CHARACTERIZATION OF BINDING SPECIFICITY PROPERTIES OF VARIANT TATA BINDING PROTEINS

by

SUSAN MAE CHEN

A THESIS

Presented to the Department of Biology
and the Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

June 1999

Approved: [redacted]

Dr. Diane K. Hawley
Advisor

An Abstract of the Thesis of

Susan Mae Chen for the degree of Bachelor of Arts

in the Department of Biology to be taken June 1999

Title: CHARACTERIZATION OF BINDING SPECIFICITY PROPERTIES OF VARIANT TATA BINDING PROTEINS

Approved: [redacted]
Dr. Diane K. Hawley

The TATA Binding Protein (TBP) is required for transcription initiation by all three eukaryotic RNA polymerases. At promoters of most genes transcribed by RNA polymerase II, TBP binds to the TATA sequence (consensus TATAAAAG) and nucleates the formation of the preinitiation complex. The preinitiation complex contains transcription factors and polymerase and is required before transcription can begin. TBP has a conserved C-terminal domain and a highly variable N-terminal domain whose function has yet to be characterized.

We have studied mutants of yeast TBP to learn more about the effect of the N-terminus on TBP behavior. We compared the *in vitro* binding behavior of full-length and N-terminal delete versions of wild-type TBP and the altered specificity double mutant I194FL205V TBP. Our observations that both full-length and N-terminal delete versions of I194FL205V TBP showed high affinity for all TATA sequences tested, while both versions of wild-type TBP showed specificity for consensus TATA only lead to the conclusion that removal of the N-terminus is not sufficient to account for binding to non-consensus TATA sequences.

We also examined the *in vivo* behavior of various TBP mutants whose binding properties had been studied *in vitro*. Reporter gene assays in the yeast *Saccharomyces cerevisiae* allowed the testing of TBP's affinity for altered TATA-box sequences. *In vivo*

findings provide functional support for previous studies in our lab that have shown the ability of I194FL205V TBP to bind better than wild-type TBP to 30G (TGTAAAAG) and 25C (TATAAACG). *In vitro* observations of binding by an N-terminal delete version of L205V TBP were supported by *in vivo* observations of reporter gene activation showing that N-terminal delete L205V TBP functions significantly better than full-length L205V on 30G and consensus TATA. These *in vivo* observations provide physiological relevance to previously observed *in vitro* behavior.

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Professor Diane K. Hawley for giving me the opportunity to do research in her lab. Her guidance in my research has helped me learn more about how to think as a scientist. Her insight, belief in my abilities, and positive outlook have certainly been an encouragement. I can only hope to find an advisor like Diane in my graduate studies. I would also like to thank Dr. Lara Baxley, whose advice and mentorship was consistently given in a patient and helpful manner. I am very fortunate to have learned the basics in molecular biology research from such a great teacher.

I appreciate the expertise of members of the G. Sprague lab. Dr. Dave Mitchell and Dr. Paul Cullen have come to my rescue numerous times to help devise methods to save my yeast stocks when they became contaminated or to answer my questions about ß-gal assays. Their advice is invaluable. I also thank Hilary Kemp, who contributed many yeast constructs used in my research.

My thanks also goes to Sara Tufa, Dr. Charles Foulds, Dr. Matt Thomas, Matt Miller, and Eric Sheagley for patiently answering all my questions about cloning, protein purification, and life in general. And to the Hawley Lab Girls, thanks for all the chocolate and Girl Scout Cookies!

Lastly, I would like to thank my high school biology teacher, Mr. Ford Morishita, for inspiring me, through his enthusiasm for science and learning, to study biology in college.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Scientists have long been fascinated with the intricacies of systems. Some focus on the interactions within components of the solar system, while others look to the relationships between organisms and the environment in the ecosystem, and still others contemplate the molecular and biochemical relationships within organisms. Despite the diversity among the subjects of each study, the goal of each investigation is to gain a glimpse into the unknown realm that lies beyond current understanding. Uncovering the unknown has been the driving force behind discoveries that have shaped the foundation of science, modern medicine, and research.

Before the mid-1800s, individuals had observed the physical manifestations of a concept we call heredity. They saw that people who were related shared physical characteristics. Out in the fields, they may have bred cattle for strength and endurance, or sheep, for quality wool. They knew that each parent contributed some essence from various parts of the body, but this concept could not provide an answer for other observations of inheritance. It was not until Mendel's experiments with pea plants that the nature of inheritance became more clear. He proposed in 1865 that the vehicle of inheritance consisted of discrete units that were passed intact through generations. Nearly one century later, with the help of modern technology, Watson and Crick were able to provide a name and propose a structure for the discrete units of inheritance previously described by Mendel. We know today that genetic information is contained in and inherited through DNA.

Understanding how this information is used by the cell is the next step for investigators. Modern research utilizes the biological paradigms found in the bacteria

Escherichia coli and the budding yeast *Saccharomyces cerevisiae* to answer questions aimed at developing current knowledge about how the information in DNA is processed. Unlocking the secrets of DNA may also provide a key to understanding genetic illnesses such as cancer in humans.

The following chapters of this thesis focus on a protein called the TATA Binding Protein (TBP) and investigate its role in initiating transcription. The thesis begins with background on transcription and TBP followed by a discussion of experiments and results pertaining to TBP's role in transcription initiation.

Transcription

An organism's DNA contains its genetic information. Not only does DNA contain genes; it also contains information about how each gene should be expressed. A specific sequence of nucleotides in front of a gene, called a promoter, signals the start site of a process called transcription. Transcription is the copying of DNA into RNA. RNA then serves as a template for the synthesis of proteins. Proteins are molecules that have essential cellular functions and are involved in growth processes and metabolism.

Transcription is carried out by proteins called RNA polymerases. Transcription can be divided into three stages. In the first stage, the polymerase, with the help of other proteins, associates with the promoter region of the gene to be transcribed. After unwinding the DNA, it moves along the DNA, creating a chain of RNA molecules that are complementary to the DNA strand. Finally, the polymerase terminates and dissociates from the DNA. The polymerase can then participate in another round of transcription.

Transcription Initiation

Eukaryotes contain three types of polymerases. Each type specifically copies, or transcribes, genes that can be classified for one type of function. Of the three eukaryotic polymerases, polymerase II (pol II) is central to the system being studied in this thesis. Pol II transcribes genes encoding messenger RNA (mRNA). mRNA provides information for synthesizing proteins. Most promoters of genes transcribed by RNA polymerase II contain a TATA element (consensus TATAAAAG). This specific sequence is recognized by the TATA-Binding Protein (TBP). TBP is a transcription factor that is a subunit of the TFIID complex and is required for the formation of the preinitiation complex. The preinitiation complex is a set of transcription factors specifically bound at the promoter and is essential for transcription initiation. In formation of the preinitation complex, TBP is the first protein to bind to the promoter. Other transcription factors are then recruited by TBP to the promoter, and finally, pol II joins to complete the preinitation complex. Transcription can now proceed (Figure 1).

TBP

TBP isolated from various eukaryotes including yeasts, vertebrates, insects, and plants have been found to share a common 180 amino acid carboxy-terminal "core" region (Hernandez, 1993). This core region, which contains two structural repeats flanking a highly basic segment known as the basic repeat, is essential for DNA binding and sufficient for basal transcription *in vitro* (Hoey et al., 1990; Horikoshi et al., 1990; Peterson et al., 1990; Lieberman et al., 1991) and activated transcription *in vivo* when core TBP is present as the only copy of TBP in the cell (Poon et al., 1991; Peterson et al., 1990). Structural studies of TBP using x-ray crystallography have provided insight into understanding its tertiary structure, DNA-binding mechanism, and binding orientation.

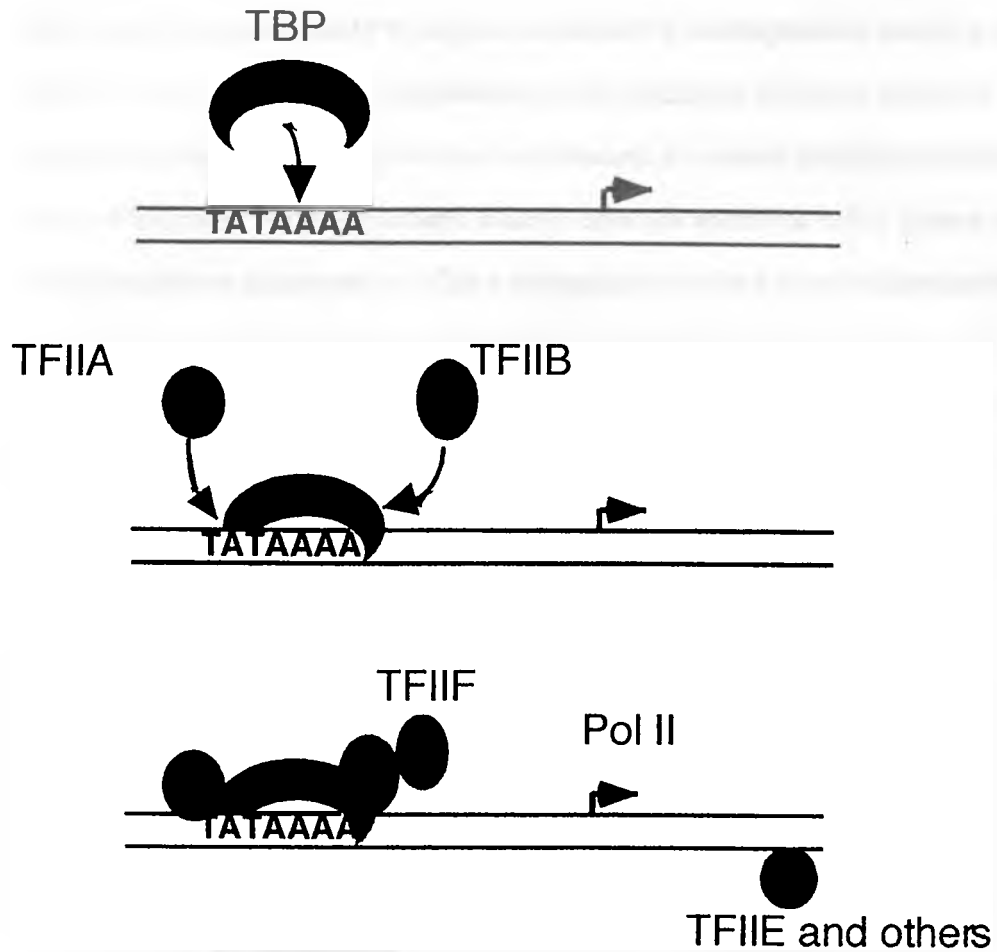

Figure 1. Schematic representation of the steps required in preinitiation complex formation where the arrow indicates the site of transcription initiation. TBP recognizes and binds to the TATA sequence (consensus TATAAAAG) at the promoter. After binding, it recruits other transcription factors and RNA polymerase II to begin transcription.

