID binds DNA. TBP binding to TATA requires that TBP open the DNA minor groove, interacting closely with the hydrophobic surface of the groove through a complementary hydrophobic surface on the underside of TBP (J. L. Kim et al. 1993; Y. Kim 1993). Due to steric constraints, this binding mechanism is compatible only with the TATA sequence, and substitution of G-C at key positions within TATA causes a severe decrease in DNA-binding affinity (Patikoglou et al. 1999). This raises the question of how TFIID and TBP specifically interact with promoter DNA at TATA-less promoters. Initial in vitro experiments showed that a human TBP mutation that compromised TBP-TATA binding decreased TBP-driven transcription but not TFIID-driven transcription from a TATA-less promoter (Martinez et al. 1995). This suggests that if TBP interacts with DNA at a TATA-less promoter, then it does so by a different mechanism compared to TBP-TATA binding.

An additional complication for TBP-DNA binding within TFIID is due to the TAND (Taf1 N-terminal) domain of Taf1. TAND tightly binds the DNA-binding surface of TBP and is likely a major contributor to the stability of TBP within the TFIID complex (Liu et al. 1998), but the TAND must be removed from TBP to allow specific DNA binding. Repression by the TAND domain can be overcome by TFIIA, which competes with TAND for TBP binding (Kokubo et al. 1998). In one study, TFIID binding to the RPS5 promoter was observed only in the presence of TFIIA (Sanders et al. 2002), consistent with the competition of TFIIA and the TAND domain for TBP.

Higher eukaryotic Taf1 contains several domains with enzymatic activity: separate N- and C-terminal protein kinase activities, histone acetyl transferase activity, and double bromo and PhD domains, the latter two interacting with acetylated and methylated histones, respectively (Matangkasombut et al. 2004). Yeast Taf1 was reported to have histone acetyl transferases (HAT) activity, although much weaker than its human counterpart (Mizzen et al. 1996), and it is not clear whether this HAT activity is functionally important in yeast. The protein kinase activities and PhD domains are not conserved between human and yeast Taf1. A yeast factor containing a double bromodomain, Bdf1, is associated with TFIID and likely represents the acetyl lysine-binding activity functionally analogous to human Taf1 (Matangkasombut et al. 2000). It is thought that interaction of the Taf1 bromodomain with doubly acetylated histone H4 flanking the promoter stabilizes TFIID binding, contributing to the stability of the PIC.

Role of Tafs in gene expression

The yeast system has led the way in understanding the in vivo role of Tafs. Experiments in yeast have shown that all of the conserved Tafs are essential for growth (Reese et al. 1994; Poon et al. 1995). To examine the role of Tafs in expression, several Tafs were initially depleted from growing cells using temperature-sensitive mutations, expression of Tafs from repressible promoters, and/or controlled protein degradation. In surprising contrast to the prevailing view at the time, most genes assayed were expressed normally upon Taf depletion (Moqtaderi et al. 1996; Walker et al. 1996). These studies continued over the next several years until microarrays were used to systematically assay genome-wide expression upon depletion of all 14 individual Tafs (Shen et al. 2003).

Analysis of genome-wide Taf function revealed that there are two types of Pol II promoters: TFIID dependent and TFIID independent (Kuras et al. 2000; Li et al. 2000; Basehoar et al. 2004; Huisinga and Pugh 2004). In cells grown in rich media, 84% of the genes assayed required at least one Taf, and 16% of genes were completely Taf independent. TFIID and SAGA share a subset of Tafs, and the most broadly required Tafs are those shared by both coactivator complexes (Shen et al. 2003).

