Transcription Annotation Guide

Purpose: The goal in this guide is to provide an overview of transcription systems from various organisms and provide guidance in annotating gene products involved in transcription, with a focus on RNA polymerase II.

Recommendation to consider RNA polymerase II-specific terms: Although mRNA transcription by RNAP II represents a relatively small proportion of the total RNA in the cell, the vast majority of the transcription literature, particularly that for eukaryotes, focuses on genes transcribed by RNA polymerase II. While researchers studying the mechanism of transcription by RNAP II will specifically talk about which RNA polymerase they are studying, the vast majority of researchers studying transcriptional regulation of their favorite (protein-coding) gene, will not mention the RNA polymerase even a single time in their paper, even though we know this must be RNAP II transcription. Thus, we ***strongly recommend*** that annotators of eukaryotic organisms (with an occasional exception for annotators of trypanosome genes, see RNA polymerase I comments below), should ***consider using the appropriate RNA polymerase II specific terms*** whenever they are annotating anything relating to transcriptional regulation of protein coding genes.

Recommendations for transcription factors that are complexes: Many "transcription factors" are complexes, not single subunit gene products. **1)** Where the functional unit of action is the complex, the MF **"*transcription factor activity*"** terms (i.e. terms with "transcription factor activity" as part of their term names) listed here are appropriate for the function of the transcription factor complex and it will often be appropriate to make annotations for all members of the complex using the "contributes to" qualifier. **2)** Some transcription factors contains subunits which, individually, possess catalytic activities. These activities should be annotated only to the subunits which possess them, not to all members of the complex.

Consider annotations to multiple terms due to has_part relationships: Note that in this guide, we have indicated appropriate terms representing overall "complex" functions of a transcription factor, and some transcription factors have more than one molecular function. These functions generally involve one or more binding interactions. Following the decision at the 2010 Annotation Camp in Geneva to use the has_part relationship to represent when a "complex" function has binding as a component of how it functions, these terms have one or more has_part relationships to represent the appropriate binding interactions between the transcription factors and nucleic acids (DNA or RNA) or proteins (RNA polymerase, transcription factors). Note that due to the decision to use the has_part relationship, annotation to a transcription factor activity term does NOT automaticaly result in annotation to a binding term via an is_a path. If you wish have the binding aspects of these functions appear in enrichment analyses, you should make direct annotations to the appropriate binding functions if appropriate. When a single subunit is shown to possess binding activity individually, e.g. TBP binds to TATA containing DNA, you should annotate that subunit individually. For binding that occurs in the context of a multisubunit complex, it might be appropriate to annotate all subunits of the multisubunit complex using "contributes to". You can get indications of appropriate binding terms to consider by following the has_part relationships from the "complex" function terms.

Guide to RNA polymerases

RNA polymerases can be single subunit or multisubunit. The multisubunit RNA polymerases are complexes of multiple gene products and are represented in the Cellular Component ontology. The single subunit RNA polymerases are not complexes and are not represented in the Cellular Component ontology. Both single subunit and multisubunit RNA polymerase activities are represented in the Molecular Function ontology, connected via a part_of relationship to a Biological Process term representing transcription from a specific type of promoter. While the basic principles are in place, not all transcription systems are completely represented in GO. If you need additional terms, please request them.

Multisubunit RNA polymerases

Nuclear (eukaryotic) RNA polymerases

RNA polymerase I - RNA polymerase I is present in all eukaryotes. In almost all species, it makes only a single transcript, the large ribosomal RNA (rRNA), in large quantities from the rDNA repeats There is one known exception to this rule. In trypanosomes, RNAP I also transcribes genes encoding the parasite's major cell-surface proteins-the variant surface glycoprotein (VSG) and procyclin (PMID:17972917). The structure of the RNAP I promoter is not well conserved between species.

RNA polymerase II - RNA polymerase II is present in all eukaryotes. In most organisms, RNAP II transcribes ALL nuclear-encoded protein coding genes (the single known exception is the protein coding genes transcribed by RNAP I in trypanosomes). RNAP II also transcribes numerous non-coding RNAs. In many species, RNAP II transcribes four of the five snRNAs (U1, U2, U4, and U5) and many/most (depending on the species) of the snoRNAs. RNAP II also transcribes various other non-coding RNAs, e.g. the RNase MRP and telomerase RNAs. However, it should be noted that there is considerable variability between species as to whether a given small non-coding RNA is transcribed by RNAP II or RNAP III.