The crystal structure of core TBP (ΔN TBP) complexed with DNA shows a twofold intramolecular symmetry that generates a saddle-shaped conformation (Kim, J. et al., 1993) (Figure 2). Residues within the saddle-shaped DNA binding region are primarily hydrophobic (Kim, Y. et al., 1993) and bind to the TATA-box with high affinity (Kim, J. et al., 1993). Unlike most DNA binding proteins, which bind in the major groove of helical DNA, TBP binds in the minor groove (Starr and Hawley, 1991). The protein undergoes subtle conformational changes upon binding DNA, but distorts the DNA to a

bend angle of approximately 90 degrees as shown by electrophoretic mobility shift assays (EMSA) (Starr et al., 1995). Sequence-specific binding in the minor groove is highly unusual because the minor groove has been thought to contain insufficient information in terms of hydrogen bonding partners, base orientation within the helix, groove width, and protein-backbone alignment to define a recognition site for a protein (Seeman et al., 1976).

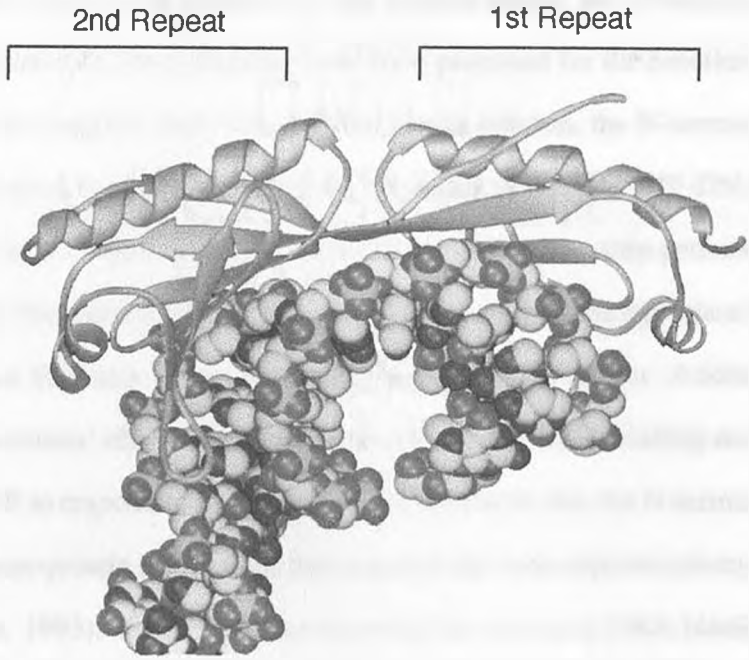

Figure 2. The second repeat of TBP (shown in green) contacts the first half the the TATA-box.

Another observation from the crystal structure is the orientation of TBP bound to the TATA sequence. The crystal structure shows the second repeat of TBP making contacts with the first half of the TATA box, and the first repeat of TBP making contacts with the second half of the TATA box (Figure 2 and Kim et al., 1993).

The amino- (N-) terminal region of TBP varies in length (18-173 residues) and sequence among different organisms. The N-terminus is believed to be unstructured. Its has not yet been characterized, but several hypotheses exist for its role in transcription initiation. While the core domain of TBP is sufficient to support transcription *in vitro* and *in vivo*, the N-terminus is thought to affect the structure and stability of the TBP-DNA

complex (Kuddus and Schmidt, 1993). At low temperatures, it has also been observed to decrease the association and dissociation rates of the TBP-DNA complex. Presence of the N-terminus also increases the already relatively high level of activation energy required for TBP-DNA complex formation (Kuddus and Schmidt, 1993). Further studies have shown the inability of core TBP (ΔN TBP) to respond to transcription activation by upstream regulators such as the human GC box binding factor, Sp1 (Peterson et al., 1990). Several models based on these findings have been proposed for the function of the N-terminus. One model suggests that when TBP is free in solution, the N-terminus is present in a conformation that interferes with the formation of a stable TBP-DNA complex (Kuddus and Schmidt, 1993). Kinetic studies that suggest a two-step pathway for TBP-DNA complex formation (Hoopes et al., 1992) also support the hypothesis that a conformational change in TBP may occur prior to stable complex formation. Additional models address the N-terminus' effect on transcription activity and DNA binding activity. The inability of core TBP to respond to Sp1 leads to the hypothesis that the N-terminus may play a role in the protein-protein interactions that regulate the transcription activity of TBP (Kuddus and Schmidt, 1993). Further, studies showing the enhanced DNA binding activity upon removal of all or part of the N-terminus by proteolysis or mutagenesis (Horikoshi et al., 1990; Lieberman et al., 1991) support a proposed function of the N-terminus in regulating the DNA binding activity of the core domain.

A compilation of these proposals leads to a working model of the N-terminus' function in transcription initiation. We hypothesize that when TBP is in solution, the N-terminus is tucked into the hydrophobic DNA binding region. A temperature-dependent conformational change that is required before DNA binding can take place (Lieberman et al., 1991) may possibly be the displacement of the N-terminus into solution upon DNA binding (Figure 3). Additional support for this model is supplied by data showing that the N-terminus is more susceptible to protease digestion when TBP is bound than when it is free in solution (Hermann and Hawley, 1999). The proposed conformation of the N-

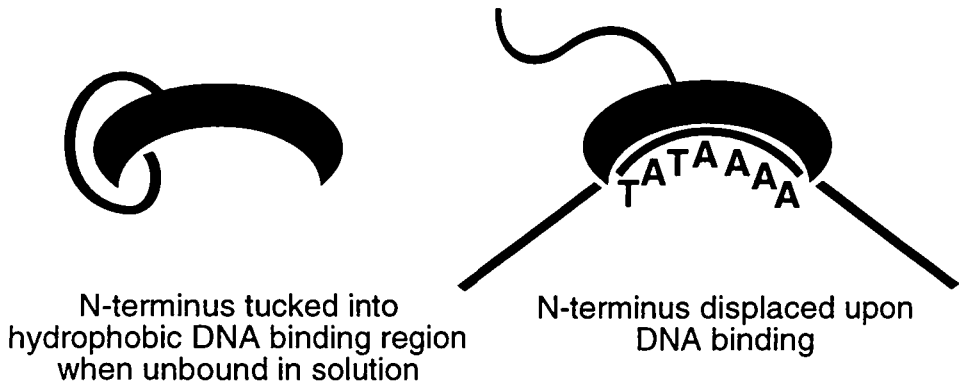

Figure 3. Model of N-terminus conformation when TBP is free (left) and when TBP is bound to DNA.

terminus is further supported by studies observing the fluorescence from a single tryptophan residue in the N-terminus of yeast TBP (yTBP) (Perez-Howard et al., 1995; Hermann and Hawley, 1999). The tryptophan of free TBP displays an emission maximum at 315 nM, which corresponds to the expected emission maximum for a tryptophan in a hydrophobic environment. When TBP binds to the TATA box, however, the emission maximum shifts to 345 nM, which corresponds to the expected emission maximum for a tryptophan in aqueous solution. Taken together, these data corroborate the proposed model of the conformation of the N-terminus in bound and free TBP.

In vitro behavior of TBP mutants

Altered specificity mutants of TBP have been examined to characterize the properties of TBP. In 1992, Strubin and Struhl conducted a genetic selection to isolate several altered specificity variants of yTBP able to promote transcription from a promoter containing the 30G TATA sequence, where the adenine in position two of the TATA sequence was changed to a guanine, giving the sequence TGTAAAAG. In our nomenclature, this sequence is called 30G because the base located 30 base pairs upstream

from the start site of transcription from the Adenovirus Major Late Promoter (AdMLP) has been changed to a G. This sequence is not recognized appreciably by wild-type TBP. The TBP from cells that survived the selection shared the same amino acid substitutions, Ile-194 to Phe (I194F), and Leu-205 to Val (L205V). These two substitutions were the minimum requirement for cell growth on 30G.

Previous studies have used the structure of TBP to predict how each of the mutations contributes to the protein's altered specificity (Kim et al., 1993a; Kim et al., 1993b). They came to the conclusion that wild-type TBP cannot bind to 30G because of steric hindrances between the leucine at position 205 and the guanine's amino group in the minor groove. Yet, the smaller side chain of L205V would remove this steric clash. The contribution of the I194F mutation is less clear from the crystal structure.

Since these studies provide a structural model to explain how the L205V mutation allows binding to 30G, an expectation may be that L205V alone is sufficient to confer binding to this altered sequence. Because both mutations are necessary *in vivo* (Strubin and Struhl, 1992), the L205V mutation may be able to bind 30G but may not be fully functional in transcription. Baxley and Hawley (1999) examined the binding ability of TBPs containing the I194FL205V mutation, the I194F mutation alone, and the L205V mutation alone on consensus TATA and 30G. In this experiment, the proteins were incubated with a labeled DNA fragment from the AdMLP that contained either consensus TATA or 30G. Reactions were then loaded onto a native polyacrylamide gel. I194FL205V TBP and I194F TBP were seen to bind to consensus TATA, while binding of L205V TBP to consensus TATA was barely detectable. The ability of I194FL205V TBP to bind 30G while the single mutants alone did not show binding (Baxley and Hawley, 1999) supports the requirement of both mutations in I194FL205V for binding to 30G *in vitro* (Baxley and Hawley, 1999). A further test for nucleotide recognition ability by the mutant TBPs was conducted with a TATA sequence called 25C that contained an altered nucleotide on the second half of the TATA box so that the sequence read: TATAAA**C**G.