Some of the first ChIP studies were designed to examine the mechanism of TBP recruitment and Taf dependence. At most promoters, there is an excellent correlation between TBP binding and transcription, with most nonexpressed genes showing low levels of TBP crosslinked to promoters and an increase in TBP crosslinking upon gene activation (Kuras and Struhl 1999; Li et al. 1999). The Taf dependence of promoters also corresponds well to whether Tafs are recruited to promoters. TFIID Taf-dependent promoters have a high ratio of crosslinked TFIID Taf/TBP while TFIID Taf-independent promoters have a much lower ratio (Kuras et al. 2000; Li et al. 2000). An unexpected finding from these studies was that upon TBP depletion, TFIID Tafs were still crosslinked to regulatory regions upon gene induction, although recruitment of the remaining general factors and Pol II was abolished (Li et al. 2000). This result suggests that the TFIID Tafs are targeted by activator, which agrees with results from humans, Drosophila, and yeast that activators can interact with TFIID subunits.
Several studies have examined which promoter elements are involved in determining TFIID dependence by swapping UAS and core promoters from TFIID-dependent or -independent genes (Shen and Green 1997; Cheng et al. 2002; Li et al. 2002). These experiments have sometimes given conflicting results with TFIID Taf dependence tracking with either the UAS or the core promoter. The most recent and comprehensive study examined the RPS5 (TFIID-dependent) and GAL1 and ADH1 (TFIID-independent) promoters (Li et al. 2002). Measurement of transcription and TFIID recruitment at these chimeric promoters showed that efficient transcription requires compatible UAS and core promoter elements. For example, a TFIID-independent UAS, e.g., GAL1, works much better with a TFIID-independent promoter. This is presumably because Gal4 does not efficiently target TFIID, but rather the SAGA coactivator. In contrast, the RPS5 UAS could activate, although with a lowered efficiency, at a TFIID-independent core promoter.

SAGA

The coactivator SAGA modulates expression of many inducible genes that are generally distinct from genes regulated by TFIID (T. I. Lee et al. 2000; Huisinga and Pugh 2004). SAGA mutations alter expression of ~10% of yeast genes, and these are usually TATA-containing, stress-regulated, and highly inducible. SAGA is targeted by gene-specific activators and functions through covalent modification of chromatin and by direct contact and recruitment of TBP. SAGA is a 1.8-mDa complex composed of 20 subunits (Baker and Grant 2007), is conserved in eukaryotes, and is orthologous to the mammalian coactivators TFTC, PCAF, and STAGA (Lee and Workman 2007). Genetic and biochemical studies implicate the sub-units Ada1, -2, -3) were found in a genetic screen discovered in a screen for suppression of Ty element insertions (Winston et al. 1984; Eisenmann et al. 1994). Several other subunits (Ada1, -2, -3) were found in a genetic screen for genes, which, when mutated, suppressed the toxic effects of an overexpressed transcription activator (Berger et al. 1992). A complex containing these and additional subunits was found in biochemical studies aimed at isolating HATs in yeast whole-cell extracts (Grant et al. 1997). Many of the yeast HAT complexes bind Ni Sepharose, and fractionation of these Ni-purified HATs led to isolation of the coactivator complexes SAGA and NuA4 containing the HATs Gcn5 and Esa1, respectively. SAGA preferentially acetylates histone H3 while NuA4 acetylates histone H4.

SAGA organization and TBP binding

Biochemical, genetic, structural, and recent mass spectrometry analysis has shown that SAGA is composed of at least five modules: (1) a core/Spt module important for SAGA integrity and TBP-binding containing Spt7, Ada1, Spt20, and the TBP-binding subunits Spt3 and Spt8 (Eisenmann et al. 1992; Grant et al. 1997, 1998a; Lee and Young 1998; Dudley et al. 1999; Sterner et al. 1999; Bhaumik and Green 2002; Wu and Winston 2002; Mohibullah and Hahn 2008; Lee et al. 2011); (2) a HAT module containing the HAT Gen5, Sgf29, and the subunits Ada2 and Ada3, which modulate substrate specificity (Grant et al. 1999; Balasubramanian et al. 2002; M. Washburn and J. Workman, personal communication); (3) a four-subunit ubiquitin protease module (Rodriguez-Navarro et al. 2004; Ingvarsdottir et al. 2005; Lee et al. 2005; Shukla et al. 2006; Kohler et al. 2008); (4) Tra1, an activator-binding module (Brown et al. 2001); and (5) a Taf module (Lee et al. 2011). The subunits Ada1 and Taf12 are thought to dimerize via histone fold domains, and SAGA appears to contain two molecules of Ada1 located within distinct regions of the complex. EM studies suggest that most of the modules are located at separate positions within SAGA (Wu et al. 2004). The distribution of the Tra1, HAT, and TBP-binding modules to different regions suggests that SAGA is composed of submodules that physically associate via a central core. The SLIK complex regulates expression from a small set of genes and is closely related to SAGA (Pray-Grant et al. 2002). SLIK contains a proteolytically processed Spt7 subunit and the subunit Rtg1 in place of Spt8 (Wu and Winston 2002).