RNA polymerase III - RNA polymerase III is present in all eukaryotes. RNAP III transcribes all nuclear-encoded tRNAs and the 5S rRNA. RNAP III generally transcribes the U6 snRNA. RNAP III also transcribes various other non-coding RNAs, e.g. the RNase P RNA and the 7SL RNA component of the signal recognition particle. However, it should be noted that there is considerable variability between species as to whether a given small non-coding RNA is transcribed by RNAP II or RNAP III.

RNA polymerases IV & V - RNA polymerases IV and V are present in plants. Pol IV and Pol V produce small-interfering RNAs (siRNAs) RNAs involved in RNA-directed DNA methylation (RdDM) in order to silence transcription of retrotransposons, endogenous repeats, invading DNA viruses, transgenes and some protein-coding genes. Pol IV and Pol V have different roles in the RdDM pathway, with Pol IV being required for siRNA biogenesis at the beginning of the pathway and Pol V transcripts being required for siRNA targeting of RdDM-affected loci (PMID:21779025).

Other multisubunit RNA polymerases

Bacterial-type RNA polymerase - The eubacterial DNA-directed RNA polymerase is a multisubunit complex with a core composed of the essential subunits beta-prime, beta, and two copies of alpha and a fifth nonessential subunit called omega. An additional subunit, a sigma factor, is required for promoter recognition and specificity.

PEP: the Plastid-Encoded Plastid RNA polymerase - An RNA polymerase complex containing polypeptides encoded by the plastid (i.e. chloroplast) genome. Plastid-encoded DNA-directed RNA polymerases resemble eubacterial multisubunit RNA polymerases, with a core composed of alpha, beta, and beta-prime subunits. Some forms contain multiple additional subunits. An additional sigma factor subunit is required for promoter recognition (PMID:21316793, PMID:10946105).

Archaeal RNA polymerase - Archaea have a single multisubunit RNA polymerase that resembles a eukaryotic nuclear multisubunit RNA polymerase (PMID:11839495). We do not currently represent Archaeal RNA polymerase or transcription in GO.

Single subunit RNA polymerases

Phage T3/T7 type RNA polymerase - The simple single subunit RNA polymerase of the T-odd bacteriophages T3 and T7 is considered a prototype single subunit RNA polymerase. (PMID:7526118).

mitochondrial RNA polymerase(s) - The mitochondrial RNA polymerases of most eukaryotic organisms are nuclear encoded enzymes with homology to the single subunit phage T3/T7 RNAP (PMID:8604305). Some plants have multiple T3/T7 type mitochondrial RNA polymerases (PMID:21316793).

chloroplast/plastid phage-type RNA polymerase(s) - In addition to containing a bacterial-type multisubunit RNA polymerase, plant chloroplasts/plastids may also contain one or more phage T3/T7 type single subunit RNA polymerases (PMID:21316793).

Defining transcription regulatory regions

Ambiguous words

promoter - Be very wary of the word "promoter" without any further qualification. The word promoter has been used in multiple ways. In E.coli, promoters were originally defined as the region bound by the RNAP and basal transcription factors (TFs). In eukaryotic transcription research, some researchers have used the word strictly as originally defined in the E. coli system, but other researchers have used the word promoter to include both the core promoter and the entire upstream regulatory region. Very commonly in papers on regulation of RNAP II genes, if you see "promoter", the researchers are referring to the "core promoter" and the "core promoter proximal region" together. This is why the term "promoter binding" has been removed, and replaced by "core promoter binding" and "core promoter proximal region DNA binding", both of which are part of the "transcription regulatory region".

enhancer - The definition of an "enhancer" suffers from a similar ambiguity in the literature as the word "promoter" in that it has been used for a number of not completely identical regulatory regions. In mammals distal ehancers are seen to act independently of location or distance. In S. cerevisiae, the word "enhancer" has been used for upstream regulatory regions that improve transcription, however, yeast, both S. cerevisiae and S. pombe, are generally considered to not contain "distal enhancers" like those of mammalian cells, though I have seen some literature regarding "regulation at a distance" in yeast. If you see the word "enhancer" in yeast literature, consider if "core promoter proximal region" would be appropriate. Bacterial transcription also uses the phrase "enhancer", but the relative position between the enhancer and the transcription start site is very different compared to the distance between an enhancer for RNAP II transcription in a mammalian cell.