I194FL205V TBP was found to bind significantly better than wild-type TBP to 25C (Baxley and Hawley, 1999). Taken together, these results lend support to the ability of the I194FL205V mutations in TBP to allow the protein to bind sequences that are altered at a distance from the site where the affected amino acids are thought to interact with the DNA (Baxley and Hawley, 1999).

As stated previously, the function of the N-terminus has not been determined. To learn more about the role of the N-terminus, we have conducted functional studies with N-terminal delete versions of the previously discussed TBP mutants.

In EMSA experiments, complexes between TBP and TATA sequences for which the TBP has low affinity cannot be detected because the complex dissociates quickly during electrophoresis. It has been shown, however, that dissociation can be slowed when the N-terminus is removed, thus allowing some weak complexes to be detected (Starr et al., 1995). Baxley and Hawley (1999) used an EMSA experiment to show an N-terminal delete version of L205V (ΔN L205V) bound significantly better than full-length L205V TBP to 30G and to consensus TATA.

The working model of the N-terminus' function in transcription initiation as presented earlier can be further clarified through studies using I194FL205V TBP. The position of the mutations in the hydrophobic DNA binding region of I194FL205V TBP may alter the conformation of the N-terminus when TBP is in solution, causing the N-terminus to be displaced even when TBP is not bound to DNA. The N-terminus has been proposed to influence the increased stability or binding ability of altered specificity mutants.

To further develop the model of N-terminus function, an N-terminal delete version of I194FL205V (ΔN I194FL205V) TBP was generated for *in vitro* EMSA studies. Full-length and core versions of wild-type TBP and L205V TBP were tested along with full-length I194FL205V TBP for their *in vivo* behavior in reporter gene studies using the budding yeast *Saccharomyces cerevisiae*. According to the proposed model, displacement of the N-terminus into solution occurs under all conditions in the full-length version of this

double mutant. ΔN I194FL205V lacks an N-terminus, and therefore, would be expected to behave similarly to the full-length version. The behavior of full-length and core versions of I194FL205V TBP should resemble that of wild-type ΔN TBP because ΔN TBP lacks an N-terminus as well. When wild-type full-length TBP and ΔN TBP behave differently, the difference may be attributed to a contribution from the N-terminus. Under these conditions, behavior of full-length and core I194FL205V TBP should be different compared to that of wild-type full-length TBP.

Data presented here show that the previously observed *in vitro* behavior of TBP variants can serve as a predictor of the functional characteristics of the protein *in vivo*. Although the contribution of the N-terminus to the protein's binding properties on a mechanistic level is not investigated here, it is apparent that the presence of the N-terminus alters the binding ability of various TBP mutants.

CHAPTER II

REMOVAL OF THE N-TERMINUS IS NOT SUFFICIENT TO ACCOUNT FOR ALTERED SPECIFICITY

Introduction

Transcription initiation by RNA polymerase II requires the formation of the preinitiation complex at the promoter of the gene. This complex contains transcription initiation factors and the polymerase (reviewed in Zawel and Reinberg, 1995). In preinitiation complex formation, the transcription factor TFIID binds to the promoter. TFIID is composed of TBP and TBP-Associated Factors (TAFs). It is TBP that recognizes and binds to the TATA sequence (consensus TATAAAAG) at the promoter (Zawel and Reinberg, 1995). Other transcription factors are then recruited to the promoter and finally, pol II binds.

TBP contains a conserved C-terminal domain and a variable N-terminal domain (Hernandez, 1993). The C-terminal or "core" domain consists of two imperfect repeats that fold into a semi-symmetric structure (Nikolov et al., 1992; Chasman et al., 1993; Kim et al., 1993a; Kim et al., 1993b; Nikolov et al., 1996). This structure has a saddle-shaped conformation (Kim, J. et al., 1993) where the hydrophobic convex bottom region of the protein is the DNA binding region. Most of the contacts made by TBP in the minor groove of DNA (Starr and Hawley, 1995) are hydrophobic contacts (Kim et al., 1993a; Kim et al., 1993b; Nikolov et al., 1996).

The N-terminal domain varies in length and sequence among different organisms. It is believed to be unstructured and its function is not yet fully understood. Previous studies have shown that ΔN TBP without the N-terminus can not only promote basal

transcription *in vitro* (Hoey et al., 1990; Horikoshi et al., 1990; Peterson et al., 1990; Lieberman et al., 1991), but can also participate in activated transcription *in vivo* when ΔN TBP is expressed as the only copy of TBP in the cell (Poon et al., 1991; Peterson et al., 1990).

Several hypotheses have been suggested for its effect on TBP behavior in transcription. The N-terminus is thought to decrease the association and dissociation rates of the TBP-DNA complex (Kuddus and Schmidt, 1993). Other studies support the belief that the N-terminus may play a role in the protein-protein interactions that regulate the transcription activity of TBP (Kuddus and Schmidt, 1993). Another group suggests the N-terminus functions in regulating the DNA binding activity of the core domain (Horikoshi et al., 1990; Lieberman et al., 1991). Based on these data and experiments conduced in our lab, we hypothesize that when TBP is in solution, the N-terminus is present in a conformation where it is tucked into the hydrophobic DNA binding region. The conformation of the N-terminus must change before TBP can bind DNA (Lieberman et al., 1991). The observed conformational change (Lieberman et al., 1991) may be the displacement of the N-terminus into solution upon DNA binding. Studies of the N-terminus using tryptophan fluorescence and protease digestion (Hermann and Hawley, 1999) provide support for N-terminus displacement upon binding.

The behavior of the altered specificity double mutant I194FL205V TBP may provide insight into the effect of the N-terminus on TBP behavior. The amino acid mutations at positions 194 and 205 in the hydrophobic DNA binding region of the double mutant may alter the interaction of the N-terminus so that the N-terminus is displaced into solution at all times. By testing the properties of the double mutant and an N-terminal delete version of the double mutant (ΔN I194FL205V TBP), we can observe whether the mutations alter the effect of the N-terminus on TBP behavior.

Results and Discussion

We performed an EMSA experiment to determine the N-terminus' contribution to TBP's ability to bind DNA sequences. In this assay, TBP is incubated with a labeled DNA fragment from the AdMLP containing the indicated TATA sequence. The reaction mix is then loaded onto a native polyacrylamide gel containing magnesium. An electric current is applied to the gel in a process called electrophoresis. In electrophoresis, negatively-charged DNA fragments migrate through the gel towards the positive charge. The nature of the gel causes larger complexes to move more slowly than smaller complexes. In this way, larger complexes of TBP bound to DNA move more slowly and are thereby separated from smaller, faster-moving DNA fragments without protein bound. Autoradiography can then be used to observe TBP binding to various TATA sequences from the intensity of the signal emitted by the labeled DNA fragments. Figure 4 shows full-length and core versions of wild-type TBP binding with high affinity to the consensus sequence. Full-length and core versions of the double mutant, however, were seen to bind with little specificity for and high affinity to all sequences. Binding specificity behavior between the full-length and core versions of wild-type TBP was similar. Behavior of full-length and core versions of the double mutant did not differ significantly in the presence or absence of the N-terminus. If the position of the N-terminus were important in DNA recognition, both versions of the double mutant should behave like core TBP. Based on these observations, removal of the N-terminus is not sufficient to account for binding to non-consensus TATA boxes. Taken together, these observations support the view that the N-terminus does not confer binding specificity properties.