Like Mediator, SAGA has a complex role in gene regulation since mutation of SAGA subunits can either increase or decrease gene expression. For example, genome-wide characterization of Gcn5-responsive genes reveals that Gcn5 functions as both a coactivator and a corepressor in S. cerevisiae and Schizosaccharomyces pombe (Xue-Franzen et al. 2010). Gene expression studies have also shown that Gcn5 and subunits of the SAGA/TBP-binding module have opposing roles; S. cerevisiae SPT3 and GCN5 mutations were found to have opposite effects in transcriptional regulation of the HO and STE11 genes (Yu et al. 2003; Helmlinger et al. 2008), although it is possible that some of these phenotypes are due to indirect effects. SAGA can also repress basal expression of some genes in vitro and in vivo (Belotserkovskaya et al. 2000; Warfield et al. 2004).

Although both TFIID and SAGA share a subset of Tafs and function to recruit TBP, they have very different biochemical activities in TBP binding. TFIID stably binds TBP under certain conditions, while purified SAGA contains little TBP (Sterner et al. 1999; Sanders et al. 2002; Laprade et al. 2007). Genetic and biochemical studies implicate the subunits Spt3 and Spt8 in TBP binding. These Spt subunits genetically interact with TBP, and mutations in Spt3 can suppress TBP mutations (Eisenmann et al. 1992, 1994; Laprade et al. 2007). In one case, an Spt3 mutation was found to dramatically increase the amount of TBP copurifying with SAGA (Laprade et al. 2007). Structure modeling strongly suggests that the N- and C-terminal ends of Spt3 interact via a noncanonical histone fold domain, homologous to the Taf 11–13 structure (Birck et al. 1998). Mutations that result in an Spt3 phenotype lie primarily along one face of this model. (Eisenmann et al. 1992; Laprade et al. 2007). These
mutations are presumably altered in their interaction with TBP because several can be suppressed in an allele-specific fashion by mutations in TBP. However, the interaction of SAGA and TBP is complex because mutations in Spt3 and Spt8 can have both positive and negative effects on gene expression (Yu et al. 2003; Helmlinger et al. 2008). Both Spt3 and Spt8 crosslink to TBP in PICs (Mohibullah and Hahn 2008), and mutations in Spt3 and Spt8 show defects in TBP recruitment in vivo (Dudley et al. 1999; Larschan and Winston 2001; Bbaumik and Green 2002; Barbaric et al. 2003; Mohibullah and Hahn 2008). Supporting the direct interaction of Spt3 and TBP is the finding that the site of TBP-Spt3 crosslinking lies very close to TBP mutations that suppress mutations in Spt3. However, the mechanism of TBP-Spt3 binding is not well understood because purified TBP and Spt3 do not interact (Madison and Winston 1997; Sterner et al. 1999; Sermwittayawong and Tan 2006).

**Tra1 has multiple functions within SAGA and NuA4**

The largest SAGA subunit is Tra1, a nearly 4000-residue essential protein that is shared with the NuA4 complex (Grant et al. 1998b; Saleh et al. 1998; Allard et al. 1999). Tra1 and its human homolog, TRRAP, are members of the P13-related protein kinase family, but Tra1 and TRRAP have specifically lost kinase activity (Mutiu et al. 2007). Biochemical and genetic experiments showed that several activators, including Gcn4 and Gal4, interact with Tra1 and that this interaction is important for activated transcription of SAGA-dependent genes (Brown et al. 2001; Fishburn et al. 2005; Reeves and Hahn 2005). Tra1 is also likely responsible for activator recruitment of the NuA4 coactivator. Most of Tra1 is composed of short repeated motifs, including HEAT and TPR repeats, which are N-terminal to the P13 kinase-like domain (Knutson and Hahn 2011). Systematic mutation of Tra1 has shown that about two-thirds of the protein is essential for growth. All lethal Tra1 mutations isolated to date abolish association of both SAGA and NuA4 subunits, showing that Tra1 uses identical regions to contact the subunits of both complexes. Nonlethal Tra1 mutations fall into three classes: (1) defective in activator-dependent promoter recruitment, (2) defective in HAT module recruitment, and (3) normal for HAT recruitment but defective for in vivo HAT activity. These latter two categories show that Tra1 is important for stability of the NuA4 and SAGA HAT modules as well as somehow being involved in activity or specificity of the HAT.