Phrases used in GO Terms

transcription regulatory region - The "transcription regulatory region" is the phrase used to describe the nucleic acid sequence that contains regulatory elements that regulate transcription. When referring to DNA sequence, e.g. "transcription regulatory region DNA binding", the transcription regulatory region for RNA polymerases which undergo lots of diverse regulation, i.e. bacterial-type RNAP and RNAP II, contains the "core promoter", the "core promoter proximal region", and possibly also an "enhancer" region. For RNA polymerases I and III, the terminology for the various elements of the transcription regulatory region is different from that for RNAP II and unique to that RNAP system.

core promoter - The "core promoter" is, by definition, where the basal transcription machinery binds. The concept of basal transcription is most relevant to RNA polymerases which are highly regulated to transcribe different sets of genes in response to different conditions or stimuli, i.e. bacterial-type RNAPs and RNA polymerase II. For bacterial type RNAPs, the core promoter is the binding site for the holo-RNAP complex, i.e. core RNAP plus sigma. For RNA polymerase II, the core promoter is also the binding site for the basal transcription machinery, which includes the basal (or general) transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH), and RNA pol II itself. If you are considering the binding site of a transcription regulatory protein for RNAP II that is NOT part of the basal transcription machinery, "core promoter" is probably NOT the appropriate region; consider "core promoter proximal region". It is not clear that the phrase "basal transcription" is meaningful for either RNA polymerase I or RNA polyerase III. However, the promoter for RNAP I has an element referred to as the "CORE" element.

core promoter proximal region - The "core promoter proximal region" is the region proximal to and upstream of the "core promoter". This is the region where the "classic" DNA binding "transcription factors", e.g. Gal4 in S. cerevisiae and x in mammalian cells, bind to specific DNA sequences and activate (or repress) transcription. In S. cerevisiae, the word "enhancer" has been used for upstream regulatory regions that improve transcription, however, yeast, both S. cerevisiae and S. pombe, are generally considered to not contain "distal enhancers" like those of mammalian cells, though I have seen some literature regarding "regulation at a distance" in yeast. If you see the word "enhancer" in yeast literature, consider if "core promoter proximal region" would be appropriate. Particularly at the most general level, there is no absolute distance that can be used as the outer limit of the "core promoter proximal region". Partly this is due to the fact that bacterial transcription is very different from eukaryotic transcription and even comparing RNAP II from yeast to RNAP II from mammals, there are significant differences in the distance of the transcription regulatory elements from the transcription start site.

distal enhancer - Distal enhancers are present in some eukaryotic systems. They are defined as acting in a manner that is independent of distance and orientation. Yeast (pombe and cerevisiae) do not have distal enhancers, as far as we are aware. Perhaps there should be a taxon constraint on a term for "distal enhancer".

RNA polymerase I enhancer - Some early work described enhancers for RNA polymerase I. It is not completely clear whether these sites truly behave as enhancers in the same sense as distal enhancers for RNA polymerase II in mammalian systems.

bacterial-type RNA polymerase enhancer - In E. coli, enhancers are common in sigma-54 type promoters, and found in some sigma-70 type promoters, and act at least a minimum distance from the core promoter, but the bacterial literature does not use the phrase "distal enhancer" and the bacterial researchers we have consulted do not like the idea of using this phrase for bacterial enhancers.

RNA polymerase II basal transcription machinery

The basal transcription machinery is typically described as the minimum set of factors required for preinitiation complex (PIC) formation. For RNA polymerase II (RNAP II or Pol II), the basal machinery includes RNAP II itself as well as a number of "basal transcription factors" (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH), most of which are complexes (except for TFIIB). The basal transcription factors are often described as having multiple roles, many of which are related to binding and which occur in the context of the _entire_ transcription factor, which again, is usually a complex. Some of the transcription factors also contain catalytic activities, however these activities are generally possessed by individual subunits. This section shows the summaries from two reviews and indicates the appropriate term(s) that correspond to each "function". Note that since researchers often use the word "function" in a looser way than the GO meaning of "molecular function" (indicated in orange), some of these "functions" correspond to terms in "biological process" (indicated in blue) in GO. Some of the indicated functions are catalytic activities contained by individual subunits within a given complex. These catalytic activities were outwith the scope of the transcription overhaul (txnOH). We have indicated that (in yellow), but also included a list of already existing catalytic terms to consider. There are a couple places, highlighted in green, where we suspect a new term is warranted, but have not yet implemented it. Feel free to give us input on these.