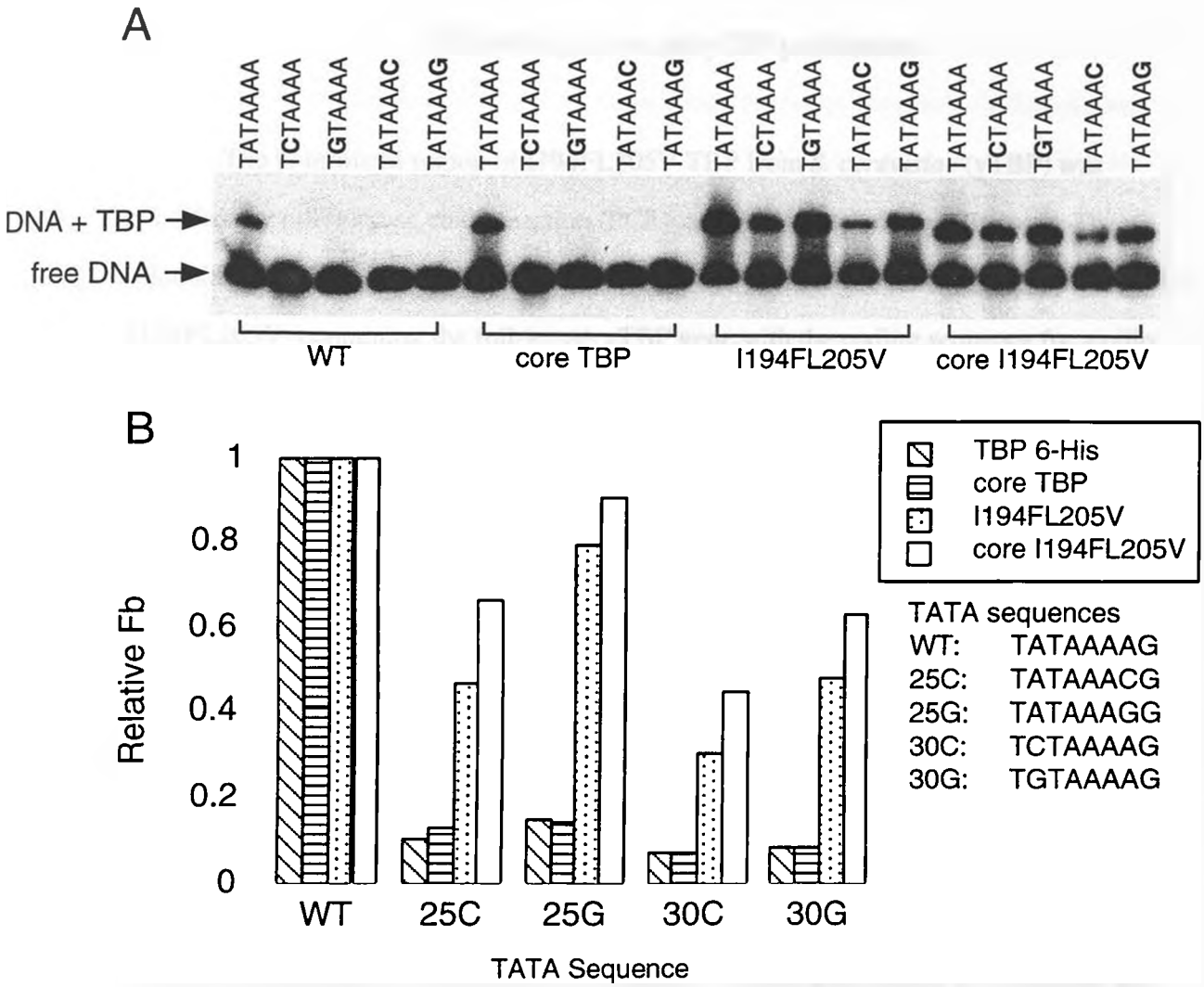

Figure 4. Full-length and N-terminal delete versions of I194FL205V TBP bind TATA sequences with high affinity. (A) Electrophoretic Mobility Shift Assay (EMSA) experiment where full-length and N-terminal delete versions of wild-type TBP and I194FL205V TBP were incubated with ^{32}P-labeled HindIII-EcoRI fragment from AdMLP containing either consensus TATA, 25C, 25G, 30C, or 30G. Reactions were performed as described in the experimental procedures. (B) The fraction of DNA bound to TBP was calculated as described in the experimental procedures and normalized to the value of each TBP for the consensus TATA-box for each TBP tested.

Experimental Procedures

DNA Manipulation and yTBP purification

The C-terminal region of I194FL205V TBP from *S. cerevisiae* (yTBP) was modified by polymerase chain reaction (PCR) using oligonucleotides (Genosys, The Woodlands, TX) in the concentration of 50 μM/100 μL reaction mix. Plasmid pTZyD I194FL205V, containing the full-length yTBP gene with the coding sequence for amino acids at positions 194 and 205 changed from ATC to TTC and TTA to GTA respectively, was used as the DNA template. An NdeI restriction site was introduced outside the core coding region so the sequence of the encoded protein was not altered. This resulted in NdeI and BamHI restriction sites flanking the coding region. PCR products were purified using QIAGEN's QIAquick PCR purification kit. PCR fragments were subcloned into expression vector pET15b between NdeI and BamHI recognition sites, replacing wild-type yTBP sequence with ΔN TBP gene sequence. pET15b contains sequence for six histidine residues N-terminal to the yTBP gene. Both fragment and vector were then digested with NdeI for two hours at 37 °C, adding 2 μL NdeI every 30 minutes, and incubated overnight. Two-hour digestion with BamHI at 37 °C followed a phenol:chloroform extraction and ethanol precipitation of NdeI digested fragments. Digests were run on a 1% agarose gel and purified with QIAGEN's QIAEXII Get Extraction Kit. Overnight ligation of pET15b and BamHI-NdeI fragment was transformed into DH5α, a strain of competent *Escherichia coli*, plated on LB-H solid agar medium containing 4x ampicillin (amp), and incubated overnight at 37 °C. Colonies were miniprepped and DNA sequence analysis used to identify candidates containing the correct incorporation of the truncated yTBP gene.

Cloned yTBP was overexpressed in and purified from *E. coli* strain BL21. In brief, one liter of cells were grown in LB-H Broth (5 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone, 1M 1N NaOH) and induced for 3 hours with 500 mM IPTG (isopropyl-ß-D-

thiogalactopyranoside) at 37 °C. Cells were pelleted at 5K for 10 minutes in a Beckman JA-20 Centrifuge and resuspended in 30 mL LB100 (20 mM Tris [Tris(hydroxymethyl)aminomethane] [pH 7.9] at 4 °C, 10% glycerol, 100 mM NaCl, 20 mM 2-mercaptoethanol (ßME), 1 mM phenylmethylsulfonyl fluoride). Cell lysate was extracted by French Press followed by separation from cellular debris by centrifugation at 11K for 40 minutes at 4 °C. The supernatant was dialyzed at 4 °C overnight into nickel binding buffer (20 mM Tris [pH 7.9] at 4°, 10% glycerol, 5 mM Imidizole, 500 mM NaCl) and loaded onto a 6 mL nickel (Ni^{2+}) column that had been equilibrated according to the manufacturer's instruction. The column was washed with nickel wash buffer (identical to nickel binding buffer but with 60 mM imidizole) and yTBP bound by interactions between histidine residues and Ni^{2+} ions was eluted with nickel elution buffer (identical to nickel binding buffer but with 500 mM imidizole and containing 10 mM ßME). Recovered yTBP in flow-through was dialyzed for 4 hours each against 500 mL LB500 Dialysis buffer (20 mM Tris [pH 7.9] at 4 °C, 20% glycerol, 500 mM NaCl, 10 mM ßME) at 4 °C. Samples of purified protein were loaded on a 12% denaturing sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel [4 mL 30% acrylamide:8% bis-acrylamide, 3.7 mL 1M Tris [pH 8.8], 100 µL 10% SDS, 2.2 mL H_2O, 100 µL 20% APS (ammonium persulfate), 5 µL TEMED (N, N, N', N'-Tetramethylethylenediamine)] with a 5% stacker (670 µL 30% acrylamide:8% bis-acrylamide, 1.85 mL Tris [pH 6.8], 50 µL 10% SDS, 2.4 mL APS, 5 µL TEMED) and electrophoresed in 1x SDS-PAGE Running Buffer (10x stock in 1 L: 30 g Tris, 144 g glycine, 100 mL 10% SDS) at 25 mA. Protein bands were visualized by the Coomassie staining procedure.

DNA Binding Experiments

Binding ability of wild-type TBP, WT ΔN TBP, full length I194FL205V, and ΔN I194FL205V was assayed under conditions of altered TATA-box sequences and varying

temperature. Labeled DNA fragments containing the TATA sequence were made by initially cutting at the HindIII site in pWRM18 (described in Hoopes *et al.*, 1992) and end-labeling with the Klenow fill-in reaction using α-^{32}P dATP. A 128-bp fragment was generated by cleaving at the EcoRI site and was isolated following separation in a 6% polyacrylamide gel (4 mL 30% acrylamide: 8% bis-acrylamide, 4 mL 5x TBE (54 g/L Tris base, 27.5 g/L boric acid, 20 mL/L 0.5M EDTA [pH 8.0]), H_2O to 20 mL, followed by 133 µL APS and 12 µL TEMED). yTBP was added to reaction mix containing $BC100_{0.1}$ (20 mM Tris [pH 7.9] at 4° with 10 µL 10 mg/mL bovine serum albumin and 1 µL 1M dithiothreitol, 20% glycerol, 0.2 mM EDTA (ethylenediamine tetraacetic acid), 100 mM KCl), TE (10 mM Tris-HCl [pH 8], 1 mM EDTA), and MHD (500 mM $MgCl_2$, 1M HEPES (4-(2-Hydroethyl)-1-piperazineethanesulfonic acid), labeled DNA fragment containing one of the following sequences: consensus TATA, 30G TATA, 25G TATA, or 25C TATA, and TE and incubated for 30 minutes at 30 °C. 1.5 µL of 0.2 mg/mL competitor poly [d(I-C)] was added followed by a 30 second incubation at 30 °C. Reaction mix was loaded onto a bandshift gel (7.5 mL 40% acrylamide, 7.5 mL 10x TGE (250 mM Tris base, 1.9 M glycine, 10 mM EDTA [pH 8.5]), 15 mL 50% glycerol, 750 µL 500 mM $MgCl_2$, H_2O to 75 mL, followed by addition of 37.5 µL 1M DTT, 500 µL APS, 44 µL TEMED), which had been prerun at 30 mA for 30 minutes, and electrophoresed in TGEM buffer (1x TGE with 500 mM $MgCl_2$) for 90 minutes at 30 mA. Following electrophoresis, the gel was dried onto Whatmann paper and exposed to film. Signals were detected by autoradiography using Molecular Dynamics' (Sunnyvale, CA) Storm 860 phosphoimager. The amount of complex formation was determined in ImageQuant version 1.2 by drawing boxes around the bands corresponding to bound and free DNA. Fraction bound was determined from: Fb = (counts bound)/(counts bound + counts free). Data were graphed in KaleidaGraph data analysis/graphing application by Synergy Software.

CHAPTER III

PHYSIOLOGICAL RELEVANCE OF *IN VITRO* BEHAVIOR

Introduction

TBP's role in transcription initiation has earned it the name of the "commitment factor" because its binding to the TATA sequence at the promoter nucleates the formation of the preinitiation complex (Zawel and Reinberg, 1995), and multiple rounds of transcription may occur before TBP dissociates from the DNA.