Although there has been significant progress on understanding function, there is not much known about how SAGA is regulated, organized, and interfaces with activators and the transcription machinery. Important areas for future work are (1) understanding the architecture of SAGA and the organization and functional relationships of the different SAGA modules; (2) understanding the mechanism of how SAGA interfaces with TBP, whether this binding is regulated, and if SAGA-TBP interaction is limiting at promoters; and (3) the mechanism of activator–Tra1 interaction.

**Transcription Activation Mechanisms**

Over the past 30 years, the yeast system has been used to answer a number of fundamental questions that are important for understanding gene control in all eukaryotes: (1) What mechanisms result in transcription stimulation? (2) How is formation of transcription preinitiation complexes regulated and what are the roles of coactivators in this process? (3) What is the nature of activation domains and what are their direct targets? and (4) How do activators interact with targets and are these interactions specific? Yeast has been an especially powerful system to dissect gene regulatory mechanisms and, in many cases, has led the way in understanding the fundamental mechanisms of transcriptional regulation.

In theory, transcription could be modulated by a number of different mechanisms (Hahn 1998; Keaveney and Struhl 1998) including (1) recruitment of coactivators and general transcription factors to promoters, (2) conformational changes in the transcription machinery leading to increased activity, (3) modification of chromatin structure by ATP-dependent remodelers or through covalent nucleosome modifications, and (4) by enhancing steps that occur after preinitiation complex formation. Each of these steps plays a role in eukaryotic regulation, although it is not yet clear if all of these mechanisms are used in yeast.

**Activation by recruitment**

Activation by recruitment is one of the best-studied regulatory mechanisms (Ptashne and Gann 2002), and there is overwhelming evidence that this is a major, but not the only, means of transcription stimulation in eukaryotes. Early support for the recruitment model was the finding in yeast of activation by "artificial recruitment" (Chatterjee and Struhl 1995; Klages and Strubin 1995; Xiao et al. 1995), where targeting an appropriate coactivator subunit or general transcription factor to a promoter by fusion to a DBD dramatically enhances transcription. The first artificial recruitment study enhanced transcription by fusion of TBP to the LexA DBD. These studies have been successfully repeated using Taf subunits, Mediator and SAGA subunits, and TFIIB (Gonzalez-Couto et al. 1997; Keaveney and Struhl 1998). Taken together, the artificial recruitment studies show that, under some circumstances, transcription can be enhanced by direct transcription of the transcription machinery, but do not by themselves prove that natural activators work by recruitment. Similar conclusions were reached in a genetic study where a mutation (Gal11p) in the Mediator subunit Gal11/Med15 was found that created a new protein–protein interaction with the Gal4 dimerization domain and stimulated transcription in the absence of a Gal4 activation domain (Barberis et al. 1995; Hidalgo et al. 2001).

Interpretation of artificial recruitment experiments is not always straightforward because the ability of these protein fusions to activate is very dependent on the architecture of the reporter and whether the reporter is located on a plasmid.
or integrated into the chromosome, on the precise way in which the protein fusion is constructed, and on whether the fusion is highly overexpressed (Gaudreau et al. 1999; Cheng et al. 2004; Wang et al. 2010). A recent study found that fusion of a DBD to any of the three Mediator tail subunits could activate transcription (Wang et al. 2010). One consistent conclusion from these experiments is that the Mediator tail module, and Gal11/Med15 in particular, is an especially good target for stimulation by artificial recruitment. It is probably not a coincidence that Gal11/Med15 is a common target of several activation domains. Part of the variability in artificial recruitment experiments may be due to the absence of targeted chromatin modification and the differential requirement of promoters for chromatin remodeling. For example, Morse and colleagues have demonstrated that artificial recruitment does not activate transcription at promoters where a nucleosome blocks access to the promoter (Ryan et al. 2000).

The most conclusive evidence for the recruitment model is studies in yeast using ChIP to examine the level of factors at gene regulatory regions before and after gene activation (Kuras and Struhl 1999; Li et al. 1999). There are now numerous examples showing that the level of coactivators, chromatin remodelers, and the general transcription factors crosslinking to UAS and promoter regions significantly increase upon transcription stimulation (Green 2005; Weake and Workman 2010). Although there are a few exceptions to this finding, it seems like the recruitment mechanism is involved in transcription activation at nearly all Pol II-transcribed genes.

**Other activation mechanisms**

Activator-induced conformational change is another mechanism that has been proposed to contribute to transcription stimulation (Taatjes et al. 2002). The best evidence for this comes from mammalian systems, where activators binding to Mediator caused dramatic activator-specific changes in Mediator structure (Taatjes et al. 2002, 2004; Meyer et al. 2010). It is not yet proven that these conformational changes occur in functional transcription complexes, but once the mechanisms of these changes are understood, it should be possible to genetically manipulate Mediator conformation and test whether and how it contributes to activation. It will also be very informative to do similar EM studies with yeast Mediator to test if it undergoes conformational changes in response to activators. An additional possibility is that Mediator itself is the target of signaling pathways that directly modulate Mediator activity though covalent modification. Studies to examine Mediator modification upon activation of various signaling pathways should begin to reveal if this mechanism is important for gene regulation.

Numerous studies in many systems have demonstrated that chromatin modification and remodeling directed by transcription factors is a key mechanism for gene activation (Narlikar et al. 2002; Li et al. 2007; Weake and Workman 2010). An early example of the importance of chromatin remodeling in yeast was found at the yeast PHO5 promoter (Svaren and Horz 1997). In another example, it was first thought that nucleosome remodeling was unimportant in the mechanism of Gal4-mediated activation because transcription of genes such as GAL1,10 were not affected by mutation of the chromatin remodeler SWI/SNF. However, recent studies showed that SWI/SNF is important for the normal rapid induction kinetics of the GAL1,10 genes (Bryant et al. 2008). It has been argued that activation involving nucleosome remodeling is another example of the recruitment mechanism because known chromatin-modifying coactivators such as SWI/SNF and NuA4 are recruited to regulatory regions by direct interaction with activators (Ptashne and Gann 2002). However, these chromatin-modifying coactivators are not by themselves sufficient for activation since artificial recruitment of factors with only chromatin-remodeling activity has not been observed to stimulate transcription (Green 2005).

Post-initiation mechanisms have most clearly been demonstrated in higher eukaryotes. The best-studied example is regulation of Pol II pausing, shortly after initiation (Buratowski 2009; Fuda et al. 2009). In genome-wide studies of mammalian and insect cells, it was found that pausing is a key and widely used mechanism of gene control and cell identity (Zeitlinger et al. 2007; Core et al. 2008; Nechaev et al. 2010). However, Pol II pausing in yeast has not been found to be a major mechanism of gene regulation. NELF, a key component of the metazoan regulatory circuit, is not conserved in yeast, and the yeast elongation factor Spt4–5 seems to have only a positive function in contrast to its metazoan counterparts. Second, studies mapping the distribution of Pol II along coding sequences have not found many instances where there is an abundance of Pol II confined to the gene 5′ end (Steinmetz et al. 2006).

**Cooperativity between coactivators**

It is clear from numerous studies in many systems that activator function ultimately results in the recruitment of a functional PIC to the promoter (Ptashne and Gann 2002; Green 2005). Measurement of factors crosslinked to promoters after induction showed that factors are recruited in an order specific to the gene under study (Cosma 2002). For example, at the yeast GAL1 promoter, the coactivator SAGA is initially recruited, followed shortly after by Mediator and subsequently by rapid binding of TBP, Pol II, and other general factors (Bryant and Ptashne 2003). Mediator recruitment appears blocked if SAGA is disrupted, suggesting that SAGA cooperatively recruits Mediator (Bhaumik et al. 2004). Variations of this recruitment mechanism are observed at other promoters. At several Gcn4–dependent promoters, SAGA and Mediator are recruited simultaneously in an interdependent fashion (Govind et al. 2005; Qiu et al. 2005). A common theme in these and other studies is that the targets of activators function cooperatively to generate an active PIC. There is little known about how different
coactivators interact, and investigating the mechanism and function of coactivator cooperativity should reveal much about how signaling pathways converge to modulate transcription.

**Activation domains**

Activators are typically bipartite with separate DBD and activation domains. Most activation domains studied do not fold into well-ordered structures in the absence of a binding target, in contrast to the well-defined structure of DNA-binding motifs. For example, the activation domains of VP16, CREB, and Gcn4 all appear to be unstructured in the absence of a binding partner (Huth et al. 1997; Radhakrishnan et al. 1997; Uesugi et al. 1997; Dames et al. 2002; Freedman et al. 2002; Brzovic et al. 2011). In contrast to DNA-binding motifs, activation domain sequences are often not highly conserved (Martchenko et al. 2007). The mechanism of how activation domains specifically identify and bind their relevant multiple targets is a major unanswered question in the transcription field and a key for understanding the mechanism and specificity of many activators.

Acidic activation domains (enriched in acidic residues) are an important class that universally stimulate transcription in all eukaryotes tested (Ptashne and Gann 1990). However, as discussed below, the critical residues of these activation domains are typically hydrophobic while the function of the acidic residues is not yet clear. Originally recognized in yeast Gal4 and Gcn4, these activators encompass most of the well-characterized yeast activation domains and include strong mammalian and viral activators such as p53 and VP16. P53 contains two tandem activation domains, TAD1 and TAD2, and several structures containing the p53 activation domains have been determined (Kussie et al. 1996; Bochkareva et al. 2005; Di Lello et al. 2008; Feng et al. 2009). These structures all involve binding of one to two short α-helices to the target protein mediated primarily by hydrophobic interactions as well as some charged and polar interactions. While these structures are an important advance, the basis of activator-target specificity is not yet understood. Important questions yet to be understood include the following: (1) What are the common features of activator-binding domains? (2) How is activator-target specificity determined? and (3) How does the interaction of activators with these targets contribute to transcription activation?

**Activator targets**

A major question since the discovery of activation domains has been the identity of the relevant activator targets. Over the past 10 years, a combination of biochemical, genetic, and structural experiments has conclusively identified relevant targets for some mammalian and yeast activators (e.g., Stevens et al. 2002; Yang et al. 2004; Green 2005; Sampietro et al. 2006; Waters et al. 2006; Thakur et al. 2009; Herbig et al. 2010; Jedidi et al. 2010). In general, these targets are located in coactivator complexes and chromatin-remodeling or -modifying factors rather than in subunits of the general transcription factors. For example, protein crosslinkers positioned within the activation domains of Gcn4 and Gal4 have conclusively identified three common activator targets (Gal11, Tra1, and Taf12) that are subunits of four coactivator complexes (Mediator, SAGA, NuA4, and TFIID) (Fishburn et al. 2005; Reeves and Hahn 2005; Herbig et al. 2010). The activator–Tra1 interaction was also revealed by in vivo FRET analysis (Bhaumik et al. 2004) and by in vitro interaction and functional studies in vivo (Brown et al. 2001). Similarly, two subunits of the chromatin remodeler SWI/SNF interact with Gcn4 in functional assays (Prochasson et al. 2003). Recent studies indicate that Rap1, a yeast factor that participates in activation of most ribosomal protein genes as well as many other genes, functionally interacts with several subunits of TFIID (Layer et al. 2010; Papai et al. 2010).

**Mechanism of Gcn4–Gal11 interaction**

Activators interact with many of the same coactivator subunits, yet the sequences of the activation domains are not well conserved and the activator-binding subunits are not obviously related in sequence. This raises the question of how an activator can interact with multiple unrelated targets and whether activator–coactivator binding is specific. One of the best-studied examples of activator–target binding is that of the activator Gcn4 binding to the Mediator subunit.
Gal11/Med15 (Park et al. 2000; Herbig et al. 2010; Jedidi et al. 2010). Gal11 contains five conserved domains, four of which are involved in activator binding. Several studies have found that at least three of these Gal11 domains interact with both of the Gcn4 activation domains (residues 1–100 and 101–134) (Park et al. 2000; Majmudar et al. 2009; Herbig et al. 2010; Jedidi et al. 2010). Surprisingly, functional studies showed that each of these four Gal11 activator-binding domains contributes additively to activation by Gcn4 (Park et al. 2000; Herbig et al. 2010; Jedidi et al. 2010). NMR analysis has shown that the binding of Gcn4-Gal11 is unstable with a half-life of less than a millisecond (Brzovic et al. 2011). This property can explain why Gal11 has multiple activator-binding domains and Gcn4 has tandem activation domains that bind to multiple sites on Gal11 (Figure 11). The model proposed to explain activator–target binding in this system is that the two Gcn4 activation domains rapidly sample multiple activator-binding domains on Gal11. These rapidly cycling interactions are capable of allowing Gcn4 to recruit Mediator to gene regulatory regions in the absence of a stable protein–protein interaction.

NMR structural analysis of the Gcn4 central activation domain bound to one Gal11 activator-binding domain (residues 158–238) has revealed much about the nature of activator–target binding and how activators can functionally interact with different unrelated coactivators (Brzovic et al. 2011). Upon binding to Gal11, about eight formerly unstructured Gcn4 residues form a helix that interacts with Gal11. The Gcn4-Gal11 protein interface is extremely simple and is purely hydrophobic, with no observed contribution from charged or polar interactions. Because of this simple interface, the Gcn4 backbone is highly flexible and is predicted by NMR to exist in multiple conformations, and, surprisingly, at least two of these conformations bind in approximately opposite orientations on Gal11. This activator–target complex is an example of a so-called “fuzzy complex” (Tompa and Fuhrer 2008) where the structure of a protein complex cannot be described by a single conformational state. These properties probably explain how one class of activators interacts with multiple unrelated targets, and it is likely that this is a commonly used mechanism for activator–target interactions. Acidic residues in the activation domain could specifically interact with the coactivator target, contribute to a nonspecific long-range electrostatic attraction, or play no role. For the Gcn4 central activation domain, mutation of all 10 acidic residues to Ala had little effect, suggesting that, in this case, acidic residues do not play an important role (Brzovic et al. 2011). It is certainly possible that acidic residues in other activation domains play important functional roles, and it will be important to address this in future work.

**Perspective**

The S. cerevisiae system has made many invaluable and groundbreaking contributions to the understanding of gene control in eukaryotes. Although a few aspects of gene regulation occur only in higher eukaryotes, most of the fundamental mechanisms of transcriptional regulation have been conserved from yeast to humans. Because of the powerful combination of genetics, molecular biology, biochemistry, and genome-wide methods that can be utilized, the yeast system has often been at the forefront in discovering and understanding fundamental regulatory mechanisms. It is certain that in the next decade and beyond the yeast system will be at the forefront of fundamental discoveries in transcriptional regulation and serve as an excellent model for understanding regulatory mechanisms in other eukaryotes.

**Literature Cited**


Allard, S., R. T. Utlej, J. Savard, A. Clarke, P. Grant et al., 1999 Nua4, an essential transcription adaptor/histone H4 acetyltransferase complex containing esalp and the ATM-related cofactor tralp. [In Process Citation] EMBO J. 18: 5108–5119.


Bali, M., B. Zhang, K. A. Morano, and C. A. Michels, 2003 The Hsp90 molecular chaperone complex regulates maltose induc-


Birch, C., O. Poch, C. Romier, M. Ruff, G. Mengus et al., 1998 Human TAF(H)28 and TAF(H)18 interact through a histone fold encoded by atypical evolutionarily conserved motifs also found in the SPT3 family. Cell 94: 239–249.


Young, E. T., K. M. Dombek, C. Tachibana, and T. Ideker, 2003 Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. J. Biol. Chem. 278: 26146–26158.


Communicating editor: A. Hinnebusch