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (Table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
TFIIA (2-3 subunits)	H. sapiens: p35 (alpha), p19 (beta) & p12 (gamma) S. cerevisiae: TOA1 & TOA2	antirepressor	functions to counteract repressive effects of negative cofactors like NC2	may need a new MF term for this? OR just a process term unless we know the mechanism.
		stabilizes TBP-TATA complex		(contributes to) MF GO:0001129 - TBP-class protein binding RNA polymerase II transcription factor activity involved in preinitiation complex assembly
		coactivator	acts as a coactivator by interacting with activators and components of the basal initiation machinery	(contributes to) MF GO:0001128 - RNA polymerase II transcription coactivator activity involved in preinitiation complex assembly

RNA polymerase II basal transcription machinery, continued

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (Table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
TFIIB (1 subunit)	H. sapiens: p33	start site selection	directs accurate start site selection	BP GO:0001174 - transcriptional start site selection at RNA polymerase II promoter
	S. cerevisiae: SUA7	stabilizes TBP-TATA complex	stabilizes TFIID-promoter binding	MF GO:0001129 - TBP-class protein binding RNA polymerase II transcription factor activity involved in preinitiation complex (PIC) assembly
		pol II/TFIIF recruitment	aids in recruitment of TFIIF/Pol II to the promoter	probably need a new term for recruitment of TFIIF/Pol II complex together
TFIID (14 subunits including TBP and TAFs)	H. sapiens: TBP + TAFs (TAF1-TAF14)	core promoter-binding factor	nucleates PIC assembly either through TBP binding to TATA sequences or TAF binding to other promoter sequences	(contributes to) MF GO:0001075 - sequence-specific core promoter binding RNA polymerase II transcription factor activity involved in preinitiation complex assembly
	S. cerevisiae: TBP (SPT15) + TAFs (TAF1-TAF14)	coactivator	coactivator activity through direct interaction of TAFs and gene specific activators	(contributes to) MF GO:0001128 - RNA polymerase II transcription coactivator activity involved in preinitiation complex assembly
		protein kinase		annotate individual subunits as appropriate for catalytic activities (not addressed by txnOH)
		ubiquitin-activating/ conjugating activity		
		histone acetyltransferase		

RNA polymerase II basal transcription machinery, continued

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (Table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
TFIIE (2 subunits)	H. sapiens: p56 (alpha) & p34 (beta) S. cerevisiae: TFA1 & TFA2	recruits TFIIH	helps recruit TFIIH to promoters	(contributes to) MF GO:0001138 - TFIIH-class transcription factor recruiting transcription factor
		facilitates formation of an initiation-competent pol II		
			stimulates helicase and kinase activities of TFIIH	outwith scope of txnOH, consider BP annotations to regulation terms (GO:0033674 or GO:0051096) with TFIIH as a target in column 16
			binds ssDNA	MF GO:0003697 - single-stranded DNA binding (consider whether to annotate as complex or individual subunit)
			and is essential for promoter melting	BP GO:0001113 - transcriptional open complex formation at RNA polymerase II promoter
		involved in promoter clearance		BP GO:0001111 - promoter clearance from RNA polymerase II promoter

RNA polymerase II basal transcription machinery, continued

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (Table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
TFIIF (2-3 subunits)	H. sapiens: RAP30 & RAP74	binds pol II and facilitates pol II recruitment to the promoter	tightly associates with RNApII; enhances affinity of RNApII for TBP-TFIIB-promoter complex	(contributes to) MF GO:0001139 - core RNA polymerase II recruiting transcription factor activity
		recruits TFIIE and TFIIH	necessary for recruitment of TFIIE/TFIIH to the PIC	(contributes to) MF GO:0001136 - TFIIE-class transcription factor recruiting transcription factor activity AND (contributes to) MF GO:0001138 - TFIIH-class transcription factor recruiting transcription factor activity
	S. cerevisiae: TFG1, TFG2, & TAF14	functions with TFIIB and pol II in start site selection	aids in start site selection	BP GO:0001111 - promoter clearance from RNA polymerase II promoter
		facilitates pol II promoter escape	and promoter escape	BP GO:0001111 - promoter clearance from RNA polymerase II promoter
		enhances the efficiency of pol II elongation	enhances elongation efficiency	BP GO:0032968 - positive regulation of transcription elongation from RNA polymerase II promoter