The identification and subsequent study of mutant TBPs that display altered binding properties have facilitated the characterization of TBP's nucleotide recognition and binding mechanisms. In a study of yTBP, Strubin and Struhl (1992) randomly mutagenized a region of a plasmid-borne TBP gene that they believed to be important for DNA binding. The plasmid was transformed into a strain of yeast in which the TATA sequence preceding the HIS3 gene had been changed at the second position so that it read TGTA (30G). Cells were grown in selective medium that lacked the amino acid histidine. Because wild-type TBP cannot recognize 30G, cells can grow only if they contain a TBP mutant that can bind 30G and nucleate preinitiation complex formation. This selection produced a TBP containing two mutations (I194FL205V). Each of the mutations alone could not promote growth on this altered TATA sequence.

The contributions from each mutation have been studied in our lab. Baxley and Hawley (1999) have reported the *in vitro* behavior of full-length and N-terminal delete versions of I194FL205V, L205V, and wild-type TBP. In EMSA experiments, both I194FL205V TBP and wild-type TBP showed binding to the consensus TATA sequence.

Binding by L205V TBP was faint. However, if the N-terminus was removed, ΔN L205V was seen to bind better to both consensus TATA and 30G (Baxley and Hawley, 1999). Of the other mutants tested, only I194FL205V TBP was able to bind detectably to 30G.

These *in vitro* observations provide a framework for understanding the effect of each mutation. Yet, conducting experiments *in vitro*, or outside of the cell, poses some concerns regarding the physiological relevance of observations. Most conditions in which *in vitro* experiments are conducted vary greatly from the biological environment. For example, the strength of a chemical reagent, concentration of protein, or amount of DNA in the reaction may alter the behavior of the observed protein and produce artifacts. With these considerations in mind, we conducted the following experiments to test if TBP's behavior *in vivo* correlated with TBP's observed behavior *in vitro*.

Results and Discussion

To further develop the understanding of TBP's binding specificity, we conducted reporter gene assays in *S. cerevisiae* to examine the ability of altered TBPs to function in promoting transcription from altered TATA sequences. The expression system in *S. cerevisiae* allows the use of plasmid-borne genes whose expression can be regulated (Figure 5). Yeast constructs used in this assay contain two plasmids, both regulated by galactose, a carbon source for the cell (see Experimental Procedures). pSc3801 contains a TATA sequence in the promoter of a reporter gene called LacZ. LacZ functions as a reporter gene because its expression can be observed through the activity of its protein product, ß-galactosidase (ß-gal). ß-gal is an enzyme that cleaves substrates to yield a color change. pCEF1A contains the TBP gene and a galactose-inducible promoter that can control the expression of the gene. In the presence of galactose, the expression system is considered to be "induced." Induction refers to the "turned-on" expression of the plasmid-borne TBP gene. The TBP protein can then bind the TATA sequence located on pSc3801

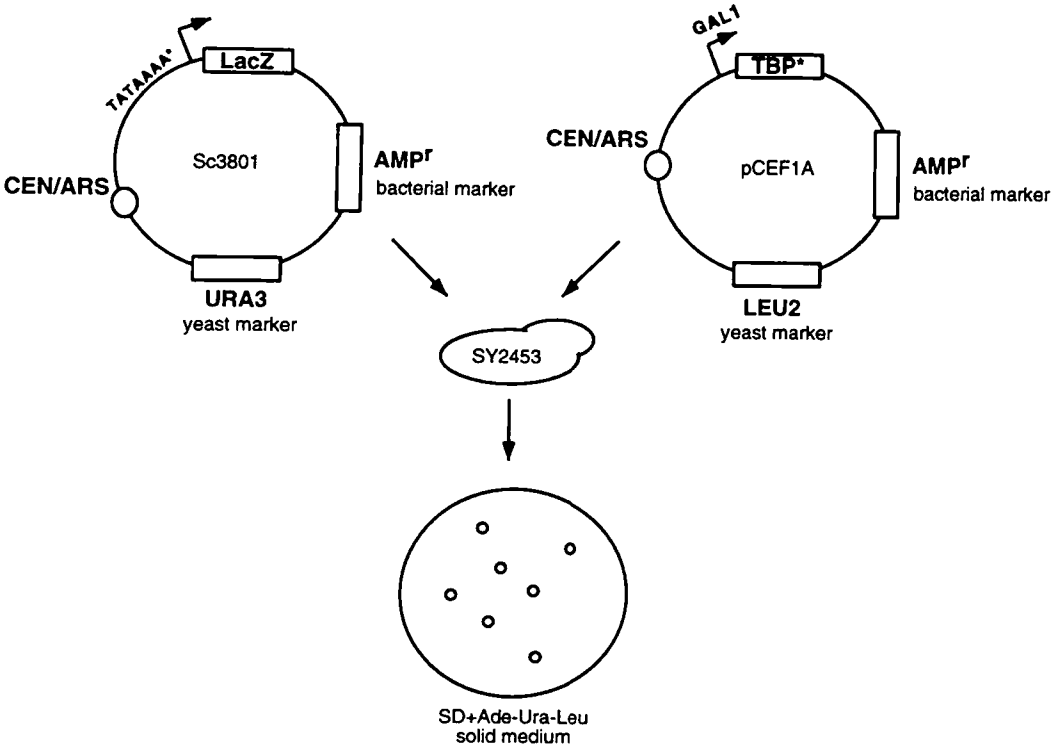

Figure 5. *In vivo* reporter gene expression system in *S. cerevisiae* strain SY2453.

to induce the expression of LacZ (Figure 6). The induction is stopped after a three hour period and the cells broken open for the next reaction. In this step, the cells are exposed to a compound called ONPG (see Experimental Procedures), which turns yellow upon cleavage by ß-gal. The intensity of color change can be quantitated by spectrophotometry.

Figure 7 depicts the ß-gal activity of toluenized cells following growth on different media. Following a two-day growth period in raffinose-containing medium (which neither induces nor represses transcription from the GAL1 promoter), background levels of ß-gal activity can be detected. Transfer from raffinose-containing medium to glucose-containing medium (which represses transcription from the GAL1 promoter and inhibits transcription factor binding to the GAL1 promoter that contains the TATA-sequence in Sc3801) resulted in ß-gal activity levels observed at or below background levels. Cells transferred into galactose-containing medium exhibited elevated levels of ß-gal activity.

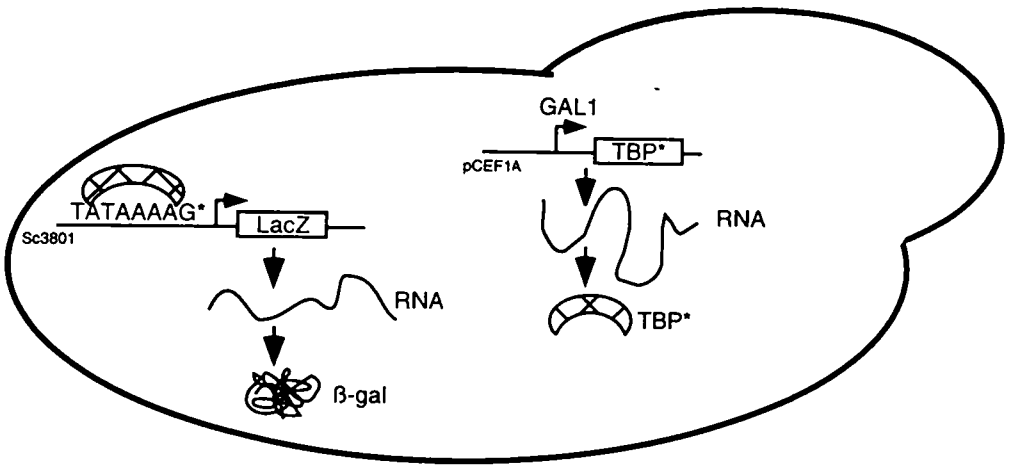

Figure 6. Representation of induced expression system where the presence of galactose induces the expression of the pCEF1A TBP gene. The TBP protein is then able to bind to the TATA-box in Sc3801 to activate expression of the LacZ gene. Activity of ß-gal can then be assayed.

Four TBP mutants, ΔN TBP, L205V, ΔN L205V, and I194FL205V, along with wild-type TBP, were tested for their function on three TATA sequences: consensus TATA, 25C, and 30G. Background levels of relative ß-gal activity from consensus TATA-sequence are shown in Figure 7a. These levels were calculated by normalizing ß-gal activity to ß-gal activity under activated conditions from the pGALLEU control construct which does not contain a TBP gene. The simultaneous presence of overexpressed copies of plasmid-borne TBP as well as endogenous TBP did not appear to alter the amount of reporter gene product expressed. This trend was consistent over the four trials conducted.

In vivo binding specificity for 25C showed a noticeable (2.5-3-fold) increase in ß-gal activity in constructs where I194FL205V TBP was overexpressed (Fig 7b). Relative ß-gal activity of constructs containing the remaining TBP mutants remained at background levels when normalized to ß-gal activity from the pGALLEU construct under conditions of activation. These *in vivo* findings correlate with the observed *in vitro* behavior of I194FL205V TBP which has been seen to bind at higher levels than wild-type TBP to 25C (Baxley and Hawley, 1999).

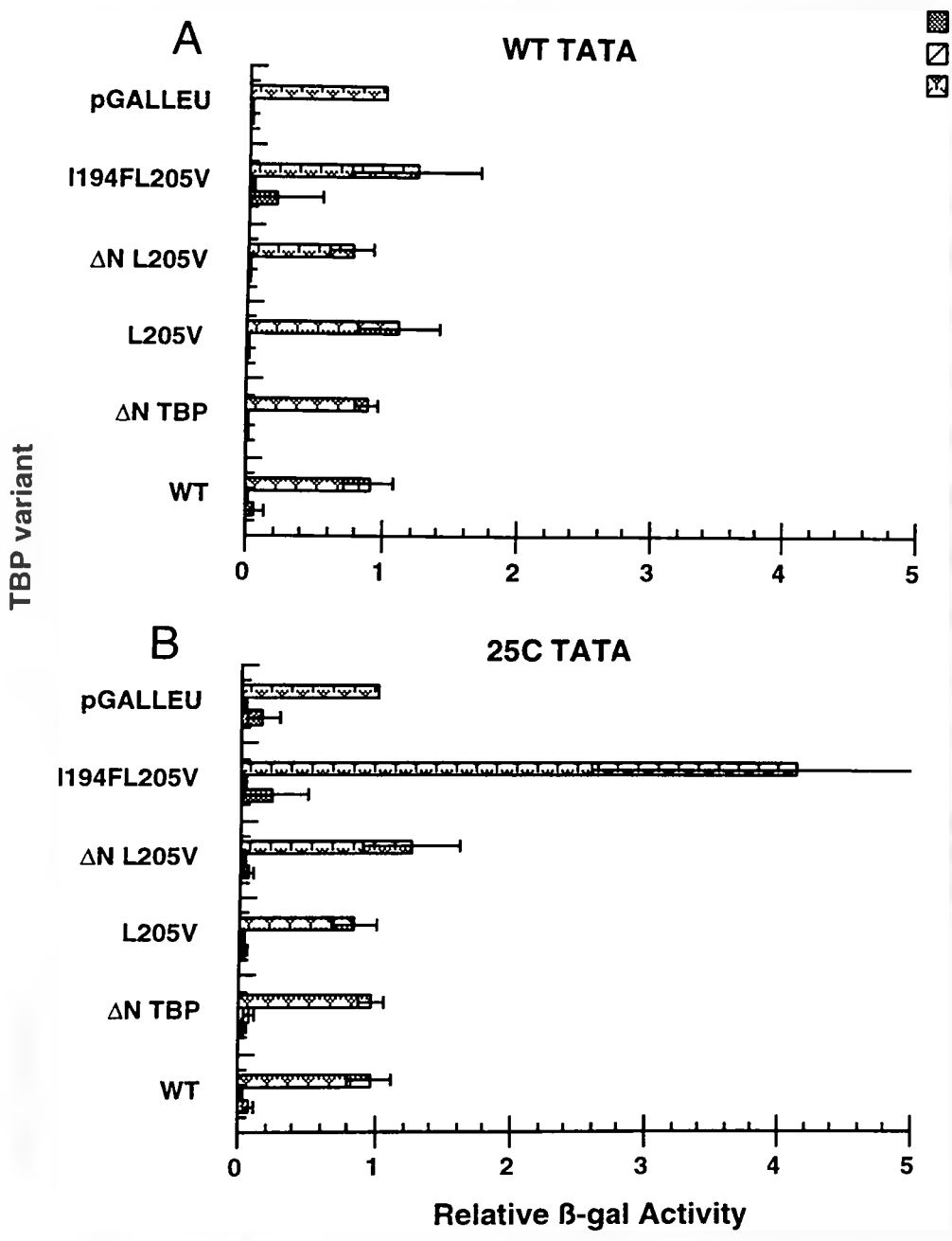

Figure 7. ß-galactosidase activity from constructs containing (A) consensus TATA and (B) 25C TATA. Cells were grown for two days in selective raffinose-containing liquid media and then diluted 1:100 in rich media and incubated for two hours at 30 °C. Two 2-mL aliquots were each grown for three hours in galactose (inducing expression) or glucose (repressing expression). Membranes were broken open with sarkosyl and toluene, and then exposed to a reaction mix containing ONPG. Reactions were quenched with 500 µL 1M Na_2CO_3 when a yellow color change was evident. ß-gal activity was determined by spectrophotometry at OD600.

Prior to normalization, data from constructs containing 30G showed high amounts (33-68 Miller Units) of ß-gal activity under conditions of I194FL205V TBP overexpression; overexpression of wild-type, ΔN TBP, and L205V TBP yielded ß-gal activity less than 7 Miller Units (data not shown). The behavior of ΔN L205V ranged widely among the eight trials conducted. To better compare the ability of TBP variants to function on the 30G TATA sequence, the behavior of I194FL205V TBP was established as an internal standard. The double mutant was chosen because it exhibited the most consistent behavior among TBP variants tested and functioned significantly above background levels. As shown in Figure 8, the behavior of ΔN L205V TBP relative to I194FL205V TBP can be seen to vary markedly, inducing the expression of low and higher levels of ß-gal (error bars, Fig. 8). Relative ß-gal activity in constructs overexpressing L205V, ΔN TBP, and wild-type TBP remained below 0.12 units (Fig. 8). The higher affinity of I194FL205V TBP for 30G as seen in these *in vivo* results correlates

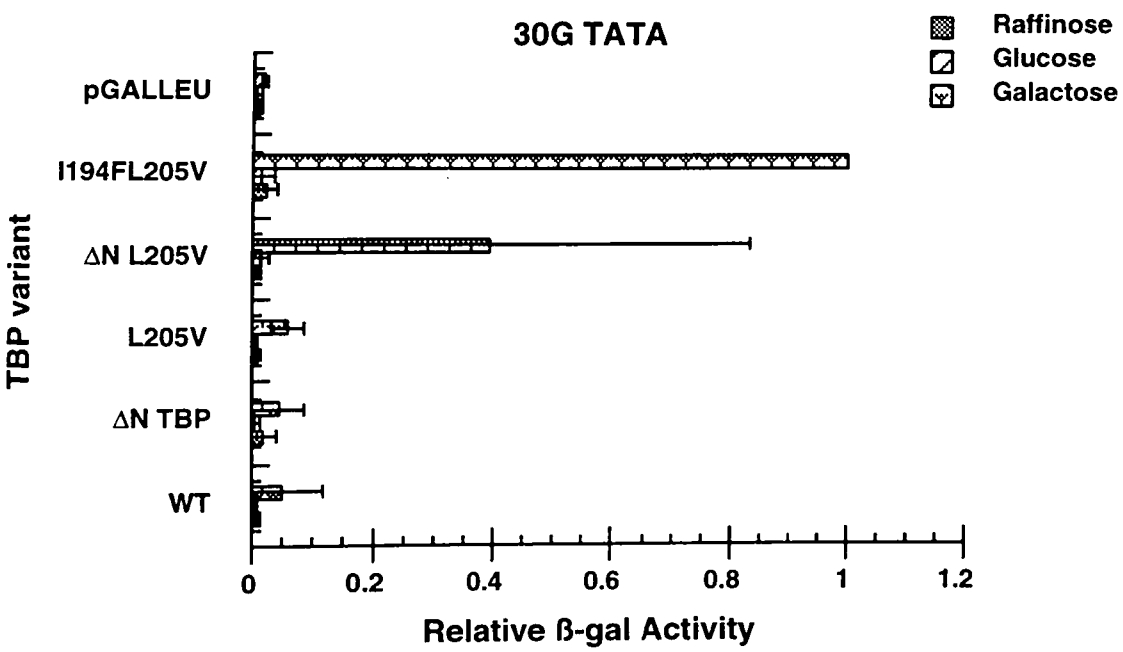

Figure 8. ß-galactosidase activity from a 30G-containing construct. Cells were induced for three hours in galactose. Membranes were broken open with sarkosyl and toluene and cells exposed to a reaction mix containing ONPG. Reactions were quenched with 500 µL 1M Na_2CO_3 when a yellow color change was evident. ß-gal activity was normalized to levels of ß-gal activation from induced I194FL205V TBP.

with *in vitro* observations that I194FL205V binds better than L205V TBP or wild-type TBP to the 30G TATA sequence (Strubin and Struhl, 1992; Baxley and Hawley, 1999).

Dosage-dependent effects were explored in reporter gene assays conducted on selective solid medium. Colonies were initially patched onto a master plate on selective synthetic glucose-containing medium and grown for two days at 30 °C. They were then transferred by replica plating onto a nitrocellulose filter placed on synthetic raffinose-containing medium and grown at 30 °C for 24 hours. The filter was then transferred onto synthetic galactose-containing medium and incubated at 30 °C for 12-24 hours. Exposure

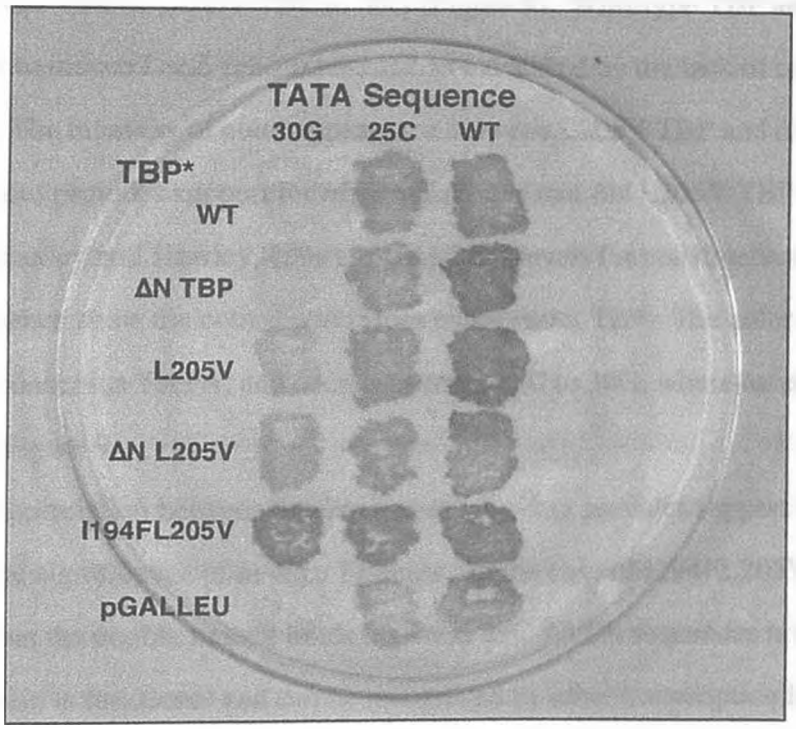

Figure 9. *In vivo* reporter gene expression system on selective solid medium. Cells were induced for 24 hours in galactose-containing selective medium, lysed in liquid nitrogen, and exposed to X-gal for 40 minutes.

to the ß-gal substrate X-gal results in a color change that allows qualitative ordering of the specific TBP variant's affinity for an altered TATA sequence. A representative picture is shown in Figure 9. A previous trial in which cells were grown in raffinose-containing medium for two days before a two-day induction showed the loss of viability of cells overexpressing ΔN TBP (data not shown). Overexpressed ΔN TBP may interfere with

cellular growth processes when present with endogenous TBP over time. Cells containing I194FL205V, ΔN L205V, and L205V TBP mutants were all viable on galactose-containing medium. Cells containing consensus TATA all underwent a color change within the 40-minute reaction time, with cells containing pCEF1A (wild-type TBP) and pGALLEU appearing blue within 10 minutes (data not shown). Cells containing 25C exhibited a color change at a slower rate among all the constructs tested. At the end of the reaction, I194FL205V TBP showed the most intense color. Observations from 30G constructs provide functional support for *in vitro* observations that I194FL205V TBP can bind better than wild-type TBP or L205V TBP to 30G (Figure 9). Wild-type TBP and ΔN TBP are both unable to induce LacZ gene expression as evidenced by the lack of color change (Figure 9). The intensity of color appearance between L205V TBP and ΔN L205V TBP (Figure 9) also provides support for *in vitro* findings that ΔN L205V TBP is better able to bind 30G (Baxley and Hawley, 1999). pGALLEU levels for constructs containing each TATA sequence show the contribution from endogenous TBP. The color intensity is greatest in consensus TATA, and decreases from 25C to 30G, where no color change is detectable (Figure 9).

The correlation between *in vitro* and *in vivo* data provides support for the physiological significance of *in vitro* findings. In the case of I194FL205V TBP, *in vivo* data show that the double mutant binds 30G and 25C TATA sequences in a manner such that the protein is functional and can be recognized by other transcription factors. Experimental results from *in vitro* and *in vivo* studies can, therefore, be a direct indication of TBP's physiological behavior.

Experimental Procedures

Yeast Methods

Standard methods were used for growth, transformation, and genetic manipulation of *S. cerevisiae*. Rich medium (YEP-Dextrose) and synthetic medium (synthetic dextrose [SD], synthetic raffinose [S-raff], and synthetic galactose [S-gal] were prepared as described.

Cloning of Altered TATA sequences and yTBP Altered Specificity Mutants

Yeast TBP expression vectors for ΔN TBP, L205V, ΔN L205V, and I194FL205V, were made by L. Baxley, H. Kemp, and S.M.C. using PCR mutagenesis on yeast expression vector pCEF1A (C. Foulds, unpublished data) to generate altered DNA sequence containing the substituted amino acid residue. pCEF1A contains a CEN/ARS (centromere/autonomously replicating sequence) element which ensures the propagation of one plasmid per cell, the LEU2 gene for yeast maintenance, amp resistance (amp^r) gene for bacterial maintenance, and full-length HA- (hemagglutinin-) tagged yTBP gene flanked by Blp1 and Xba1 recognition sites positioned behind a galactose (gal) inducible promoter. pGALLEU (LEU2) plasmid contains a galactose-inducible promoter driving non-coding sequence. Promoter constructs containing a 100 bp TATA-box fragment in Sc3801 (Singer et al., 1990) were made by L. Baxley using enzyme digestion to remove the TATA fragment, kinasing and hybridizing synthesized oligonucleotides to incorporate altered TATA sequence, followed by ligation and transformation. Sc3801 contains a CEN/ARS element, URA3 gene, amp^r gene, and TATA-box promoter upstream of the LacZ reporter gene. Candidates were screened through restriction digests and sequencing.

In vivo constructs were made by H. Kemp using The Boss Lithium-Acetate method of yeast transformation (C. Boone and J. Horecka, 1992) into a competent yeast strain SY2453, which has a mutated mating locus, and is unable to synthesize amino acids leucine (LEU) and histidine (HIS) and the nucleotide uracil (URA). Yeast colonies were maintained on YEPD (20 g/L peptone, 10 g/L yeast extract, 40 mL/L 50% glucose) media supplying all required amino acids except leucine and/or the nucleotide uracil. Cells viable on this medium must be able to synthesize their own leucine and/or uracil from the gene(s) provided by the plasmid(s). Viable cells were replica-plated onto SD + Ade - Ura - Leu - His plates to select against prototrophic phenotypes.

Phenotypic analysis

Yeast colonies for X-gal filter assays were replica-plated from SD + Ade - Ura - Leu master plates onto nitrocellulose filter (Millipore, pore size 0.45 µM) and placed on S-Raff + Ade -Ura -Leu (20 g/L Bifco agar, 6.7 g/L yeast nitrogenous base, 800 mL/L H_2O, 100 mL/L 2% raffinose, 10 mL/L 2 mg/mL adenine, 100 mL/L dropout) solid media and grown overnight at 30 °C. Filters were then removed to S-gal (pure) + Ade -Ura -Leu plates and incubated at 30 °C for 24 hours. ß-gal activity was visualized by exposing lysed cells on filters previously placed in liquid nitrogen for 30 seconds to Whatmann paper soaked in 1x Z-Buffer + ßME (100 mM $NaPO_4$ [pH 7.0], 10 mM KCl, 1 mM $MgSO_4$, 50 mM ßME) and 300 µL X-gal and incubated at 30 °C for 30-60 minutes. Cells for ß-gal assays were grown for 48 hours at 30 °C in 10 mL S-Raff + Ade -Ura -Leu (50 mL/L 2xS Broth (13.4 g/L yeast nitrogenous base), 4% raffinose, 1x-Leu-Ura dropout, 10 mL/L adenine). YEP-4% Raffinose cultures were inoculated 1:100 with overnight culture and incubated at 30 °C for 2 hours. A 2 mL aliquot of cells was induced with 2% galactose and another aliquot repressed with 2% glucose preceding 30 °C incubation for 3 hours. Cells were prepared by resuspending pellets in Z-Buffer + ßME, from which a 50 µL

resuspension was added to 1 mL SF (0.15 M NaCl, 3.7% formaldehyde) for cell density measurements using a Beckman DU-40 Spectrophotometer at reading OD_{600}. Toluene, which was evaporated at 37 °C for 30 minutes prior to assaying ß-gal activity, and 5% sarkosyl were added in 2 µL aliquots each to break open the remaining cells. Small test tubes containing 1 mg/mL ONPG (2-Nitrophenyl-ß-D-galactopyranoside) in Z-Buffer and 0.5 mL Z-Buffer minus the volume to be assayed were pre-warmed in a 28 °C water bath. Assay amount of 50 µL toluenized cells was added to each reaction tube, and duration of time from reaction start to stop, as indicated by medium-light yellow appearance of solution or elapse of 120 minutes, was recorded. Reactions were terminated with 0.5 mL 1M Na_2CO_3, and supernatent from pelleted cells was measured at OD_{420} for activity determination. Modified Miller units were calculated with the following equation:

$$\text{Units} = \frac{1000 \times OD_{420} \times d}{t \times v \times OD_{600}}$$

where t is time of reaction in minutes, v equals volume of cells assayed in mL, and d equals dilution factor used for cell density determinations. OD_{420} is the density of the reaction supernatent, and OD_{600} is the density of the cells added to the assay. Microsoft Excel 4.0 computer spreadsheet program was used for calculations.

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

Understanding and characterizing the ability of a protein to recognize its substrate and bind it preferentially requires a knowledge of the mechanism(s) essential to substrate recognition and the amino acid residues of the protein primarily involved in protein-DNA interactions. In the case of TBP, a DNA binding protein which preferentially binds the consensus TATA sequence, the limited number of contacts available for nucleotide recognition in the minor groove of DNA may raise the question of TBP's ability to function at other sequences that also offer limited contacts for recognition and binding. Studies presented here show that multiple amino acid substitutions in the second half of the DNA binding region of the protein confer altered specificity in nucleotide recognition and binding. While one amino acid substitution alone has not yet been identified to confer functionality in binding or transcription of a reporter gene above background levels, removal of the N-terminus in conjunction with the single amino acid substitution has been observed *in vitro* to increase binding at various altered TATA sequences (Baxley and Hawley, 1999) and *in vivo* to increase transcription from these sequences. The C-terminal domain of TBP is sufficient for transcription *in vitro* (Hoey et al., 1990; Horikoshi et al., 1990; Peterson et al., 1990; Lieberman et al., 1991) and *in vivo* (Poon et al., 1991; Peterson et al., 1990). Removal of the N-terminus in a TBP variant where the leucine at position 205 has been changed to a valine (ΔN L205V TBP) may result in the protein being able to form a more stable TBP-DNA complex. *In vitro* studies suggest that ΔN L205V functions significantly better than full-length L205V on 30G and consensus TATA (Baxley and Hawley, 1999). More conclusive physiological support for ΔN L205V TBP's *in vivo* behavior requires additional experiments.

Because removal of the N-terminus is not sufficient to account for TBP's altered specificity, the amino acid residues in the DNA-binding region of the C-terminal domain must be sufficient for nucleotide sequence recognition. The N-terminus may function downstream of DNA binding, either in recruiting transcription factors and/or in stabilizing or destabilizing the TBP-DNA complex following formation of the preinitiation complex and transcription initiation. The ability of the core domain of TBP to function in activated transcription *in vivo* has been shown in previous studies. However, when ΔN TBP is overexpressed in a cell containing endogenous TBP, ΔN TBP expression has been observed to lead to a dominant lethal phenotype.

To better design a model to explain this phenotype, additional experiments are required. Reporter gene assays with a yeast strain that contains a chromosomally-integrated version of ΔN TBP in place of the full-length endogenous protein will allow the examination the ß-gal activity from consensus TATA sequences in constructs containing a plasmid-borne copy of pGALLEU, ΔN TBP, or WT TBP. ß-gal levels from pGALLEU constructs will indicate the ability of endogenous ΔN TBP to function on the consensus TATA sequence. This level can be compared to the ß-gal activity in cells expressing plasmid-borne ΔN TBP. If these cells exhibit a lethal phenotype, interactions solely between ΔN TBP and WT TBP cannot account for the previously observed lethality in constructs expressing endogenous TBP and plasmid-borne ΔN TBP. The ß-gal activity of constructs overexpressing WT TBP is also important in developing our understanding of ΔN TBP's behavior. The presence of plasmid-borne WT TBP will provide information regarding ΔN TBP and WT TBP interaction. The phenotype of this construct, in conjunction with other findings, will likely shape the direction of further studies. For example, western blot analyses of constructs will allow for steady-state protein level analyses of endogenous TBP and plasmid-borne TBP using anti-yeast TBP and HA-antibody respectively. Data from these analyses can provide information regarding

expression levels within the cell. This information may be useful in exploring the contribution of dosage-dependent effects on phenotype.

The presence of a second amino acid substitution, where isoleucine is replaced by a phenylalanine at position 194 (I194F), in addition to L205V generates the double mutant I194FL205V that exhibits a loss of binding specificity. I194FL205V functions within background levels of transcription on consensus TATA, but causes increased transcription from altered TATA sequences of 25C and 30G. These altered TATA sequences test TBP's ability to recognize nucleotides on either half of the TATA-box. 25C contains a nucleotide substitution on the downstream, second half of the TATA-box. This half is contacted by the first repeat of TBP, while it is the second repeat of TBP that contains the amino acid substitutions. The ability of the double mutant to function at elevated levels as compared to other TBP variants on 25C supports the belief that I194F and L205V confer greater stability to the TBP-DNA complex. I194FL205V also functions better on 30G where steric interactions may be improved between the residues of the amino acid substitutions in the second repeat of TBP and the altered nucleotide sequence in the first half of the TATA-box.

The range of ß-gal activity from all TBPs tested on consensus TATA sequence fell within background levels of activation by endogenous TBP. This may indicate that the overexpression of plasmid-borne TBP does not significantly alter the level of LacZ gene expression. Such similar levels of ß-gal activity are not observed in 25C TATA or 30G TATA constructs. The inability of wild-type TBP to function significantly on these altered TATA sequences allows for observation of altered TBP function.

Further insight into binding specificity properties of TBP can be obtained through binding assays using altered specificity variants of TBP that contain analogous amino acid substitutions in the first repeat of the protein. Observations from these TBP variants may be combined with *in vitro* trends to advance the understanding of TBP's binding specificity. Development of a model for the mechanism of nucleotide recognition and binding by TBP can indeed further the understanding of TBP's function in promoter

recognition and binding. Utilization of this model may prove useful in the regulation of proliferating cells.

GLOSSARY

basal: in transcription: the fundamental, or basic, level of transcription from general transcription machinery.

carboxy terminus: in proteins, region of amino acids that ends in a carboxy (-COOH) group. The carboxy terminus varies in number and type of amino acids for individual proteins.

column affinity chromatography: method of separation in which a mixture of substances in the mobile phase is passed over a porous solid matrix called the stationary phase. Interactions between the mobile and stationary phases cause molecules to migrate at different rates.

complementary: complementary bases in both RNA and DNA exhibit the ability to pair through intermolecular interactions that cause two bases to adhere to each other. In DNA, adenine (A) interacts with thymine (T), and guanine (G) interacts with cytosine (C).

DNA (deoxyribonucleic acid): describes the structure of molecules that comprise genes, units of genetic information, that are passed from generation to generation.

eukaryote: organism whose cells possess membrane-bound nuclei and other organelles.

genetic selection: in yeast, involves growth under conditions which select for cells containing the characteristic of interest.

in vitro: the environment outside a cell

in vivo: the environment within a cell

nucleotide: a monomer of DNA or RNA.

mutagenesis: process or act of introducing a mutation or DNA base change.

plasmid: small circular DNA molecule that carries genes separate from the main bacterial DNA.

polymerase: enzyme that catalyzes the formation of a polymer, a molecule built from repeating subunits of the same general type, such as amino acids in proteins, nucleotides in nucleic acid, or sugar units in polysaccharides.

prototroph: a strain of organism that is able to grow on a defined minimal medium from which all of the more complex biological molecules required can be synthesized.

proteolysis: cleavage of a protein.

RNA (ribonucleic acid): nucleic acids that function as intermediates when information in DNA is directed to produce proteins.

transcription: the synthesis of RNA from a DNA template.

transcription initiation: the first step which begins the three-step process of transcription.

transcription factor: a DNA binding protein that is involved in the regulation of transcription.

translation: conversion of information provided by messenger RNA into a specific sequence of amino acids in a polypeptide chain.

BIBLIOGRAPHY

Baxley, L. E. and Hawley, D. K. (1999).The two repeats of TBP have distinct contributions to binding. In preparation.

Cavallini, B., Faus, I., Matthes, H., Chipoulet, J. M., Winsor, B., Egly, J. M., and Chambon, P. (1989). Cloning of the gene encoding the yeast protein BTF1Ywhich can substitute for the human TATA box-binding factor. Proc. Natl. Acad. Sci. USA 86, 9803-9807.

Chasman, D. I., Flaherty, K. M., Sharp, P. A., and Kornberg, R. D. (1993). Crystal Structure of yeast TATA-binding protein and model for interaction with DNA. Proc. Natl. Acad. Sci. USA 90, 8174-8178.

Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R. G., and Chua, N. H. (1990). Arabidopsis thaliana contains two genes for TFIID. Nature 346, 390-394.

Hahn, S., Buratowski, S., Shart, P. A., and Guarente, L. (1989). Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 supressor of Ty element insertion. Cell 58, 1173-1186.

Hermann, S. and Hawley, D. K. (1999). Evidence for a multi-step binding pathway in the formation of TATA binding protein-DNA complexes. In preparation.

Hernandez, N. (1993). TBP, a universal transcription factor? Genes Dev. 7, 1291-1308.

Hoey, T., Dynlacht, B. D., Peterson, M. G., Puch, B. F., and Tijan, R. (1990). Isolation and characterization of the Drosophila gene encoding the TATA box binding protein, TFIID. Cell 61, 1179-1186.

Hoopes, B. C., LeBlanc, J. F., and Hawley, D. K. (1992). Kinetic Analysis of Yeast TFIID-TATA Box Complex Formation Suggests a Multi-step Pathway. J. Biol. Chem. 267, 11539-11547.

Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A., and Roeder, R. G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. Nature 341, 299-303.

Horikoshi, M., Yamamoto, T., Ohkuma, Y., Weil, P. A., and Roeder, R. G. (1990). Analysis of structure-function relationships of yeast TATA box binding factor. Cell 61, 1171-1178.

Hoffmann, A., Horikoshi, M., Wang, C. K., Schroeder, S., Weil, P. A., and Roeder, R. G., (1990). Cloning of the *Schizosaccharomyces pombe* TFIID gene reveals a strong conservation of functional domains present in *Saccharomyces cerevisiae* TFIID. Genes Dev. 4, 1141-1148.

Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R. G. (1990). Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). Nature 346, 387-390.

Kao, C. C., Lieberman, P. M., Schmidt, M. C., Zou, Q., Pei, R., and Berk, A. J. (1990). Cloning of a transcriptionally active human TATA binding factor. Science 248, 1646-1650.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365, 520-527.

Kim, Y., Geiger, J. H., Hahn, S., Sigler, P. B. (1993). Crystal structure of a yeast TBP/TATA-box complex. Nature 365, 512-521.

Kuddus, R., and Schmidt, M. C. (1993). Effect of the non-conserved N-terminus on the DNA binding activity of the yeast TATA binding protein. Nucl. Acids. Res. 21, 1789-1796.

Lieberman, P. M., Schmidt, M. C., Kao, C. C., and Berk, A. J. (1991). Two distinct domains in the yeast transcription factor IID and evidence for a TATA box-induced conformational change. Mol. Cell. Biol. 11, 63-74.

Muhich, M. L., Iida, C. T., Horikoshi, M., Roeder, R. G., and Parker, C. S. (1990). cDNA clone encoding Drosophila transcription factor TFIID. Proc. Natl. Acad. Sci. USA 87, 9148-9152.

Nikolov, D. B., Hu, S. H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N. H., Roeder, R. G., and Burley, S. K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature 360, 40-46.

Nikolov, D. B., Chen, H., Halay, E. D., Hoffmann, A., Roeder, R. G., and Burley, S. K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. Proc. Natl. Acad. Sci. USA 93, 4862-4867.

Peterson, M. G., Tanese, N., Pugh, B. F., and Tijan, R. (1990). Functional Domains and Upstream Activation Properties of Cloned Human TATA Binding Protein. Science 248, 1625-1630.

Poon, D., Schroeder, S., Wang, C. K., Yamamoto, T., Horikoshi, M., Roeder, R. G., and Weil, P. A. (1991). The Conserved Carboxy-Terminal Domain of *Saccharomyces cerevisiae* TFIID is Sufficient to Support Normal Cell Growth. Mol. Cell. Biol. 11, 4809-4821.

Schmidt, M. C., Kao, C. C., Pei, R., and Berk, A. J. (1989). Yeast TATA-box transcription factor gene. Proc. Natl. Acad. Sci. USA 86, 7785-7789.

Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA 73, 804-808.

Singer, V. L., Wobbe, C. R., and Struhl, K. (1990). A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes Dev. 4, 636-645.

Starr, D. B., and Hawley, D. K. (1991). TFIID Binds in the Minor Groove of the TATA Box. Cell 67, 1231-1240.

Starr, D. B., Hoopes, B. C., Hawley, D. K. (1995). DNA Bending is an Important Component of Site-specific Recognition by the TATA Binding Protein. J. Mol. Biol. 250, 434-446.

Strubin, M., and Struhl, K. (1992). Yeast and Human TFIID with Altered DNA-Binding Specificity for TATA Elements. Cell 68, 721-730.

Zawel, L., and Reinberg, D. (1995). Common themes in assembly and function of eukaryotic transcription complexes. Ann. Rev. Biochem. 64, 2313-2323.