RNA polymerase II basal transcription machinery, continued

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (Table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
TFIIH (10 subunits)	H. sapiens: p89/XPB, p80/xPD, p62, p52, p44, p40/CDK7, p38/Cyclin H, p34, p32/MAT1, & p8/TFB5	ATPase activity for transcription initiation and promoter clearance; helicase activity for promoter opening	ATPase/helicase activity for promoter opening and promoter clearance	annotate individual subunits as appropriate for catalytic activities (not addressed by txnOH)
				BP GO:0001113 - transcriptional open complex formation at RNA polymerase II promoter
			facilitates transition from initiation to elongation	BP GO:0001111 - promoter clearance from RNA polymerase II promoter
	S. cerevisiae: CCL1, KIN28, RAD3, SSL1, SSL2, TFB1, TFB2, TFB3, TFB4, & TFB5	transcription-coupled nucleotide excision repair	helicase activity for transcription coupled DNA repair	out of scope of txnOH, but consider BP GO:0006283 - transcription-coupled nucleotide-excision repair, or child terms
		kinase activity for phosphorylating pol II CTD	kinase activity required for phosphorylation of RNApII CTD	annotate individual subunits as appropriate for catalytic activities (not addressed by txnOH)
		E3 ubiquitin ligase activity		
RNAP II or pol II (12 subunits)	H. sapiens: RPB1-RPB12	transcription initiation, elongation, termination	catalyzes transcription of all mRNAs and a subset of noncoding RNAs including [most] snoRNAs and miRNAs	(contributes to) MF GO:0001055 - RNA polymerase II activity
				BP GO:0006366 - transcription from RNA polymerase II promoter
	S. cerevisiae: RPO21/RPB1, RPB2, RPB3, RPB4, RPB5, RPO26/RBP6, RPB7, RPB8, RPB,9 RPB10, RPB11, & RPC10/RPB12	recruitment of mRNA capping enzymes		out of scope of txnOH, not currently represented in GO
		transcription-coupled recruitment of splicing and 3' end processing factors		BP GO:0034402 - recruitment of 3'-end processing factors to RNA polymerase II holoenzyme complex
		CTD phosphorylation, glycosylation, and ubitination		[occurs on RNAP II, not something RNAP II does]

RNA polymerase II cofactors

A number of transcription cofactors are characterized as being involved in or required for transcription from RNAP II promoters. A cofactor is defined as "interacting seletively and non-covalently with a regulatory transcription factor and also with the basal transcription machinery in order to modulate transcription." The phrase "cofactor activity" is invoked a lot in the literature, and it was difficult to pin down exactly how they function beyond this general definition. I think it may turn out that there are multiple different types of cofactors that act in slightly different ways. Some of the basal transcription factors, e.g. TFIIA and TFIID, are described as having "cofactor activity". These two large complexes have not been classified as part of the basal transcription machinery, and have been purified numerous different ways with slighly differing compostions. In addition, there is some evidence that these complexes may be involved in "basal" transcription as well as in "activator-dependent" transcription, adding to the confusion. Nevertheless, these complexes are described as having "cofactor activity", as defined above. Thus we have represented that level of detail, including the positive (coactivator) and negative (corepressor) versions of cofactors. If our understanding advances, it may be appropriate to make more specific terms representing more precise mechanisms of action. The functions given for the cofactor complexes in this section of the guide are derived only from the Sikorski & Buratowski review; again we have listed appropriate MF (orange) or BP terms (blue) that correspond with the described functions.

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
Mediator (at least 24 subunits)	multiple complexes have been identified; functional differences between them are not always clear		bridges interaction between activators and basal factors	(contributes to) MF GO:0001128 - RNA polymerase II transcription coactivator activity involved in preinitiation complex assembly
			stimulates both activator dependent and basal transcription	BP GO:0006357 - regulation of transcription from RNA polymerase II promoter
			required for transcription from most RNApII dependent promoters	BP GO:0006366 - transcription from RNA polymerase II promoter
SAGA (20 subunits)	multiple complexes have been identified; functional differences between them are not always clear		interacts with activators, histone H3, and TBP	(contributes to) MF GO:0001128 - RNA polymerase II transcription coactivator activity involved in preinitiation complex assembly
			histone acetyltransferase activity	annotate individual subunits as appropriate for catalytic activities (not addressed by txnOH)
			deubiquitinating activity	

Catalytic activities of subunits of RNAP II basal txn factors

The transcription overhaul dealt with the functions of transcription factors as a whole. Thus, for the basal transcription factors, most of which are complexes, we dealt only with the functions of the complexes. Some of the basal transcription factors and cofactors contain subunits which individually have catalytic activities, e.g. kinase, helicase, histone acetyltransferase, etc. This section indicates some of the terms which exist for these catalytic activities. These terms, or their child terms, may be appropriate for annotation of the individual subunits which possess them. Note that the transcription overhaul did not address any of these terms, that most of these are not specific to transcription. As this was not the focus of the transcription overhaul, we have selected some high level terms as a guide; you should consider if there are more specific terms appropriate for a given gene.

Factor	Protein composition	"Function" from Thomas & Chiang 2006 (table 1)	from Sikorski & Buratowski (2009)	Suggested GO term(s)
		protein kinase		GO:0004672 - protein kinase activity (or child terms)
		kinase activity for phosphorylating pol II CTD	kinase activity required for phosphorylation of RNApII CTD	MF GO:0008353 - RNA polymerase II carboxy-terminal domain kinase activity
		ATPase activity [for transcription initiation and promoter clearance]	ATPase/helicase activity [for promoter opening and promoter clearance]	GO:0016887 - ATPase activity
		helicase activity [for promoter opening]		GO:0004386 - helicase activity
		histone acetyltransferase		GO:0004402 - histone acetyltransferase activity
		ubiquitin-activating/ conjugating activity		GO:0004842 - ubiquitin-protein ligase activity
		E3 ubiquitin ligase activity		
		deubiquitinating activity		GO:0004843 - ubiquitin-specific protease activity

RNA polymerase II "regulatory" transcription factors

Type of factor, by DNA binding	Examples	Note	Suggested GO term(s)
Factors which bind DNA sequences in a sequence specific mannor (generally upstream of and) proximal to the core promoter. Note that proximal is used loosely, as opposed to "distal enhancers". These are the classic "regulatory transcription factors" which bind to a specific DNA sequence.	Mammalian: myoD, NFκB, etc.	This term is the most general, requiring only that you know that it binds to a specific sequence and in doing so, modulates transcription	MF GO:0000982 - RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity
		A great many of the "regulatory" transcription factors act by improving the rate of PIC formation. In practice though, David has usually not seen evidence for this.	MF GO:0001074 - RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in preinitiation complex assembly
	S. cerevisiae: GAL4, GCN4, etc.	In practice, David has been finding it easy to find evidence for a positive effect, or a negative effect, on transcription	MF GO:0001077 - RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription
			MF GO:0001078 - RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription

RNA polymerase II "regulatory" transcription factors, cont.

Type of factor, by DNA binding	Examples	Note	Suggested GO term(s)
Factors which bind enhancer DNA sequences distal to the core promoter	David and I kept these enhancer terms because, particularly for mammalian, people talk about "distal enhancers" that function independently of distance or orientation, which is what the term appears to have been originally intended for. However, David has come across cases where the term enhancer is used very loosely to mean things other than these distal regulatory sites. Sometimes for the same site, some groups indicate that it is proximal to the core promoter, while others refer to it as an enhancer. Karen has come across use of the word "enhancer" to refer to the UAS region in S. cerevisiae while other researchers refer to it in ways that would suggest that it is part of the core promoter proximal region. Be cautious of using these terms unless you are really sure that it is an enhancer; this may take reading more than a single paper. Please feel free to provide input on improving the usage of "enhancer".		MF GO:0003705 - RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity
			MF GO:0001205 - RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription
			MF GO:0001206 - RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription
Factors which bind DNA sequences of uncertain type	When you want to indicate that a transcription factor is binding DNA in a sequence specific manner and thus regulating RNA polymerase II transcription, but you can not determine what region of the transcription regulatory region that the factor is binding to, consider these terms.		MF GO:0000981 - sequence-specific DNA binding RNA polymerase II transcription factor activity
			MF GO:0001228 - polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription
			MF GO:0001227 - RNA polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription