Title
Analysis of Drosophila melanogaster snRNA activating protein complex binding to the U1 gene promoter

Permalink
https://escholarship.org/uc/item/7r9982g8

Author
Lai, Hsien-Tsung

Publication Date
2007

Peer reviewed|Thesis/dissertation
Analysis of *Drosophila melanogaster* snRNA activating protein complex binding to the U1 gene promoter

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biology by Hsien-Tsung Lai

Committee in Charge:

University of California, San Diego

Professor Tracy Johnson
Professor James W. Posakony

San Diego State University

Professor William E. Stumph, Chair
Professor Sanford I. Bernstein
Professor Terrence G. Frey

2007
The dissertation of Hsien-Tsung Lai is approved, and it is acceptable in quality and form for publication on microfilm:

_____________________________________________________________________
_____________________________________________________________________
_____________________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

Chair

University of California, San Diego
San Diego State University

2007
Dedicated to my beloved wife and rest of my family

For their love, support and guidance
# TABLE OF CONTENTS

Signature Page ........................................................................................................ iii
Dedication Page ....................................................................................................... iv
Table of Contents ...................................................................................................... v
List of Figures .......................................................................................................... vii
Acknowledgement .................................................................................................. ix
Vita .......................................................................................................................... x
Publications and Abstracts ..................................................................................... xi
Fields of Study ....................................................................................................... xii
Abstract .................................................................................................................. xiii

General Introduction ............................................................................................... 1
  A: References ........................................................................................................ 13

Chapter 1

The PSEA promoter element of the *Drosophila* U1 snRNA gene is sufficient to bring DmSNAPc into contact with 20 base pairs of downstream DNA .............................................................................. 19
  A: Abstract ......................................................................................................... 20
  B: Introduction .................................................................................................. 20
  C: Material and Methods .................................................................................. 21
  D: Results ........................................................................................................... 23
  E: Discussion .................................................................................................... 25
  F: References .................................................................................................... 26

Chapter 2

Stoichiometry of the subunits of the *Drosophila melanogaster* small nuclear RNA activating protein complex .............................................................................. 28
  A: Abstract ......................................................................................................... 29
  B: Introduction .................................................................................................. 30
  C: Material and Methods .................................................................................. 32
Chapter 3

A new method for localizing sites within a protein that cross-link to specific nucleotide positions within a DNA recognition sequence: Application to determining the orientation of the large subunit of DmSNAPc bound to a U1 snRNA gene promoter

A: Abstract
B: Introduction
C: Material and Methods
D: Results
E: Discussion
F: References

Concluding Remarks

A: References

Appendix

A: DNA constructs for expressing various forms of the DmSNAP subunits in S2 cells
B: Stably transfected S2 cell lines
C: Purification of DmSNAPs from S2 cells by nickel chelate chromatography
D: Detailed protocol for band-shift and super-shift assay
E: Detailed protocol for western blot
F: Site-specific protein-DNA photo-cross-linking of DmSNAP/PSEA using azidophenacyl bromide as cross-linking agent
G: Hydroxylamine digestion after photo-cross-linking reactions
H: Detailed protocol for dumbbell probe preparation
LIST OF FIGURES

General Introduction

Figure G.1: Schematic representation of cis-acting elements in the 5’-flanking DNA of a variety of snRNA genes................................. 5
Figure G.2: Conserved structure of Drosophila snRNA gene promoters transcribed by RNA polymerase II and RNA polymerase III........... 6
Figure G.3: (A) Summary of the photo-cross-linking results (B) Schematic model for the differential interaction of DmSNAPc with U1 and U6 PSEAs................................................................................. 9
Figure G.4: Illustration of alternative DmSNAP43 binding to wild type or mutant U1 promoter ................................................................. 12

Chapter 1

Figure 1.1: (A) Organization of promoter sequences of spliceosomal snRNA genes. (B) Interaction of the subunits of DmSNAPc with U1 and U6 promoters. (C) A listing of the sequences of the PSEBs of the known U1-U4 and U5 genes of D. melanogaster................................................................................. 21
Figure 1.2: Site-specific protein–DNA photo-cross-linking of DmSNAPc subunits to phosphate positions downstream of the U1 PSEA in probes that contain either the wild-type or a mutant PSEB........ 22
Figure 1.3: PSEB mutations reduce luciferase reporter gene expression ~2-fold.................................................................................................. 24
Figure 1.4: Primer extension assays indicate that mutations in the PSEB reduce U1 promoter activity 2- to 4-fold........................................ 24
Figure 1.5: The PSEB cannot effectively substitute for the TATA-box in the U6 promoter ................................................................. 25
Figure 1.6: A speculative model of TBP and DmSNAP43 interacting simultaneously with the PSEB............................................................... 26

Chapter 2

Figure 2.1: Stoichiometry of DmSNAP190 in DmSNAPc................................. 35
Figure 2.2: Stoichiometry of DmSNAP50 in DmSNAPc................................. 39
Figure 2.3: Stoichiometry of DmSNAP43 in DmSNAPc................................. 41
Chapter 3

Figure 3.1: Schematic structure of DmSNAP190 .............................................. 50

Figure 3.2: Photo-cross-linking of DmSNAPc to positions 1-25 of a U1 gene PSEA .................................................................................... 53

Figure 3.3: DmSNAP190 fragments predicted from hydroxylamine digestion ........................................................................................ 56

Figure 3.4: Phosphate position 24 of the U1 PSEA cross-links to the N-terminus of DmSNAP190 .............................................................. 58

Figure 3.5: Phosphate position 12 of the U1 PSEA cross-links to the N-terminus of DmSNAP190 .............................................................. 60

Figure 3.6: Phosphate position 1 of the U1 PSEA most likely cross-links to the C-terminus of DmSNAP190 .............................................................. 62

Concluding Remarks

Figure C.1: (A) DmSNAP190 amino acid sequence. Red font indicates the evolutionarily conserved Myb repeats. (B) An alignment of the Myb repeats (Rh, Ra, Rb, Rc and Rd) ......................................................... 70

Figure C.2: Diagram showing the positions of the naturally occurring hydroxylamine cutting sites (N-G) in all three DmSNAPc subunits .......................................................................................... 72
ACKNOWLEDGEMENTS

I would like to thank Dr. Stumph for giving me the opportunity to work in his lab and for his guidance, sincere help and constant encouragement throughout the projects. I thank Dr. Bernstein, Dr. Frey, Dr. Johnson and Dr. Posakony for serving on my dissertation committee. I am grateful to Kathleen McNamara-Schroeder for excellent technical assistance and Michelle Chen, Mitch Titus, Ko-Hsuan Alex Hung, Shu-Chi Kiki Chiang, Nermeen Barakat, Anna Luy, Genaro Hernandez, Gale A. Harding, and all the other members of the Stumph lab for making an exciting work environment. I would like to thank Cheng Li who gave me many helpful hints and suggestions as well as passing on a detailed photo-cross-linking laboratory protocol.

The text of Chapter one, in full, is a reprint of the material as it appears in: Lai, H. T., H. Chen, C. Li, K. J. McNamara-Schroeder, and W. E. Stumph. 2005. The PSEA promoter element of the Drosophila U1 snRNA gene is sufficient to bring DmSNAPc into contact with 20 base pairs of downstream DNA. Nucleic Acids Res 33:6579-86. I was the primary researcher and author listed in this publication and Dr. William E. Stumph directed and supervised the research which forms the basis for this chapter.
VITA

1990-1994  B.S., Biology
Fu Jen Catholic University
Hsin-Chuang, Taiwan

1994-1996  M.S., Institute of Radiation Biology
National Tsing-Hua University
Hsin-Chu, Taiwan

1996-1998  Military service in Army
Taiwan

1998-2001  Research Associate
Institute of Zoology, Academia Sinica
Taipei, Taiwan

2001-2004  Teaching Associate
Department of Biology, San Diego State University
San Diego, California

2004-2007  Pre-doctoral Candidate
Department of Biology (Cell and Molecular Biology)
University of California, San Diego and
San Diego State University
San Diego, California

2007  Ph.D. in Biology (Cell and Molecular Biology)
University of California, San Diego and
San Diego State University
San Diego, California
PUBLICATIONS AND ABSTRACTS


FIELDS OF STUDY

Major Field: Cell and Molecular Biology

Studies in Transcriptional Regulation of Eukaryotic Gene Expression
Professor William E. Stumph
ABSTRACT OF THE DISSERTATION

Analysis of Drosophila melanogaster snRNA activating protein complex binding to the U1 gene promoter

by

Hsien-Tsung Lai

Doctor of Philosophy in Biology
University of California, San Diego, 2007
San Diego State University, 2007

Professor William E. Stumph, Chair

In animals, the U1, U2, U4 and U5 small nuclear RNA (snRNA) genes are transcribed by RNA polymerase (RNAP) II, but U6 snRNA genes are transcribed by RNAP III. Transcription of both classes of genes is dependent upon a 21 base pair (bp) sequence termed the PSEA located ~40-60bp upstream of the transcription start site. Other promoter elements consist of a TATA box (in U6) and a PSEB (in U1-U5).

The PSEAs of both classes of Drosophila snRNA genes are recognized by the same transcription factor, DmSNAPc (Drosophila melanogaster snRNA activating protein complex), which comprises three distinct subunits (DmSNAP43, DmSNAP50 and DmSNAP190). A striking previous finding was that the DmSNAP43 subunit cross-links to DNA more than 20 bp downstream of the U1 PSEA (a region that
includes the PSEB). These findings raise the question of whether the PSEB contributes to the cross-linking pattern downstream of the U1 PSEA. To investigate this, the photo-cross-linking patterns from wild type or mutant PSEB probes were compared. Both sets of probes produced a similar, although not identical, photo-cross-linking pattern. These results indicate that the PSEA itself can bring DmSNAP43 into close proximity to the downstream DNA regardless of the PSEB sequence.

A second part of this study focused on the stoichiometry of the subunits of DmSNAPc bound to DNA. To investigate this, identical subunits were tagged with different epitopes and co-expressed in Drosophila S2 cells with each other and the other two subunits. Following purification of the tagged DmSNAPc the presence of differently tagged subunits in DmSNAPc bound to DNA was investigated by band-shift and super-shift assays. The results indicate that each of the subunits is present in only a single copy in DmSNAPc bound to DNA.

A third part of this study focused on the N- and C-terminal orientation of the largest subunit, DmSNAP190, when bound to the U1 promoter. By combining the photo-cross-linking assay with chemical digestion of the protein, I have been able to demonstrate that the N-terminal half of DmSNAP190 contacts the 3’ end of PSEA and most likely the C-terminal half contacts the 5’ end of the PSEA.
GENERAL INTRODUCTION
Characteristic features of U-snRNAs

The small nuclear RNAs (snRNAs) are a metabolically stable class of RNA molecules in eukaryotic nuclei. “U” snRNAs were named according to their being rich in uridylic acid and different from messenger RNA (mRNA) or ribosomal RNA (rRNA) or transfer RNA (tRNA). As more snRNAs were discovered, though, some of these were not found to contain a high proportion of uridylic acid. The important function of the U-snRNAs is reflected by the fact that the predicted secondary structures among homologous U-snRNAs from distant species are virtually identical (1).

Function of the spliceosomal snRNAs

There are five snRNAs U1, U2, U4, U5 and U6 involved in splicing the precursors of mRNAs (2-4). In vivo, one U-snRNA is usually associated with 6-10 polypeptides and comprises a small nuclear ribonucleoprotein particle (snRNP) (1,5-7). The U4, U5, and U6 snRNPs interact to form a single particle called the U4-U5-U6 tri-snRNP complex (8,9).

Role of the U1-U6 snRNAs in spliceosome assembly and catalysis

Spliceosome formation begins when U1 snRNP recognizes and binds specifically to the 5' splice site of a mRNA precursor, via the 5' terminal sequence of U1 snRNA which is complementary to the 5' splice site of intron-exon junctions. This is followed by the specific binding of the U2 snRNP to the branch site, which also involves specific base pairing that "bulges out" the branch point adenosine (10,11). After U1 and U2 snRNPs bind to the pre-mRNA, the U4-U5-U6 tri-snRNP joins the complex. The U5 snRNA initially binds to a conserved exon sequence adjacent to the
5' splice site (12). Following an ATP-dependent conformational change, the active site is formed that allows catalysis.

The first step in catalysis involves the nucleophilic substitution at the 5' splice junction by the 2'-hydroxyl group of the branch point adenosine. The resulting products are the lariat intermediate and the free 5' exon. During this process, the U5 snRNA shifts and makes a stronger link to the 5' exon. At the same time, pairing between U1 snRNA and the intron sequence weakens. Then U6 snRNA takes the place of U1 in binding to these sequences (9). The U6 snRNA base pairs to the 5' splice site (13,14). The U1 and U4 snRNPs then leave the spliceosome.

The second step of catalysis involves the nucleophilic substitution at the 3' splice junction by the 3' OH group of the newly freed 5' exon. Here, U5 retains its contact with the free 5' exon and also establishes a new contact with a 3' exon sequence immediately downstream from the 3' splice junction (15). The precise mechanisms of the catalysis remain unknown and are the subject of extensive investigation.

**Structure and expression of genes coding for the snRNAs**

In *Drosophila melanogaster*, there are several copies of each of the snRNA genes (16). More specifically, there are 5 copies of true U1 snRNA genes, 5 copies of U2 snRNA genes, 4 copies of U4 snRNA genes, 7 copies of U5 snRNA genes and 3 copies of U6 snRNA genes in *D. melanogaster* (17). In humans, there are about 30 copies per haploid genome of true U1 snRNA genes, and ~10 copies of the U2 genes (18-21). There are 5 copies of U5 genes (22). There are 9 full-length U6 loci in the human genome, five of them being true genes (23).
The snRNA genes represent a unique class of transcription units

Interestingly, the U1-U5 snRNAs are synthesized by RNAP II, but U6 is synthesized by RNAP III (24-26). Furthermore snRNA genes in higher plants and animals have promoter structures and cis-acting elements that functionally distinguish them from classical RNAP II or RNAP III transcription units.

Fig. G.1 summarizes work on a variety of higher eukaryotes and shows regulatory elements identified in the basal promoter regions of snRNA genes transcribed by either RNAP II or RNAP III. A promoter element (designated PSE, PSEA or USE in Fig. G.1) is located approximately 40 to 75 bp upstream of the start site and is essential for the initiation of snRNA gene transcription (24,27-30).

Besides the PSE, many snRNA genes contain a second promoter element located approximately 25-30 bp upstream of the transcription start site. An exception is that vertebrate RNAP II-transcribed snRNA genes lack any well-conserved element in this region (Fig. G.1). In contrast, vertebrate U6 genes, transcribed by RNAP III, contain a conserved TATA box. Mutation of the U6 TATA box to an unrelated sequence changed the promoter specificity from RNAP III to RNAP II (31,32). Conversely, the introduction of a TATA sequence into the U1 or U2 promoters altered their specificity to RNAP III (31,32). These findings led to the idea that the TATA box acts as a dominant element in determining the RNAP III specificity of vertebrate U6 promoters (33).
Figure G.1. Schematic representation of cis-acting elements in the 5’-flanking DNA of a variety of snRNA genes. All the snRNA genes shown contain an essential element upstream of position -40. Most snRNA genes also contain a regulatory element near position -30 or -25, which probably represents a site of interaction with TBP or a TBP-related factor.

In plants, both classes of snRNA genes contain a TATA box. In this case, polymerase specificity is determined by a 10 bp difference in spacing between the TATA box and the USE (32-36 bp spacing for RNAP II versus 23-26 bp for RNAP III, Fig. G.1) (34,35). In vertebrates and plants (as well as sea urchins), the PSE (or USE) was found to be functionally interchangeable among the U1, U2, and U6 genes (32,34-37). The conclusion from those studies was that the PSE or USE itself does not contribute directly to RNA polymerase specificity in those organisms. There are little or no mechanistic data about how RNA polymerase specificity is determined by the presence vs. absence of a TATA box in vertebrate snRNA gene promoters, or by the difference in spacing between the USE and TATA box in plants.
In the fruit fly and other insects, the promoter elements of the various snRNA genes are more conserved with regard to both sequence and location than generally observed in other organisms (Fig. G.2) (17). All known *D. melanogaster* snRNA genes contain a 21 bp PSEA that is well conserved in sequence (38-40). All *D. melanogaster* snRNA genes transcribed by RNAP II also contain a well-conserved 8 bp PSEB that is necessary for efficient initiation of transcription (28). A separation of 8 bp is strictly conserved between the PSEA and PSEB in RNAP II-transcribed snRNA genes and 12 bp between the PSEA and TATA box of U6 gene promoters (17,27,38,39).

Figure G.2. Conserved structure of *Drosophila* snRNA gene promoters transcribed by RNA polymerase II and RNA polymerase III. The wild type U1 and U6 PSEA sequences used in our studies differ at only 5 of 21 positions (bold and underlined). The PSEB and TATA box differ at 5 of 8 positions.

**The PSEA is a dominant element for determining the RNA polymerase specificity of *D. melanogaster* snRNA gene promoters**

The results of *in vitro* transcription experiments using mix-and-match templates that contained all possible combinations of U1 or U6 PSEA, 8 or 12 bp
spacing, and PSEB or TATA box has been previously reported by our lab (39). The U1 and U6 PSEAs differed at only the five nucleotide positions shown in Fig. G.2. Constructs that contained the U1 PSEA were transcribed by RNAP II, and those that contained the U6 PSEA were transcribed by RNAP III. The PSEB and TATA elements, as well as the 8 vs. 12 bp spacing, affected transcription efficiency but did not directly affect the choice of RNA polymerase in vitro (39).

Analogous in vivo experiments were carried out with reporter constructs that contained U1 and U6 promoters with “swapped” PSEAs (41). Substitution of the U6 PSEA into the U1 promoter, or substitution of the U1 PSEA into the U6 promoter suppressed transcription in vivo. The results clearly indicated that the U1 PSEA cannot function for RNAP III transcription and the U6 PSEA cannot function for RNAP II transcription, even through they differ at only 5 of 21 nucleotide positions.

**Characterization of the *Drosophila* PSEA-binding Protein**

The PSE-binding protein (PBP) was first identified in the human system in HeLa cell extracts (42). It was further characterized and variously termed proximal transcription factor (PTF) (43,44) or snRNA activating protein complex (SNAPc) (45,46), which has become the most popular name for the factor. SNAPc/PTF was capable of activating both RNAP II and RNAP III transcription from snRNA promoters (42,44,47-49). The human protein contains integral polypeptide subunits with apparent molecular weights of approximately 19, 43, 45, 50, and 190 kDa. The genes for each have been cloned (44,46-48,50-52). These proteins, and the genes that encode them, are termed SNAP19, SNAP43, SNAP45, SNAP50, and SNAP190.
The *Drosophila melanogaster* PSEA-binding protein (DmPBP, more recently re-termed DmSNAPc) has been characterized in our lab (26,53-55). DmSNAPc (DmPBP) binds to the U1 and U6 PSEAs and can activate transcription of the *Drosophila* U1 and U6 snRNA genes *in vitro* (26). Our lab also showed that DmSNAPc contains three distinct polypeptides: DmPBP45, DmPBP49, and DmPBP95, designations based upon their apparent molecular weights on SDS PAGE (55). These three subunits were originally identified by site-specific protein-DNA photo-cross-linking (55). The cloning and characterization of the genes for the three subunits of DmSNAPc has also been carried out in our lab (54). These three genes encode proteins with similarity to the SNAP43, SNAP50, and SNAP190 subunits of human SNAPc (54). Therefore, we named these the DmSNAP43, DmSNAP50 and DmSNAP190 genes and now refer to the encoded proteins by the same names (i.e., DmSNAP43=DmPBP45, DmSNAP50= DmPBP49 and DmSNAP190= DmPBP95).

The photo-cross-linking assay also gave us important information about how the protein complex is situated on the DNA (Fig. G.3), when the DNA is oriented as shown in Fig. G.3A. DmSNAP43 interacts primarily with the “upper” face of the DNA helix. DmSNAP50 approaches the “front” surface of the DNA. DmSNAP190 interacts with the front face of the helix in the 5’ half of the PSEA, but interacts with the lower surface of the DNA toward the 3’ end of the PSEA.
Figure G.3. (A) Summary of the photo-cross-linking results from Wang and Stumph (1998) and Li et al. (2004). Colored spheres and colored backbone (blue, green and yellow for DmSNAP43, DmSNAP50 and DmSNAP190 respectively) indicate positions of significant cross-linking of DmSNAP subunits to the DNA, where the positions are relative to the rectangles shown at the top of the figure. Red spheres indicate the positions of very strongest cross-linking. (B) Schematic model for the differential interaction of DmSNAPc with U1 and U6 PSEAs. The U1 and U6 PSEAs are bent similarly toward the face of the helix that contacts DmSNAP43. The different DNA sequences are believed to act as differential allosteric effectors of the conformation of DmSNAPc.
Finally, it is important to note that the cross-linking pattern of the DmSNAPs to the U1 PSEA and U6 PSEA was different (Fig. G.3A). The smallest subunit, DmSNAP43 displayed profoundly different cross-linking patterns when DmSNAPc bound to U1 vs. U6 PSEAs. On a U1 PSEA, DmSNAP43 cross-links to phosphate positions between 16 and 40, but on a U6 PSEA it cross-links to phosphate positions between 11 and 25. DmSNAP50 also exhibited significant differences in its cross-linking pattern to U1 and U6 PSEAs (Fig. G.3A).

Overall, the results of the photo-cross-linking assay indicate that the conformation of the protein-DNA complex is different when DmSNAPc binds to a U1 PSEA vs. a U6 PSEA. Previous results from our lab indicated that the U1 and U6 PSEAs are both bent by a similar degree toward the face of the DNA contacted by the DmSNAP43 subunits (56). We therefore believe that the conformational differences observed in the U1 and U6 DNA-DmSNAP complexes exist primarily at the level of the protein. A schematic arrangement of the DmSNAPc-PSEA complex is shown in Fig. G.3B.

**Subject matter of this dissertation**

The wild type *D. melanogaster* U1 and U6 promoters each contain a (functionally distinct) PSEA, and each contains a dissimilar element downstream of the PSEA, either a PSEB (U1) or a TATA box (U6). It is interesting that DmSNAP43 cross-links to nucleotides up to 20 bp downstream of the PSEA when DmSNAPc is bound to a U1 PSEA (Fig. G.3). Interestingly, these strong cross-links occur within and near to the PSEB. However, substitution of the U6 PSEA for the U1 PSEA abrogated these downstream cross-links (54). Those findings raise the distinct
possibility that the PSEB cooperates with the U1 PSEA (but not with the U6 PSEA) to permit strong cross-linking of DmSNAP43 to the PSEB and nearby nucleotides.

In work described in Chapter 1, I carried out photo-cross-linking with probes that contain a mutant PSEB downstream of the U1 PSEA. The PSEB mutation was one previously found to reduce U1 in vitro transcription more than 8-fold (28). There were two alternative results possible. First, the mutant probe could have a pattern of cross-linking identical to that seen with the wild type probe (compare Fig. G.4A and 4B). This would indicate that the U1 PSEA by itself is capable of bringing DmSNAP43 into close proximity to the downstream DNA, regardless of the PSEB sequence. (The U6 PSEA does not possess this ability). The alternative result is that the downstream cross-linking could be absent or different as a result of mutating the PSEB (compare Fig. G.4C and 4A). This result would indicate that the sequence of the PSEB is important for DmSNAP43 cross-linking. Furthermore, it would suggest that DmSNAP43 has some sequence specificity for DNA binding, and this would possibly explain why the PSEB sequence is conserved in fly snRNA gene promoters transcribed by RNAP II.
My work demonstrates that mutation of the PSEB does not abolish the cross-linking of DmSNAP43 to the PSEB. Thus the U1 PSEA alone is capable of bringing DmSNAP43 into close contact with this downstream DNA. However, mutation of the PSEB perturbed the cross-linking pattern. In concordance with these findings, PSEB mutations resulted in a 2 to 4-fold reduction in U1 promoter activity when assayed by transient transfection.

Although the preceding illustrations are drawn to suggest that there is only a single copy of each subunit in DmSNAP bound to DNA, there has been no information that directly addresses the stoichiometry of the subunits in DmSNAPc. For example, it is possible that one or more of the subunits could be present in two or more copies. To address this question, in Chapter 2 I co-expressed differently tagged
versions of the same subunit within the same cells: following purification of the differently tagged complexes, I employed electrophoretic mobility shift and super-shift assays. From these results I was able to deduce that each of the subunits is present in only one copy when DmSNAPc is bound to DNA.

In Chapter 3, I developed a new methodology in which I combined site-specific protein-DNA photo-cross-linking with chemical digestion of N- and C-terminally tagged DmSNAP190. Results reveal that the N-terminal half of DmSNAP190 contacts the 3’ end of the PSE and suggest that the C-terminal half of DmSNAP190 contacts the 5’ end of the PSE. This orients the N-terminal half of DmSNAP190 toward the transcription start site and the C-terminal half toward the upstream DNA.

REFERENCES


CHAPTER 1

The PSEA promoter element of the Drosophila U1 snRNA gene is sufficient to bring DmSNAPc into contact with 20 base pairs of downstream DNA
The PSEA promoter element of the *Drosophila* U1 snRNA gene is sufficient to bring DmSNAPc into contact with 20 base pairs of downstream DNA

Hsien-Tsung Lai, Hsiang Chen, Cheng Li\(^1\), Kathleen J. McNamara-Schroeder\(^1\) and William E. Stumph\(^1,*\)

Department of Biology and \(^1\)Department of Chemistry and Biochemistry, San Diego State University, San Diego, CA 92182-1030, USA

Received September 11, 2005; Revised and Accepted November 3, 2005

**ABSTRACT**

Most of the major splicedosomal small nuclear RNAs (snRNAs) (i.e. U1, U2, U4 and U5) are synthesized by RNA polymerase II (Pol II). In *Drosophila melanogaster*, the 5’-flanking DNA of these genes contains two conserved elements: the proximal sequence element A (PSEA) and the proximal sequence element B (PSEB). The PSEA is essential for transcription and is recognized by DmSNAPc, a multi-subunit protein complex. Previous site-specific protein–DNA photo-cross-linking assays demonstrated that one of the subunits of DmSNAPc, DmSNAP43, remains in close contact with the DNA for 20 bp beyond the 3’ end of the PSEA, a region that contains the PSEB. The current work demonstrates that mutation of the PSEB does not abolish the cross-linking of DmSNAP43 to the PSEB. Thus the U1 PSEA alone is capable of bringing DmSNAP43 into close contact with this downstream DNA. However, mutation of the PSEB perturbs the cross-linking pattern. In concordance with these findings, PSEB mutations result in a 2- to 4-fold reduction in U1 promoter activity when assayed by transient transfection.

**INTRODUCTION**

The genes that code for the splicedosomal small nuclear RNAs (snRNAs) are unusual transcription units in that most of the snRNA genes (U1, U2, U4 and U5 genes) are transcribed by RNA polymerase II (Pol II), but U6 genes are transcribed by RNA polymerase III (Pol III). However, the promoters for both classes of genes are similar and both utilize a unique multi-subunit transcription factor that has variously been termed the snRNA activating protein complex (SNAPc) (1,2), PTF (3,4) or PB P (5,6). In vertebrates, this protein complex recognizes a proximal sequence element, or PSE, centered ~50–55 bp upstream of the transcription start site of both classes of snRNA genes (7–10). U6 gene promoters additionally contain a TATA-box that is required for the Pol III specificity of the vertebrate U6 genes (11,12). The occurrence of a TATA-box 25–30 bp upstream of the transcription start site of U6 genes appears to be a conserved feature that is maintained throughout all branches of metazoan evolution. The vertebrate snRNA genes that are transcribed by Pol II, on the other hand, lack TATA-boxes or any other recognizable conserved element in the ~25 to ~30 region. However, TATA-binding protein (TBP) is still required for the transcription of these genes (1,13).

Our lab is studying the expression of the snRNA genes of the fruit fly *Drosophila melanogaster*. Interestingly, the *D.melanogaster* snRNA genes transcribed by Pol II contain two conserved elements in their 5’-flanking DNAs that are termed proximal sequence element A (PSEA) and proximal sequence element B (PSEB) separated by 8 bp of non-conserved sequence (Figure 1A). The three *D.melanogaster* U6 genes on the other hand each contain a PSEA separated by 12 bp from a downstream TATA-box.

The upstream PSEAs of both classes of fruit fly genes are recognized by DmSNAPc, a multi-subunit protein complex formerly called DmPBP (14–16). It contains three distinct subunits that are homologous to subunits of human SNAPc: DmSNAP43, DmSNAP50 and DmSNAP190. Site-specific protein–DNA photo-cross-linking studies have established the architectural arrangement of these three subunits on the U1 and U6 PSEAs (15,16). This is shown schematically for the
U1 and U6 gene promoters in Figure 1B. On both promoters DmSNAP43 contacts the PSEA, but on the U1 promoter the DmSNAP43 subunit further contacts nucleotides for 20 bp beyond the PSEA. This finding raised the possibility that the PSEA might provide some sequence-specificity for the binding of DmSNAP43 to DNA sequences downstream of the U1 PSEA. The PSEA sequences of the D.melanogaster U1, U2, U4 and U5 genes are compared in Figure 1C. A consensus sequence is shown at the bottom of the column.

The DNA contacts that DmSNAP43 makes near and within the PSEA seem to be important because substitution of a U6 PSEA for the U1 PSEA in the U1 promoter both (i) abolished the contacts of DmSNAP43 with the PSEA and (ii) suppressed transcription from the U1 promoter by Pol II (16,17). Photo-cross-linking and partial proteolysis assays indicate that DmSNAP43 binds in altered conformations to the U1 and U6 PSEAs (15,16,18). In our working model (Figure 1B), the different sequences of the U1 and U6 PSEAs act as differential allosteric effectors of DmSNAP43, which then adopts conformations compatible with the formation of only Pol II or Pol III transcription initiation complexes.

Considerable work has been done on the function of the U1 and U6 PSEAs, but the function of the PSEB in transcription of fruit fly snRNA genes by Pol II is not understood. In an earlier study, we found that site-specific mutation of the PSEB resulted in an 8-fold decrease in promoter activity when assayed by transcription in vitro (19). However, its effect on transcription in living cells had not been investigated. Moreover, it seemed important to determine whether the PSEB sequence itself was required for DmSNAP43 to make contact to the DNA downstream of the U1 PSEA (Figure 1B). Here we report the results of such experiments.

**MATERIALS AND METHODS**

**Site-specific protein–DNA photo-cross-linking assays**

Forty-four different probes, each containing cross-linking reagent at a unique position, were used to scan for DmSNAP43 interactions with the DNA downstream of the U1 PSEA on the template and non-template strands. Two series of probes were prepared: one having a U1 PSEA with a wild-type PSEB (CATGAAA) downstream, and an analogous series with an 8 bp mutation in the PSEB (AGGCTCTT) (19). Cross-linker was incorporated at every second phosphate position as indicated by the asterisks in Figure 2. Probes were prepared as described previously (15,16), based upon earlier methods as described by Yang and Nash (20) and Lague et al. (21). Briefly, DNA oligonucleotides 23 to 38 bases long were synthesized with phosphorothioate incorporated 5' of the third nucleotide from the 5' end. The phosphorothioate-incorporated oligonucleotides were derivatized with azido-phenacyl bromide and then radiolabeled at the 5' end by using [γ-32P]ATP and T4 polynucleotide kinase.

To make the photo-cross-linking probes double-stranded, a derivatized oligonucleotide and an unmodified upstream
Figure 2. Site-specific protein-DNA photo-cross-linking of DmSNAPc subunits to phosphate positions downstream of the U1 PSEA in probes that contain either the wild-type or a mutant PSEA. (A) Sequences of the relevant areas of the U1 photo-cross-linking probes. The sequences of the 21 bp U1 PSEA of the PSEB and of the mutant PSEB are shown in boldface. Except in the region of the PSEB, the probes were identical to ensure that any differences in the cross-linking pattern were due solely to the substitutions in the PSEB. The individual phosphate positions at which a cross-linker is incorporated are indicated by the asterisks and numbers above and below the sequences. (B) Forty-four different radiolabeled site-specific probes were incubated in separate reaction mixtures with DmSNAPc. Following UV irradiation and nuclease digestion, polypeptides that cross-linked to the DNA were detected by SDS-PAGE and autoradiography. Only the regions of the gels that correspond to DmSNAPc3 and DmSNAPc30 are shown. The left panels show the results of cross-linking to the template strand, and the right panels show cross-linking to the non-template strand. In each case, the upper panel illustrates the pattern for the wild-type PSEB and the lower panel illustrates the pattern for the mutant PSEB. When cross-linked to positions 37-41 on the non-template strand, a portion of DmSNAPc3 runs with an anomalous mobility, creating a doublet. (Evidence that the lower band is due to DmSNAPc3 comes from photo-cross-linking experiments performed with DmSNAPc3 that contains a higher molecular weight tagged DmSNAPc3. The tagged DmSNAPc3 shifts the lower band (as well as the upper band) upward in the gel [data not shown]. The anomalous mobility is possibly due to the position of nucleotide linkage to the polypeptide chain.)

Wild-type and mutant reporter constructs

Constructs that contained the promoter of the U1.95Ca gene (formerly referred to as the U1.95.1 gene) fused to the firefly luciferase gene have been described previously (17). Each of these constructs contained 381 bp of 5'-flanking DNA plus 32 bp of the U1 coding region. Constructs that contained the U6-2 maxi-gene together with 409 bp of 5'-flanking DNA have also been described (17). The U6 maxi-gene contained a 10 bp insertion between 66 and 67 nt that allowed for the annealing of a complementary maxi-gene specific oligonucleotide for primer extension analysis without interference from endogenous U6 snRNA. Mutations were introduced either by cloning synthetic oligonucleotides between pre-existing restriction sites in either template or by using Stratagene’s Quick-Change Site-Directed Mutagenesis kit. The wild-type U1 PSEA was 5'-TAAATCCCACTGTTTTAGC-3'; the U6 PSEA substitution was 5'-TAAATCCCACTGTTTTAGC-3'; the mutant PSEA sequence was 5'-CCCTAGGTGATCGAC-3'; the wild-type PSEB was 5'-CATGGAAA-3'; the mutant PSEB sequence was 5'-AGGTTCTC-3'; and the TATA substitution that replaced the PSEB was 5'-TTTATAA-3'.
The U6 PSEA and TATA sequences used for these substitutions were derived from the *D. melanogaster* U6-2 gene, one of three U6 genes in the *D. melanogaster* genome (22,23).

**Transient expression assays**

*D. melanogaster* S2 tissue culture cells were grown and transfected as described previously (17). All transfections were performed in triplicate, and each trial was carried out as a distinct experiment on a different day. In experiments that measured luciferase activity, the U1 (firefly) luciferase constructs were co-transfected with a Renilla luciferase expression plasmid as an internal control for transfection efficiency. Assays were carried out by using Promega’s Dual-Luciferase Reporter Assay System and a Turner TD-20/20 luminometer.

For primer extensions to measure wild-type and mutant U1 and U6 promoter activities, total RNA was isolated from transfected cells as described previously (17). RNA aliquots (25 μL each) were co-precipitated with a 32P-labeled 5′ oligonucleotide recovery standard before the extension reaction. Extensions were carried out with Promega’s Primer Extension System with Invertigo’s SuperScript II reverse transcriptase substituted for Promega’s AMV reverse transcriptase. The oligonucleotide primers used for the extension assays have been described previously (17). Briefly, the U1-luciferase primer was complementary to a region near the 3′ end of the luciferase gene and yielded an 84 nt extension product; the U6 primer was partially complementary to the U6 maxi-gene insert and yielded an 89 nt product. Following gel electrophoresis, the intensities of the extension products were quantified by phosphorimager analysis.

**RESULTS**

**Mutation of the PSEB does not abolish the cross-linking of DmSNAP43 to DNA downstream of the U1 PSEA**

When DmSNAP43 binds to DNA that contains a U1 PSEA and a PSEB, the DmSNAP43 subunit not only contacts the 3′ end of the PSEA but also contacts the DNA near and within the PSEB (Figure 1B) (16). This conclusion was based upon results of high resolution site-specific protein–DNA photo-cross-linking experiments (15,16). The fact that DmSNAP43 contacts the PSEB in the U1 promoter raised the question of whether the sequence of the PSEB is necessary for these contacts to occur. Alternatively, it is possible that the U1 PSEA (but not the U6 PSEA) provides all the necessary information to place DmSNAP43 in contact with the DNA in the vicinity of the PSEB.

To examine this question, site-specific protein–DNA photo-cross-linking assays were carried out with DmSNAP43 and radiolabeled DNA probes that contained either a wild-type or mutant PSEB downstream of the wild-type U1 PSEA (Figure 2A). Forty-four individual probes were prepared. Each probe contained cross-linker at an unique phosphate position either on the lower (template) strand or on the upper (non-template) strand at the individual positions indicated by asterisks in Figure 2A. A 32P-radiolabel was incorporated at the adjacent phosphate in each case. Each probe was incubated individually with DmSNAP43 from *D. melanogaster* embryos. The reaction mixtures were then irradiated with UV light to activate the cross-linking agent, subjected to extensive nuclease digestion and run on SDS–PAGE. Protein bands that contained cross-linked radiolabel were detected by autoradiography.

Figure 2B (left side) shows the results of cross-linking DmSNAP43 to the template strand at positions 20 through 42 with probes that contain the wild-type PSEB (upper panel) or mutant PSEB (lower panel). The subunit known as DmSNAP50 cross-linked to both types of probes at positions 20 and 22, but this subunit did not cross-link to the DNA further beyond that point. In contrast, DmSNAP43, as seen previously (16), cross-linked to the wild-type probe as far as position 40, two turns of the DNA helix beyond the PSEA. The strongest cross-links were at positions 28 and 30.

Interestingly, probes that contained the mutant PSEB substitution also cross-linked to DmSNAP43 at positions downstream of the PSEB (Figure 2B, lower left panel). The differences in the pattern of relative cross-linking intensities could be noted relative to the wild-type probes. For example, when the cross-linking at positions 20 and 22 were used for purposes of normalization, the wild-type probes cross-linked more strongly at positions 28 and 30 than did probes that contained the PSEB mutation. Differences were sometimes observed at positions 34, 36 and 40, but results with these weaker bands were less consistent from experiment to experiment.

When cross-linking experiments were carried out with the template strand, a similar set of results was obtained (Figure 2B, right side). SNAP43 cross-linked downstream of the U1 PSEA regardless of whether the probe contained the wild-type or mutant PSEB. At some positions, differences in the cross-linking intensity could be observed. For example, positions 29, 31 and 33 cross-linked more intensely in the wild-type probe than did the DmSNAP43-mutant PSEB. Conversely, positions 27 and 35 cross-linked with relatively greater intensity in the mutant PSEB probe.

In summary, the results of the photo-cross-linking experiments indicate that the U1 PSEA by itself is capable of bringing DmSNAP43 into close proximity to the DNA that encompasses the PSEB. Although the PSEB is not required for these contacts, the PSEB appears to subtly affect the way in which DmSNAP43 interacts with this region of the DNA.

**The PSEB contributes to U1 promoter efficiency in vivo**

Because mutation of the PSEB did not abolish the interaction of DmSNAP43 with the DNA downstream of the U1 PSEA, we wished to determine the extent to which the PSEB contributes to transcriptional efficiency in living cells. To do this, we carried out transient transfections in triplicate of *D. melanogaster* S2 tissue culture cells using U1 promoter constructs that contained substitutions in either the PSEB or the PSEA (Figure 3). We used luciferase as a reporter gene because previous work indicated that the *D. melanogaster* U1 promoter efficiently drives the synthesis of functional mRNAs (17). The results of luciferase assays are shown in the column at the right side of Figure 3. As noted previously (17), mutation of the PSEA to an unrelated sequence (mutant PSEA) or to a U6 PSEA (which normally promotes efficient transcription by Pol III in the context of the U6 promoter) resulted in a reduction in reporter gene expression to 1–2% of the wild-type level (constructs B and C).
These results re-affirmed the critical importance of the U1 PSEA in promoting transcription by Pol II.

Mutation of the PSEB, on the other hand, to a GC-rich sequence (Construct D) reduced luciferase activity only by a factor of ~2. Interestingly, the replacement of the PSEB by a TATA-box (Construct E) similarly reduced expression ~2-fold. It is notable that the TATA-box worked only slightly better as a promoter element than the random GC-rich sequence (53 versus 45% relative expression efficiency). Although the contribution of the PSEB to U1 promoter activity was not dramatic, these results indicate that the PSEB represents a preferred sequence for U1 promoter activity.

The results with the final construct in Figure 3, which contains a TATA-box together with a mutant PSEA, indicates that the TATA-box alone, in the context of the U1 promoter, is insufficient for transcription in vivo. Thus, a TATA-box cannot compensate in vivo for inactivation of the PSEA. This contrasts with in vitro data in which substitution of a TATA-box for the PSEB increased transcription to 4.5-fold and made the high level of transcription independent of the PSEA (19).

To confirm the results of the luciferase assays, primer extension reactions were subsequently carried out. An autoradiogram of typical results is shown in Figure 4. Lane 3 illustrates the level of transcription obtained from the wild-type U1 gene promoter. In agreement with the results of the luciferase assays, an extensive mutation of the PSEA, as well as the conversion of the U1 PSEA to a U6 PSEA, reduced transcription from the U1 promoter to nearly undetectable levels (lanes 1 and 2). On the other hand, when the PSEB was changed either to a GC-rich sequence or to a TATA-box, a more modest reduction in transcription was observed (lanes 4 and 5). Furthermore, when the TATA substitution was combined with an inactive PSEA, transcription was almost negligible (lane 6). Thus, as observed with the luciferase assays, the PSEB was still required for transcription in vivo even when the PSEB was converted to a TATA-box.

A quantitation of primer extension data from three independent transfections are shown in the column at the right of Figure 4. The quantitative results of the primer extension assays were in close agreement with the results of the luciferase assays (Figure 3). The greatest discrepancy was with Construct D, but even in this case only a 2-fold difference was observed between the two types of assays. Though the effects of mutating the PSEB were relatively modest, the PSEB was reproducibly a 2- to 4-fold better U1 promoter element in vivo than either a mutant PSEB or a TATA sequence.

Can the PSEB substitute for a TATA-box in the U6 promoter?

Since mutation of the PSEB to a TATA-box had only a modest effect on U1 promoter activity in vivo, it became of interest to
Figure 5. The PSEB cannot effectively substitute for the TATA-box in the U6 promoter. The U6 promoter was used to drive expression of a reporter U6 maxi-gene. The wild-type U6 promoter construct (G) and three promoter substitution constructs (H, I and J) were used to transiently transfected S2 cells. RNA was isolated and used for primer extension assays. An autoradiograph of typical results is shown in the panel at the left. The relative promoter activities determined by phosphorimaging are shown in the column at the right. Errors are the standard deviation of the mean for three separation experiments.

determine whether a PSEB might substitute effectively for the TATA-box of the U6 snRNA gene promoter. To examine this possibility, the U6 maxi-gene reporter constructs diagrammed in Figure 5 were prepared and used for transfection of S2 cells.

RNA was prepared and U6 promoter activity was measured by primer extension assays using a DNA primer specific for RNA transcribed from the maxi-gene constructs (17).

The results of a typical primer extension assay are shown in the panel at the left side of Figure 5. Lane 1 depicts the level of transcription obtained from the wild-type U6 promoter. When the 8 bp TATA sequence was changed to the sequence of the PSEB (CA1GGAAA), transcription was still evident upon a long exposure but was dramatically reduced (Lane 2). As a baseline for comparison, mutation of the U6 PSEA either to a U1 PSEA or to an unrelated sequence reduced transcription nearly to undetectable levels (lanes 3 and 4). The transfections were performed in triplicate and the results of the primer extensions were quantified by phosphorimaging. These data are shown in the column at the right side of Figure 5. Either a complete mutation of the U6 PSEA or its substitution with a U1 PSEA reduced transcription more than 100-fold. Mutation of the TATA-box to the PSEB reduced transcription ~30-fold. Thus, although mutations of the TATA-box to a PSEB may not be quite as detrimental as mutations in the U6 PSEA, the PSEB is a very poor substitute for a TATA sequence. Consequently, the TATA-box appears to play a more significant role in expression of the D.melanogaster U6 gene than the PSEB does in expression of the U1 gene.

DISCUSSION

Earlier photo-cross-linking studies indicated that the U1 PSEA, but not the U6 PSEA, was able to bring the DmSNAP43 subunit of DmSNAPc into contact with the downstream PSEB. However, those studies did not address whether the PSEB played a role in those protein–DNA interactions. A mutation that changes every nucleotide in the PSEB indicates that the U1 PSEA alone contains sufficient information to bring DmSNAP43 into close contact with the downstream DNA. However, the PSEB mutation did produce more or less subtle changes in the cross-linking pattern. This suggests that the PSEB contributes some measure of specificity to the interactions that take place.

The results of the expression assays are in accord with this concept. The PSEB is not an essential element for transcription; rather it seems to act as a modifier of transcriptional activity. Although not essential, the wild-type PSEB provides a 2- to 4-fold higher transcription efficiency than a random GC-rich sequence or a TATA-box. The interaction of DmSNAP43 (as part of DmSNAPc) with the PSEB may in some ways be analogous to the interaction of TFIIIB with sequences just upstream of the TATA-box of mRNA genes. TFIIIB has a weak but non-essential preference for a consensus sequence (the BRE) adjacent to the 5' edge of the TATA-box; the more optimal BRE sequences result in stronger promoter activities (24).

To our knowledge, conserved elements analogous to the PSEB have not been identified in the Pol II-transcribed snRNA genes of other metazoans in which functional studies have been carried out. However, we have noted that snRNA genes of other insects contain conserved nucleotides in this location (unpublished data). It is possible that fruit flies (and other insects) have taken advantage of utilizing the specificity of the PSEB to modulate the strength of snRNA promoters over an evolutionary time scale. For example, three variant U5 genes, which are probably expressed at low levels, have PSEBs that are among the most divergent from the consensus PSEB (Figure 1C).

Interestingly, substitution mutations in the −33 to −20 region of a human U2 gene have been found to have minor effects on the transcription start site (25). It is therefore possible that the D.melanogaster PSEB, which is located within this region, may play a role in helping to establish the correct start site. Earlier work from our lab (19) indicated that transcription of the D.melanogaster U1 gene requires the TBP. Due to the location of the PSEB (~25 to −32) and its 8 bp length, it seems possible that the PSEB may be a site of DNA interaction with TBP. The PSEB may represent a ‘compromise’ sequence that allows it to be co-occupied simultaneously both by DmSNAP43 and TBP.

To see if this might be possible, we modeled TBP bound to DNA as if the PSEB were a TATA-box (Figure 6) (26,27). Then, taking into consideration that the PSEA is separated from the PSEB by exactly 8 bp, we identified the sites in red color where DmSNAP43 would cross-link with the DNA. Two different orientations are shown in the figure to fully reveal the sites of DmSNAP43 cross-linking relative to the bound TBP. The modeling illustrates that the phosphates that cross-link to DmSNAP43 are not occluded by TBP, and further suggests that DmSNAP43 could interact with the DNA both ‘behind’ and ‘beneath’ TBP. Further experiments will be required to examine the validity of this working model.
ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation Grant MCB-0113151 and in part by the California Metabolic Research Foundation. C.L. was a recipient of an Arne N. Wick Pre-doctoral Research Fellowship from the California Metabolic Research Foundation. Funding to pay the Open Access publication charges for this article was provided by the California Metabolic Research Foundation.

Conflict of interest statement. None declared.

REFERENCES

ACKNOWLEDGEMENTS:

This Chapter, in full is a reprint of material as it appears in *Nucleic Acid Research*. Lai, H. T., H. Chen, C. Li, K. J. McNamara-Schroeder, and W. E. Stumph. 2005. 33:6579-86. The dissertation author was the primary investigator and co-author of this paper. Photo-cross-linking experiments were done by this dissertation author and Hsiang Chen did luciferase activity and primer extension experiments.
CHAPTER 2

Stoichiometry of the subunits of the *Drosophila melanogaster* small nuclear RNA activating protein complex
ABSTRACT

The small nuclear RNA activating protein complex (SNAPc) in eukaryotes is a multi-subunit transcription factor required for the expression of small nuclear RNA genes. This protein binds as a complex to an essential promoter element known as the PSE that is located approximately 40-65 base pairs upstream of the transcription start site. Definitive studies to examine the stoichiometry of the subunits of metazoan SNAPc free of DNA are lacking. More importantly, the stoichiometry of the polypeptide subunits in the protein complexed with DNA has not been examined. In the fruit fly Drosophila melanogaster, DmSNAPc contains three distinct polypeptide subunits: DmSNAP190, DmSNAP50, and DmSNAP43. To investigate the stoichiometry of these polypeptides, identical subunits were tagged with different epitopes and co-expressed in D. melanogaster S2 cells with each other and with the other two subunits. The ability of differently tagged but otherwise identical subunits to associate with each other into the same protein-DNA complex was assayed by band-shift and super-shift assays. The results strongly support the concept that there is only a single copy of each of the three DmSNAP subunits in DmSNAPc bound to DNA.
INTRODUCTION

The snRNAs known as U1, U2, U4, U5, and U6 comprise a highly abundant class of metabolically stable non-polyadenylated RNA molecules that are required for pre-mRNA splicing (1,2). These snRNAs are synthesized by RNAP II, with the exception of U6, which is synthesized by RNAP III (3). The snRNA genes in metazoans have promoter structures that functionally distinguish them from classical RNA polymerase II (Pol II) or RNA polymerase III (Pol III) transcription units (3-7). In animals, a unique cis-acting element termed the Proximal Sequence Element (PSE) is located approximately 40-65 base pairs (bp) upstream of the transcription start site and is essential for basal transcription of snRNA genes by either Pol II or Pol III (8). In insects, this element is termed the PSEA to distinguish it from the PSEB, which is a less-conserved non-essential promoter sequence located approximately 25-30 bp upstream of the transcription start site of the Pol II-transcribed insect snRNA genes (4,9,10). Although the Pol III-transcribed U6 snRNA genes contain canonical TATA boxes, the Pol II-transcribed snRNA genes lack TATA boxes (5,11).

The protein that recognizes the PSE (or PSEA) in animals has been characterized in the human and fruit fly systems. It is most commonly called the small nuclear RNA activating protein complex (SNAPc) (8,12), but it has also been termed PSE Binding Protein (PBP) (13,14) and Proximal Transcription Factor (PTF) (15,16). In humans, SNAPc is a complex that contains 5 distinct subunits known as SNAP190, SNAP50, SNAP45, SNAP43, and SNAP19 (12,17-22). Orthologs of three of these subunits (DmSNAP190, DmSNAP50, and DmSNAP43) were found to be integral
components of the *Drosophila melanogaster* PSEA-binding protein, DmSNAPc (formerly called DmPBP) (23). Interestingly, a highly divergent yet homologous complex of three polypeptides is required for Pol II transcription of the Spliced Leader RNA (SL RNA) in trypanosomes. This suggests that a SNAP-like complex originated very early in eukaryotic evolution and continues to play a role in the transcription of essential small nuclear RNAs in present-day eukaryotes.

Although it has often been presumed that there is one copy of each of the protein subunits in SNAPc, there is little information that directly and precisely addresses the stoichiometry of the subunits in the complex, particularly when SNAPc is complexed to DNA. In the fruit fly, several pieces of data raise the possibility that one or more of the subunits could be present in more than one copy in DmSNAPc. For example, even though the three subunits individually add up to a molecular mass of only 169 kDa, gel exclusion chromatography indicated that DmSNAPc eluted at a position corresponding to an apparent molecular mass of 375 kDa (24), which is sufficient for two copies of each subunit. Moreover, DmSNAP190 contains 4.5 Myb repeats that have sequence similarity to the DNA binding domain of c-Myb (21,23), and some Myb domain proteins, although not most, bind to DNA as dimers (25-27). Furthermore, photo-cross-linking data indicate that DmSNAP190 can be cross-linked to a region of DNA extending over at least 27 bp (28). This represents an unusually long region of DNA to be contacted by a single subunit of a DNA-binding protein. Similarly, DmSNAP43 can be cross-linked to a 25 bp stretch of DNA when DmSNAPc binds to a U1 PSEA (23). We therefore used variously tagged proteins in electrophoretic mobility shift and super-shift assays to examine the number of copies
MATERIALS AND METHODS

Source of DmSNAPs

Constructs that code for DmSNAP43, DmSNAP50 and DmSNAP190 cloned into the inducible Drosophila expression vector, pMT/V5-His-TOPO, a component of the Drosophila Expression System from Invitrogen, have been previously described (23). These include constructs that provide termination at the natural stop codon of the protein or that have the stop codon altered so that the V5 and 6His tags provided by the vector are included at the C terminus. Constructs that contain the FLAG-Myc-His tags were prepared by first removing the V5 epitope by digesting the appropriate constructs described above with BstBI and MluI. A double-stranded synthetic oligonucleotide coding for the FLAG and Myc epitopes was then cloned between these sites. The vectors provide copper-inducible gene expression under the control of the metallothionein promoter. Drosophila S2 cells were co-transfected with a combination of four plasmids (e.g., see Figs. 1A, 2A, and 3A) as well as a plasmid (pCoBlast, Invitrogen) to provide resistance to the antibiotic blasticidin. Following selection for 3-4 weeks, stably transfected cells were induced for 24 hr with copper sulfate and lysed in CelLytic M Reagent (Sigma). The tagged DmSNAPc was then partially purified by Ni column affinity chromatography, dialyzed against BCZ-100 (20 mM HEPES, 5 mM MgCl₂, 10 μM ZnCl₂, 200 μM EDTA, 100 mM KCl, 3 mM
DTT, 0.5 mM PMSF, 20% [by volume] glycerol), and concentrated approximately 10-fold by centrifugation in an Centricon-30 centrifugal filtration device (Millipore) to a suitable concentration for electrophoretic mobility shift analysis (EMSA). The full sequences of the tagged and untagged DmSNAP subunits are shown in Appendix A.

**DNA probe for EMSA analysis**

The radiolabeled probe used for the experiments reported in this paper contained the PSEA and PSEB sequences of the *D. melanogaster* U1:95Ca gene [previously called the U1 95.1 gene (4)]. It was prepared to have a covalently closed “dumbbell structure” to make is resistant to low levels of exonucleases still present in the nickel column fractions (4,29). The following two complementary 59-base-long oligonucleotides were each radiolabeled with [\(\gamma^{32}\)P] ATP and T4 polynucleotide kinase: 5'-TTGCAATTCCCAACTGTTTTAGCTGCTCAGCCATGGAAACCTGGCTACTTTCTAGCCA-3' and 3'-GCTTGGCTTTCCCAAGCAACGTTAAGGGTTGACCCAATACTGACGAGTCGGTACCTTTGGACCAAATC-5'. They were annealed in equimolar quantities and ligated with T4 DNA ligase; the closed circular oligonucleotide was then purified by gel electrophoresis (4).

**EMSA reaction conditions**

Protein-DNA or antibody-protein-DNA complexes were formed in a 12 μl reaction volume in a final buffer composition of BCZ-100. Samples also contained 1 μg of poly (dl-dC) • poly (dl-dC). Reaction mixtures were incubated at room temperature for 30 min prior to gel electrophoresis. When antibodies were included, they were added half way through the incubation period. The monoclonal antibodies
used were the following: Anti-V5 (Invitrogen cat. # 46-0705); Anti-FLAG (Sigma M2 cat. # F1804); and anti-Myc (Sigma M4439). Samples were electrophoresed in 5% (29:1 acrylamide/bisacrylamide ratio) native gels in a running buffer consisting of 0.025 M Tris, 0.19 M Glycine, 1 mM EDTA (pH 8.3).

RESULTS

Stoichiometry of DmSNAP190 in DmSNAPc bound to DNA

To investigate the stoichiometry of the subunits of DmSNAPc, various constructs were prepared as described in Materials and Methods and co-over-expressed in homologous *Drosophila* S2 cells from the copper-inducible metallothionein promoter. To investigate the stoichiometry of the DmSNAP190 subunit, S2 cells were simultaneously co-transfected with the four constructs shown in Fig. 2.1A, together with the plasmid pCoBlast to permit the selection of stably transfected cells using the antibiotic blasticidin. Following selection and copper sulfate induction, cells were lysed, and protein extracts were applied to nickel columns to purify DmSNAPc via the 6His tags at the carboxyl termini of the DmSNAP190 subunits.
Figure 2.1. Stoichiometry of DmSNAP190 in DmSNAPc. (A) Schematic representation of co-expressed DmSNAPc subunits. (B) Electrophoretic mobility shift analysis of nickel column purified cell extract from cell line co-expressing differently tagged DmSNAP subunits illustrated in (A). Reactions shown in lane 2 to 4 and 6 to 9 contained added antibodies as indicated above each lane.
The cells co-expressed two differently tagged forms of DmSNAP190. One form had tandem FLAG and Myc tags preceding the 6His tag, whereas the other form contained a V5 tag preceding the 6His tag (Fig. 2.1A). Western blots indicated that both tagged forms of DmSNAP190 as well as the other two untagged DmSNAPs co-eluted from the nickel column in the same elution fractions (data not shown). The tagged DmSNAPc thus obtained was employed for electrophoretic mobility shift and super-shift analysis with a radiolabeled DNA fragment that contained the PSEA and PSEB of the D. melanogaster U1:95Ca gene. To super-shift the complexes, monoclonal antibodies against the V5, Myc, and FLAG epitopes were used separately or in various combinations.

Fig. 2.1B shows the results of such an analysis to examine the stoichiometry of DmSNAP190 using nickel column purified extracts obtained from cells that were co-expressing the four constructs shown in Fig. 2.1A. Fig. 2.1B lanes 1 and 5 show the position of the shifted band that results from DmSNAPc alone binding to the DNA fragment. Addition of either Myc antibody (lane 2) or V5 antibody (lane 4) resulted in the appearance of a super-shifted band, but in both cases a portion of the labeled fragment remained at the original band shift position. However, when both Myc and V5 antibodies were added to the same reaction, all of the protein-DNA complex was super-shifted and appeared at the same position (lane 3). This indicates that all of the DmSNAPc activity in the nickel column fraction as expected contains either a Myc or V5 tag.

The results shown in Fig. 2.1B are representative of the pattern expected if there is only one copy of DmSNAP190 (that has either a Myc tag or a V5 tag) in the
DmSNAPc-DNA complex. The observed pattern is not consistent with there being two or more DmSNAP190 subunits present in the complex. For example, if there were two DmSNAP190 subunits in DmSNAPc, it becomes necessary to hypothesize that the band labeled “antibody super-shift” is due to the binding of two antibodies to the complex. But if there were two copies of DmSNAP190, “heterodimers” of Myc- and V5-tagged DmSNAP190 should exist in some of the DmSNAPc-DNA complexes, and in lanes 2 and 4 these would be expected to migrate to a position intermediate between the two observed bands. No evidence appears for the existence of such complexes.

Lanes 6-9 in Fig. 2.1B show results from similar protein-DNA incubations except that in these lanes antibodies against the FLAG epitope were also included. As shown in Fig. 2.1A, the FLAG epitope is present on the same polypeptide as the Myc epitope. As expected, FLAG antibody alone gave the same pattern as the Myc antibody alone (compare lanes 6 and 2). Likewise, the FLAG antibody and the V5 antibody together gave the same result as the Myc and V5 antibodies together (compare lanes 9 and 3). However, the FLAG antibody and the Myc antibody when added together (lane 7) caused a further super-shift resulting from the binding of both FLAG antibody and Myc antibody to FLAG/Myc-tagged DmSNAP190 present in the DmSNAPc-DNA complex. This clearly demonstrates that more than one antibody can bind to provide a double super-shift. Further addition of the V5 antibody (lane 8) super-shifted the lower band observed in lane 7 (which contained only V5-tagged DmSNAP190); in contrast, the V5 antibody had no effect on the double super-shifted band that resulted from the binding of the FLAG and Myc antibodies. These experiments thus provide no evidence for the presence of more than one DmSNAP190
subunit in DmSNAPc and concurrently strongly support the concept that there is only one DmSNAP190 subunit in the protein-DNA complex.

**Stoichiometry of DmSNAP50 in DmSNAPc bound to DNA**

To investigate whether DmSNAP50 is present in one or more than one copy in DmSNAPc, the constructs shown in Fig. 2.2A were co-expressed in stably-transfected *Drosophila* S2 cells. Following partial purification by nickel affinity chromatography, the tagged DmSNAPc was used for electrophoretic mobility shift and super-shift analysis (Fig. 2.2B). Only the upper part of the gel that shows the protein-retarded bands is shown in Fig. 2.2.

Addition of Myc antibody by itself super-shifted the majority, but not all, of the DmSNAPc-DNA complex (compare lane 2 with lane 1). Reciprocally, inclusion of V5 antibody alone super-shifted a minority of the complex (lane 4). This result indicates this cell line expresses more of the Myc-tagged DmSNAP50 than of the V5-tagged DmSNAP50. When both Myc and V5 antibodies were included in the incubation prior to loading onto the gel, the entire signal from the DmSNAPc-DNA complex was super-shifted (lane 3). These results paralleled those obtained with the tagged DmSNAP190 protein shown in Fig. 2.1B above. Due to the expression ratio of the differently tagged forms of DmSNAP50, most of the V5-tagged subunits should exist in “hetero” complexes with Myc-tagged DmSNAP50 if there were more than one copy of DmSNAP50 in DmSNAPc bound to DNA. However, there was no evidence for the existence of such “hetero” complexes with intermediate mobilities.
Figure 2.2. Stoichiometry of DmSNAP50 in DmSNAPc. (A) Schematic representation of co-expressed DmSNAPc subunits. (B) Electrophoretic mobility shift analysis of nickel column purified cell extract from cell line co-expressing differently tagged DmSNAP subunits illustrated in (A). Reactions shown in lane 2 to 4 and 6 to 9 contained added antibodies as indicated above each lane.
Addition of FLAG antibody to the reaction (Fig. 2.1B, lanes 6-9) further retarded the mobility of protein-DNA complexes that contained the Myc epitope but had no effect on complexes that contained the V5 epitope. These results are basically identical to those obtained for the DmSNAP190 subunit (Fig. 2.1B). Together, these results indicate that DmSNAP50, like DmSNAP190, is present in only one copy in DmSNAPc complexed with DNA.

**Stoichiometry of DmSNAP43 in DmSNAPc bound to DNA**

Finally, to investigate the stoichiometry of DmSNAP43 in the DmSNAP complex, similar experiments were carried out with differently tagged DmSNAP43 subunits. Since DmSNAP43 contacts DNA over a 25 bp region in the U1 promoter (10,23) but has no obvious DNA binding motif, it seemed possible that DmSNAP43 could be present in more than one copy. Therefore DmSNAPc was purified from stably transfected cells that co-expressed the constructs shown in Fig. 2.3A and was used in mobility shift assays (Fig. 2.3B).

The pattern of band shifts, super-shifts, and double super-shifts with tagged DmSNAP43 was essentially the same as seen above with tagged DmSNAP190 and DmSNAP50. The Myc-tagged version of the protein was again expressed better in these cells than the V5-tagged version, but bands arising from DmSNAPc carrying the V5-tagged version of DmSNAP43 were clearly visible. (Note that the order of lanes 6-9 with respect to antibody addition is different in Fig. 2.3B compared to Figs. 1B and 2B.) From these results, we conclude that there is only a single copy of DmSNAP43 in DmSNAPc bound to DNA.
DISCUSSION

Only one copy of each SNAP subunit in the DmSNAPc-DNA complex

The stoichiometry of the subunits of SNAPc from Drosophila or other organisms had not been systematically investigated prior to this work. We previously found that DmSNAPc fractionated by gel exclusion chromatography with a relative molecular mass of ~375 kDa relative to globular protein standards (24); this mass is sufficient to accommodate two copies of each subunit. Human SNAPc/PTF possesses a relative molecular mass of ~500 kDa by gel exclusion chromatography (16) but
~200 kDa by glycerol gradient sedimentation (8,16). The latter is clearly an underestimate since there are five different subunits in human SNAPc that individually have a combined molecular mass of ~343 kDa (12,17-22). It was also reported that the subunits in purified human PTF were present in similar molar amounts, although the ratios of the subunits in immunoprecipitated PTF varied 3-fold (19). In the unicellular parasite Trypanosoma brucei, the three SNAP subunits have been reported to co-purify in approximately stoichiometric amounts (30). In another trypanosome, Leptomonas seymouri, the three subunits of SNAPc have a combined molecular mass of 139 kDa (31) and by combining the results of gel exclusion chromatography and glycerol gradient centrifugation (32) an estimated molecular mass of 122 kDa. Thus, these data were very consistent with there being one copy of each subunit in trypanosomal SNAPc free in solution. In metazoans, the data were also consistent with a single copy of each subunit, but the data also allowed for other possibilities.

Regardless of the data determined for SNAPc free in solution, a more important question was the configuration in which SNAPc binds to DNA. It remained an open question whether eukaryotic SNAPc might bind to DNA as a dimer of a heteromeric complex. For example, DmSNAPc could potentially bind to DNA as a dimer of heterotrimerers. The data presented herein essentially rule out that possibility and provide direct evidence that there is only one copy of each DmSNAP subunit in the protein-DNA complex formed on the PSEA.

**Metazoan SNAPc as an elongated asymmetrical complex**

Given that there is only one copy of each subunit in DmSNAPc (and by inference in human SNAPc), it appears that metazoan SNAPc is a highly asymmetrical
complex. This is consistent with the fact that gel exclusion chromatography overestimates the molecular mass of SNAPc (16), but glycerol gradient sedimentation underestimates the mass of SNAPc (8,16). Moreover, DmSNAPc is in close contact with at least 40 bp of DNA when it binds to a U1 gene PSEA (10,23). If DmSNAPc were a spherical molecule, it would almost certainly have to significantly wrap/bend the DNA to remain in contact with four turns of the DNA helix. However, DmSNAPc only modestly bends the U1 PSEA (at an angle certainly less than 50° and more likely less than 20°) based upon circular permutation, mini-circle binding, and ligase-catalyzed circularization assays (33). Moreover, human SNAPc/PBP was found not to bend (14). Thus metazoan SNAPc is very likely an asymmetrical protein complex that is elongated along the axis of the DNA.

REFERENCES


25. Yoshioka, S., Taniguchi, F., Miura, K., Inoue, T., Yamano, T. and Fukuzawa,


CHAPTER 3

A new method for localizing sites within a protein that cross-link to specific nucleotide positions within a DNA recognition sequence: Application to determining the orientation of the large subunit of DmSNAPc bound to a U1 snRNA gene promoter
ABSTRACT

Transcription of U1-U6 snRNA genes in *Drosophila melanogaster* is dependent upon a multi-subunit transcription factor, DmSNAPc, which recognizes a proximal sequence element (PSEA) located approximately 40-60 base pairs upstream of the transcription start site. Three distinct subunits (termed DmSNAP190, DmSNAP50, and DmSNAP43) are present in DmSNAPc and are orthologous to three of the similarly named subunits of human SNAPc. When DmSNAPc binds to a U1 gene promoter, the largest subunit, DmSNAP190, contacts DNA over a length of at least 25 base pairs. From results presented in Chapter 2, we know that there is only one copy of DmSNAP190 present in DmSNAPc bound to the U1 promoter. In this chapter, I describe experiments aimed at determining the orientation in which DmSNAP190 binds to the U1 PSEA. Toward that end, I have developed a novel method to map sub-regions within a protein that cross-link to specific nucleotide positions within a DNA recognition sequence. To do that, I combined site-specific protein-DNA photo-cross-linking with chemical digestion of N- and C-terminally tagged DmSNAP190. My results indicate that the N-terminal half of DmSNAP190 contacts the 3’ end of the PSEA and suggest that the C-terminal half of DmSNAP190 contacts the 5’ end of the PSEA. This orients the N-terminal half of DmSNAP190 toward the transcription start site and the C-terminal half toward the upstream DNA.
INTRODUCTION

The *D. melanogaster* small nuclear RNA activating protein complex (DmSNAPc) binds to a 21 base pair PSEA located about 40-65 base pairs upstream of the transcription start site of all *D. melanogaster* snRNA genes (1-3). It consists of three subunits, DmSNAP190, DmSNAP50, and DmSNAP43 that are each present in one copy in DmSNAPc. Our lab has previously found, by using site-specific protein-DNA photo-cross-linking, that the three subunits together are in close proximity to at least 40 base pairs of DNA. The largest subunit, DmSNAP190, can be cross-linked to an area that extends over at least 25 base pairs of DNA (see Fig. 3 in the General Introduction to this dissertation).

DmSNAP190 is 721 amino acid residues in length (calculated MW = 84 kDa). It contains an unusual yet evolutionarily conserved domain between residues 195 and 436 that consists of 4.5 Myb repeats named Rh, Ra, Rb, Rc, and Rd (Fig. 3.1). Canonical Myb domain proteins normally contain three Myb repeats (R1, R2, R3), and the second and third repeats are involved in sequence-specific recognition of DNA (4). Each of these two repeats forms a helix-turn-helix motif, and the recognition helices of each repeat are directly juxtaposed to each other end-to-end while lying in the major groove of the DNA (5). Other families of Myb domain proteins contain either two repeats (R2, R3) or a single Myb repeat (1R) (6-8). Proteins with a single Myb repeat generally bind to DNA as a homodimer (5,7). Authentic Myb proteins binds to pyAACNG consensus sequences in DNA, with usually two or, rarely, only one of the Myb repeats being sufficient to confer binding (10,13-16).
To our knowledge, DmSNAP190 and its orthologs from other species (9) are the only Myb domain proteins that contain 4.5 tandem Myb repeats. As mentioned above, DmSNAP190 can be cross-linked to at least 25 base pairs within and downstream of a U1 PSEA sequence. The consensus *D. melanogaster* PSEA contains a canonical Myb recognition sequence PyAACNG (10,11) near its center extending between positions 8-13 (12). Because DmSNAP190 is in close proximity to at least 25 base pairs of DNA, it would be extremely interesting to understand how the 4.5 Myb repeats are situated longitudinally along the axis of the DNA. In other words, assuming that all of the Myb repeats are involved in protein-DNA contacts (which may not be true), which repeats interact with the 5’ end of the PSEA, and which ones interact with the 3’ end of the PSEA?

Results from our lab indicate that the C-terminal domain of DmSNAP190 (residues 437-721, Fig. 3.1) is also necessary for DNA binding by DmSNAPc (Mitch Titus, submitted). Thus, it is possible that sequences C-terminal to the Myb domain may also be involved in contacting DNA. We would like to know if this is true, and if so, which part of the PSEA is contacted by this C-terminal domain that lacks Myb
repeats?

To begin to map and localize sites of protein-DNA contact within DmSNAP190, I have been developing a new method to localize sites within proteins that contact specific nucleotide positions within a DNA recognition sequence. This involves a combination of site-specific protein-DNA photo-cross-linking followed by chemical protein digestion. Protein fragments that cross-linked to the DNA are then identified by gel electrophoresis. This chapter describes initial experiments that provide evidence that the N-terminal half of DmSNAP190 interacts with the 3’ end of the PSEA and that the C-terminal half interacts with the 5’ end of the PSEA.

MATERIALS AND METHODS

Source of DmSNAPc containing untagged or N- or C-terminally tagged DmSNAP190

The overexpression of tagged forms of DmSNAPc in homologous S2 cells has been described in the previous chapter of this thesis and elsewhere (2). Briefly, stably transfected S2 cell lines were prepared that co-over-expressed all three DmSNAP subunits from the copper inducible metallothionein promoter. Three cell lines were used for the experiments described in this chapter. One cell line produced DmSNAPc with DmSNAP190 tagged at the C terminus with FLAG-Myc-6His tags that increased the molecular weight of the native DmSNAP190 by 5.658 kDa. A second cell line produced DmSNAPc with DmSNAP190 tagged at the N terminus with 6His-FLAG tags that increased the molecular weight of the native DmSNAP190 by 3.25 kDa. The
third cell line produced DmSNAPc with untagged DmSNAP190 but with the DmSNAP43 subunit carrying FLAG-Myc-6His tags at the C terminus to allow for purification of the complex. Detailed descriptions of the various DmSNAP constructs and the individual cell lines are provided in Appendices A & B.

Following induction with copper sulfate, cells were lysed in CelLytic™ M lysis buffer (Sigma). The tagged DmSNAP proteins were purified by nickel chelate affinity chromatography and were dialyzed against BCZ-100 (20mM HEPES, 100mM KCl, 5mM MgCl₂, 10μM ZnCl₂, 200μM EDTA, 3 mM DTT, 0.5mM PMSF, 10% (by volume) glycerol). Fractions were concentrated by centrifugation in a Centricon-30 centrifugal filtration device (Millipore) to a final protein concentration of approximately 0.5 to 2 mg/ml if needed. A detailed protocol for the purification and preparation of SNAPc fractions for photo-cross-linking is provided in Appendix C.

**Site-specific probe preparation**

Three different probes, each containing photo-cross-linking reagent at a unique position, were prepared to study DmSNAP190 interactions with the 5’ end, middle, and 3’ end of the U1 PSEA sequence. Phosphate positions 1, 12 and 24 were selected based upon strong DmSNAP190 photo-cross-linking signals and minimal interference from the cross-linking of other subunits. (Fig. 3.2)
Figure 3.2. Photo-cross-linking of DmSNAPc to positions 1-25 of a U1 gene PSEA (adopted from Wang et al., 1998 (3)). Phosphate positions selected for photo-cross-linking in this Chapter are position 1 from the non-template strand, and positions 12 and 24 from the template strand. Note that these positions all have a strong photo-cross-linking signal and minimal interference from cross-linking to DmSNAP50 and DmSNAP43. Their locations represent the 5’ end, middle, and 3’ end of the PSEA.
Probes were prepared as described by Wang and Stumph (3). Briefly, DNA oligonucleotides 23-25 bases long were synthesized with phosphorothioate incorporated 5’ of the third nucleotide from the 5’ end. The phosphorothioate-substituted oligonucleotides were derivatized with azidophenacyl bromide and then radiolabeled at the 5’ end by using $[^{32}\text{P}]$ ATP and T4 polynucleotide kinase. To prepare U1 probes with cross-linking reagent in the template strand or non-template strands respectively, the 79-mer (template) oligo-nucleotide 3’-CGTACTGGTAGTAATGCTTAGATAAGAATATTAAAGGTTGACCAAAATCGGCCATGGCGGTACCTTTCCATACCCTAGG -5’ and 81-mer (non-template) oligonucleotide: and 5’-ACGAATTCCATTCTTAATTTCCCAACTGTTTTTAGCGGTACCGCCATGGGAAAGGTATGGGATCCTCAATACCTCCGCGCATGCA-3’ were used. (The letters in bold represent the nucleotides of the PSEB and bold with underline represent the U1 PSEA). The oligo sequences are U1UP-1: 5’-AT*AATTCCC-AACTGGTTTAGC-3’, U1Low12: 3’-TAAGTAAGAATATTAAAAGTTGGA*CC-5’ and U1Low24: 3’-TTAAGGGTTGACCAAAATCGCCA*TG-5’ respectively (The asterisk symbol represents the replaced phosphate with phosphorothioate).

**Site-specific protein-DNA photo-cross-linking**

Photo-cross-linking reactions were carried out as previously described (1-3). Briefly, derivatized DNA fragments were incubated with partially purified DmSNAPc. After 30 min at 25°C in the dark, reaction mixtures were irradiated with UV light at 313 nm for 333 seconds followed by DNase I and S1 nuclease digestions. Normally
ten individual reactions were performed for each sample; eight reactions were used for hydroxylamine digestion and two were kept as untreated controls.

**Hydroxylamine protein digestion.**

Following the photo-cross-linking, samples were incubated in 1.8 M hydroxylamine, pH 9.0 at 45 °C for 5 hours in 12-14 kDa cut-off dialysis tubing. After hydroxylamine incubation, the same dialysis tubing was transferred to 1/10 X SDS PAGE loading buffer (5 mM Tris, pH 6.8, 250 mM urea, 0.2% SDS, 0.5% β-mercaptoethanol) for 6 hours dialysis. Samples were then collected, vacuum dried and reconstituted to 1X loading buffers with ¼ of dialyzed volume ddH2O and ¼ of dialyzed volume of 5X completion SDS PAGE loading buffer (50% glycerol, 25% β-mercaptoethanol, 0.125% bromophenol blue, 25% water). Each sample’s radioactivity was recorded before loading into the gel for cross-linking efficiency measurement. Samples, both hydroxylamine digested and untreated, were subjected to 7 or 15% SDS-PAGE. Gels were vacuum heat-dried and autoradiographed overnight. A detailed protocol for photo-cross-linking the protein-DNA complexes, carrying out the hydroxylamine digestion, and analysis by gel electrophoresis and autoradiography is included in Appendices F & G.

**RESULTS**

Hydroxylamine can be used to digest N-G (asparaginyl-glycyl) peptide bonds in proteins. There is a single N-G peptide bond in DmSNAP190 located between positions 358 and 359 (Fig. 3.1). This occurs very close to the middle of the
DmSNAP190 polypeptide and also near the middle of the Myb repeat labeled Rc. Because hydroxylamine cleaves so close to the middle of the protein, it would not be possible to distinguish the N- and C-terminal fragments directly by gel electrophoresis. Moreover, the photo-cross-linked polypeptide-DNA complex normally runs at a slightly higher molecular weight than expected due to the presence of a few cross-linked nucleotides. Finally, due to its length and flexibility, it is possible for the azidophenacyl group to cross-link to different amino acids, and this can slightly affect the mobility of the protein (1). To aid in identification of N- and C-terminal fragments, constructs were prepared that contained either an N-terminal or C-terminal tag. The fragments predicted after hydroxylamine digestion of native and of N- and C-terminally extended DmSNAP190 are shown in Fig. 3.3.

Figure 3.3. DmSNAP190 fragments predicted from hydroxylamine digestion. Red represents the evolutionarily conserved region also present in the human homologue. Green indicates non-conserved region. Blue segments represent the tags on C- or N-terminus. The N-tag increases the native form by 3.214 kDa. The C-tag increases the native form by 5.668 kDa
If the N-terminal half of the protein is radiolabeled following photo-cross-linking, then the phosphate at the defined nucleotide position must be contacting the N-terminal half of the protein. Conversely, if the C-terminal half of the protein is radiolabeled, then the C-terminal half of the protein must be close to the positionally-defined phosphate that contains the cross-linker. To aid in eliminating ambiguity, cross-linking reactions were carried out with DmSNAPc that contained untagged or either N- or C-terminally tagged DmSNAP190, and the products were run side-by-side on polyacrylamide gels following hydroxylamine digestion.

**Phosphate position 24 is contacted by the N-terminal half of DmSNAP190**

Fig. 3.4 shows the results of photo-cross-linking experiments of DmSNAPc to position 24 of the PSEA. The first three lanes show the results of cross-linking experiments prior to hydroxylamine digestion, and the last three lanes show the pattern following hydroxylamine digestion. Samples in lanes 1 and 4 were cross-linked to N-terminal tagged DmSNAP190; lanes 2 and 5 to untagged SNAP190; and lanes 3 and 6 to C-terminal tagged SNAP190.

Prior to hydroxylamine digestion, untagged DmSNAP190 (lane 2) ran with a slightly faster mobility than either of the tagged versions, and the N-terminal tagged version (lane 1) ran with a slightly faster mobility than the C-terminal tagged version (lane 3), in accordance with their relative molecular weights. Following digestion with hydroxylamine (lanes 4-6), products in the range of 41-48 kDa were expected. Importantly, the major bands on the gel appeared in that approximate molecular weight range in the gel. Significantly, the radiolabeled bands produced from untagged and C-terminal tagged DmSNAP190 ran with the same mobility (lanes 5 and 6). Even
more importantly, these bands had a faster mobility than the band produced with N-terminal tagged DmSNAP190 (lane 4). This is the pattern expected if the cross-link occurs to the N-terminal half of the protein that precedes the N-G peptide bond. If the cross-linking had occurred to the C-terminal half of the protein, faster-migrating bands would be expected in lanes 4 and 5 with a slower migrating band in lane 6. From these results, I am able to conclude that position 24 of the PSEA cross-links to the N-terminal half of DmSNAP190.

Figure 3.4. Phosphate position 24 of the U1 PSEA cross-links to the N-terminus of DmSNAP190. DmSNAPc containing N-terminal tagged DmSNAP190 (lanes 1 and 4), untagged DmSNAP190 (lanes 2 and 5), or C-terminal tagged DmSNAP190 (lanes 3 and 6) was cross-linked to position 24 of the U1 PSEA (see Fig. 3.2), run on a 15% polyacrylamide SDS gel, and radiolabeled bands were detected by autoradiography. Lanes 1-3 contain samples not subjected to hydroxylamine digestion whereas the samples in lanes 4-6 were digested with hydroxylamine. For further details, see the text.
Phosphate position 12 is also contacted by the N-terminal half of the DmSNAP190 subunit

Experiments were next performed with cross-linker incorporated at position 12 in the PSEA, and the results are shown in Fig. 3.5. Residual bands corresponding to undigested DmSNAP190 can be seen toward the top of the gel in the 95-100 kDa range. However, the more intense bands in the 45-50 kDa range represent the major products of hydroxylamine digestion. These bands exhibit the same pattern as observed with position 24: i.e., when the DmSNAPc used for cross-linking contains DmSNAP190 with an N-terminal tag, the mobility of the band is reduced (lane 1). From these results, I conclude that position 12 in the PSEA is contacted by the N-terminal half of DmSNAP190.

Much less intense bands are also visible in lanes 2 and 3 that have a mobility similar to the major band in lane 1. The origin of these bands is not certain. One possibility is that a certain fraction of the cross-link occurs to a different nearby amino acid position that results in a different mobility. We have observed such a double band before (1). It also could arise from cross-linking to a non-specific protein in the SNAPc fraction. We also cannot rule out that a small fraction of the cross-links might occur to the C-terminal half of DmSNAP190. However, if that were the case, we would expect the light band in lane 3 to have a slower mobility than the one in lane 2.
Figure 3.5. Phosphate position 12 of the U1 PSEA cross-links to the N-terminus of DmSNAP190. DmSNAPc containing N-terminal tagged DmSNAP190 (lane 1), untagged DmSNAP190 (lane 2), or C-terminal tagged DmSNAP190 (lane 3) was cross-linked to position 12 of the U1 PSEA (see Fig. 3.2), subjected to hydroxylamine digestion, run on a 15% polyacrylamide SDS gel, and radiolabeled bands detected by autoradiography. For further details, see the text.
Phosphate position 1 is most likely contacted by the C-terminal half of DmSNAP190

When cross-linking agent was incorporated at position 1 of the U1 PSEA, there was a change in the cross-linking pattern (Fig. 3.6). The radiolabeled fragment had a slower mobility with C-terminal tagged DmSNAP190 (lane 3) as compared with N-terminal tagged DmSNAP190 (lane 1). This strongly suggests position 1 cross-links to the C-terminal fragment. However, an unexpected result was obtained with the untagged DmSNAP190 in that two bands were apparent (lane 2).

To try to better correlate the autoradiography bands with the hydroxylamine digestion products, western blots were carried out with an antibody that recognizes the C-terminus of DmSNAP190. The corresponding bands are labeled in lanes 4-6. The C-terminal fragments from the hydroxylamine digestion are readily apparent in each lane, and these correspond to the bands seen in lanes 1 and 3 and the faster mobility band in lane 2. The identity of the slower migrating band in lane 2 (denoted by an asterisk) remains unknown. It clearly runs with slower mobility than the tagged N-terminal fragment (lane 7). Although better data are needed for position 1, the results in Fig. 3.6 suggest that position 1 is recognized by the C-terminal half of DmSNAP190.
Figure 3.6. Phosphate position 1 of the U1 PSEA most likely cross-links to the C-terminus of DmSNAP190. Photo-cross-linking and hydroxylamine digestion was carried out as in Fig. 3.5, except the cross-linking agent was at position 1 of the PSEA and the samples were run on a 7% polyacrylamide gel. Following gel electrophoresis protein bands were transfered to PVDF membrane for western blot analysis (lanes 4-9). Bands are labeled to indicate whether they are from the C- or N-terminus. The membrane was then used for autoradiography (lanes 1-3). In lanes 4-6, DmSNAP190 was detected using antibodies prepared against a C-terminal synthetic peptide. In lanes 7-9, anti-FLAG antibodies were used. In lane 8, the FLAG antibody is detecting tagged DmSNAP43. The origin of the band in lane 2 denoted with an asterisk is unknown.
DISCUSSION

The preliminary results presented here indicate that when DmSNAPc binds to a U1 PSEA, the DmSNAP190 subunit is oriented such that its N-terminal half interacts with the central region and 3’ end of the PSEA, and the C-terminal half interacts with the 5’ end of the PSEA. This is consistent with several other pieces of data that by themselves are rather inconclusive:

1) Recent data from our lab (Mitch Titus, submitted) indicate that DmSNAP50, which binds to the 3’ end of the PSEA, interacts with the N-terminal region of DmSNAP190. This would suggest placement of the N-terminal region of DmSNAP190 toward the 3’ end of the PSEA.

2) In the fruit fly, DmSNAP43 also binds to the 3’ end of the PSEA. Mitch Titus (submitted) has shown that DmSNAP43 interacts with the Myb domain of DmSNAP190, but the region has not yet been mapped in any greater detail. However, in the human system, HsSNAP43 has been found to interact with the N terminus of HsSNAP190.

3) Also in the human system, a small region near the N-terminus of HsSNAP190 (residues 38-84) is involved in recruiting TBP to the U1 promoter, and presumably TBP binds downstream of the PSEA closer toward the transcription start site.

4) In the human system, the Rc and Rd repeats are the most important repeats for the binding of HsSNAPc to the PSE (17), and the most conserved region of the fly PSEA is the 5’ end. Thus, it is logical that the Rc and Rd repeats might bind to the 5’
end of the PSEA.

In this work, we have made use of a naturally occurring N-G peptide bond for hydroxylamine digestion that is located very near the center of the DmSNAP190 polypeptide. It is also near the center of the Rc repeat. In future work, it would be useful to map out the interactions of the individual Myb repeats with specific nucleotide positions within the PSEA. This is discussed further in the next section of this dissertation.

REFERENCES


CONCLUDING REMARKS
Formation of a transcription pre-initiation complex on *Drosophila* snRNA genes starts when DmSNAPc binds to a PSEA. Earlier work in our lab indicated that the U1 or U6 PSEA is the primary determinant for RNA polymerase specificity and that the PSEAs are not interchangeable. Only 5 out of 21 bases of the PSEA are different between the U1 and U6 PSEAs that we use in our experiments, and by changing the 3’ end sequence we can alter RNA polymerase specificity from one polymerase to the other *in vitro*. This dominant role of the PSEA in determining RNA polymerase specificity suggested to us that DmSNAPc might bind differently to the U1 and U6 gene PSEAs. Previous site-specific protein-DNA photo-cross-linking experiments revealed that the conformations of the protein-DNA complexes were different depending upon whether DmSNAP was bound to a U1 or U6 PSEA. Perhaps most interestingly, it was found that DmSNAP43 contacts about 20 more base pairs of DNA downstream of a U1 PSEA vs. a U6 PSEA.

In Chapter 1, I described work investigating if the PSEB, downstream of a U1 PSEA, plays a role in determining whether DmSNAP43 contacts this downstream DNA. The results show only subtle changes in the cross-linking pattern upon mutation of the PSEB. This indicates the PSEB is not required for these downstream contacts. The sequence of the U1 PSEA is sufficient to bring DmSNAP43 into contact with the downstream DNA, and it seems amazing that only 3-5 nucleotides in the PSEA can determine how DmSNAPc binds to the DNA up to 20 bp away.

Fig. G.5A in the General Introduction presents a summary of the combined photo-cross-linking data of Yan Wang and Cheng Li (1,2). It is worth noting that
many of the strong contacts downstream of the U1 PSEA [denoted by blue (and red) spheres in the top line of the figure] occur near and within the PSEB. The function of the 8 bp PSEB in transcription of *Drosophila* snRNA genes is unknown; however, mutation of the PSEB to an unrelated sequence reduced U1 transcription in vitro by 7-fold (3). In vivo, however, the reduction was about 4-fold (Chapter 1).

Because TBP is required for U1 transcription in vitro (3), it is reasonable to postulate that the PSEB is a “compromise” sequence for both TBP and DmSNAP43 contact. This assumption is supported by the finding that if we superimpose the DmSNAP43 photo-cross-linking signals to a TBP-DNA 3D model (Chapter 1 Fig. 1.6), both DmSNAP43 and TBP should be able to bind simultaneously to the PSEB. This suggests that the DmSNAP43 subunit resides very close to TBP and makes us believe that DmSNAP43 is involved in recruiting TBP to the PSEB.

This model is also consistent with the finding that the human homolog, SNAP43, is capable of interacting with TBP (4). However there is not yet any direct evidence in the *Drosophila* system to support this model. It would be interesting to carry out GST pull down assays to determine if DmSNAP43 and TBP are capable of specific interactions in solution.

Chapter 2 reported work that investigated the stoichiometry of the DmSNAPc subunits. By expressing the same subunit with different tags, I was able to demonstrate that there is only one copy of each subunit in DmSNAPc bound to DNA. Both DmSNAP43 and DmSNAP190 contact about 24 bases of the U1 promoter DNA. By molecular weight alone, it seems DmSNAP43 should be present in at least two copies.
to contact a length of DNA similar to that as observed for DmSNAP190. But this
dimer hypothesis turned out not to be true.

Prior to performing the experiments reported in Chapter 3, it was important to
eliminate the possibility that DmSNAP190 might be present in 2 copies in DmSNAPc
bound to DNA. If DmSNAP190 were present in two copies, then it is possible that
position 12 (as an example) could cross-link to the N-terminus of one DmSNAP190
and to the C-terminus of the other. Or, if two DmSNAP190 subunits were arranged in
a head-to-head fashion, then position 12 could cross-link to the C-termini and
positions 1 and 24 could both cross-link to the N-termini. The non-ambiguous
interpretation of such data would be very difficult.

In Chapter 3, I presented the initial experiments to investigate the orientation
of the single copy of the largest subunit, DmSNAP190, bound to the U1 promoter. By
using hydroxylamine to cleave the protein at a single site, I was able to digest
DmSNAP190 into two nearly equal halves, the N-terminus half 1-358 and the C-
terminus half 359-721 (Fig. C.1). By making use of N- or C-terminal tags, I was able
to differentiate which half was contacting individual phosphate positions following
photo-cross-linking. The results show the 3’ end of the PSEA contacts the N-terminus
half of the protein (1-358) and the 5’ end of the PSEA most likely contacts the C-
terminus (359-721) region.

The N-G site of DmSNAP190 is near the middle of repeat 3 (Rc Fig. C.1), and
this gives a reference point to identify the orientation of the 4.5 Myb domains in this
conserved peptide region. Based upon my data, I strongly suspect that the order of
repeats is RdRcRbRaRh going from the 5’ to 3’ direction. In the future, it should be
possible to determine this at high resolution by using site-directed mutagenesis to remove the naturally occurring N-G site and introduce new N-G sites into DmSNAP190 within the linker regions that separate the individual repeats (Fig. C.1B italicized region). Introduction of an N-G site just beyond the Rd repeat would also make it possible to determine if the C-terminal domain, which is required for DmSNAPc binding to DNA (Mitch Titus, submitted) actually contacts the DNA.

Figure. C.1. (A) DmSNAP190 amino acid sequence. Red font indicates the evolutionarily conserved Myb repeats. (B) An alignment of the Myb repeats (Rh, Ra, Rb, Rc and Rd). The only N-G site for hydroxylamine digestion is located at position 358-359 (underlined). The italic regions at the end and beginning of each repeat represent areas where it may be possible to introduce new N-G sites for a high-resolution determination of which individual Myb repeats contact specific nucleotide positions in the PSEA.
From more extensive sequence comparisons (5) it appears that the conserved sequence of the *D. melanogaster* PSEA may actually extend 2-3 bases further in the 5’ direction than we have considered for the classical 21 bp PSEA. Thus it is possible that positions -3, -2, -1 or 0 upstream of position 1 of the PSEA may be suitable for photo-cross-linking, and may help resolve the ambiguities associated with position 1 cross-linking (Fig. 3.6 in Chapter 3).

It should also be possible to use these techniques to map out sites of protein-DNA contact for DmSNAP43 and DmSNAP50 (Fig. C.2). DmSNAP50 contains an unorthodox zinc finger at its C-terminus. Shu-Chi Chiang in our lab has recently shown that this region of DmSNAP50 is involved in interactions with DmSNAP190 and DmSNAP43. It would be interesting to know if this zinc finger region of DmSNAP50 is involved also in contacting DNA. Evidence from the human system (6) suggests it may.

DmSNAP43 has no recognizable DNA binding motifs, yet it interacts with a region of DNA ~24 bp in length. It would be very interesting to map out which regions of DmSNAP43 interact with the 3’ end of PSEA and which regions interact with the PSEB. In the absence of structural data from crystallography, such experiments will yield valuable information about the arrangement of DmSNAPc on the U1 (and U6) promoters.
Figure C.2. Diagram showing the positions of the naturally occurring hydroxylamine cutting sites (N-G) in all three DmSNAPc subunits. Red color represents the evolutionarily conserved regions. The molecular weights are shown for the predicted complete digestion fragments.

REFERENCES


APPENDIX

A. DNA constructs for expressing various forms of the DmSNAP subunits in S2 cells

B. Stably transfected S2 cell lines

C. Purification of DmSNAPs from S2 cells by nickel chelate chromatography

D. Detailed protocol for band-shift and super-shift assay

E. Detailed protocol for western blot

F. Site-specific protein-DNA photo-cross-linking of DmSNAP/PSEA using azidophenacyl bromide as cross-linking agent

G. Hydroxylamine digestion after photo-cross-linking reactions

H. Detailed protocol for dumbbell probe preparation
Appendix A: DNA Constructs for expressing various forms of the DmSNAP subunits in S2 cells

A. Vector used for cloning DmSNAPs

The genes that code for DmSNAP43, DmSNAP50 and DmSNAP190 have been placed into an inducible *Drosophila* expression vector, pMT/V5-His-TOPO, a component of the *Drosophila* Expression System (DES) from Invitrogen. This vector provides copper-inducible gene expression under the control of the metallothionein promoter. Individual clones have been prepared such that each protein can be expressed with or without a V5 monoclonal antibody epitope tag and 6xHis tags.

Nomenclature "with stop" or "stop" means that the gene includes the native stop codon, usually provide in the reverse PCR primer. "Without stop" or "no stop" means that we designed our reverse PCR primer to remove the native stop codon and maintain the frame through the DNA encoding the V5 and 6xHis tags. Constructs
were also prepared that had FLAG-Myc-6Hiis tags at the C-terminus or 6His-FLAG tags at the N-terminus.

**B. Sequences for the following constructs**

- pMT/V5-His-TOPO DmSNAP43STOP
- pMT/V5-His-TOPO DmSNAP43FMH
- pMT/V5-His-TOPO DmSNAP43V5H
- pMT/V5-His-TOPO DmSNAP50STOP
- pMT/V5-His-TOPO DmSNAP50FMH
- pMT/V5-His-TOPO DmSNAP50V5H
- pMT/V5-His-TOPO DmSNAP190STOP
- pMT/V5-His-TOPO DmSNAP190FMH
- pMT/V5-His-TOPO HFDmSNAP190STOP
- pMT/V5-His-TOPO DmSNAP190V5H
DmSNAP50FMH
pMT/V5-His-TOPO DmSNAP190V5H
Appendix B: Stably transfected S2 cell lines

Stable cell lines were prepared following the Invitrogen protocol with some modifications.

A 19:1 ratio of expression vector(s) to selection vector was used. Different amounts of recombinant DNA were used in different cell lines to optimize expression of the desired protein complex.

Day 1: Preparation

1. Seed 3x10⁶ S2 cells in a 100x20 plate in 10ml complete DES expression medium.
2. Grow 6-16 hours at 22-24°C until the cells reach 25-30% confluency.

Day 2: Transfection

3. Prepare the following transfection mix for a Corning 100x20 plate:
   - Promega ProFection mammalian transfection system calcium phosphate, catalog number: E1200.
   - In a microcentrifuge tube, mix together the following components. This will be Solution A.
Using cell line 202B as example

<table>
<thead>
<tr>
<th>Compound</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>2M CaCl$_2$</td>
<td>36 µl</td>
</tr>
<tr>
<td>pMTDmSNAP 190Stop (1 µg)</td>
<td>6 µl</td>
</tr>
<tr>
<td>pMTDmSNAP 50V5His (1 µg)</td>
<td>4.5 µl</td>
</tr>
<tr>
<td>pMTDmSNAP 50FlagMycHis (1 µg)</td>
<td>4.5 µl</td>
</tr>
<tr>
<td>pMTDmSNAP 43Stop (1 µg)</td>
<td>4.5 µl</td>
</tr>
<tr>
<td>pCoBlast(1 µg)</td>
<td>1 µl</td>
</tr>
<tr>
<td>Tissue culture sterile water</td>
<td></td>
</tr>
</tbody>
</table>

Bring to a final volume of 300 µl

In a second microcentrifuge tube, add 300 µl 2xHBS (50 mM HEPES, 1.5 mM Na$_2$HPO$_4$, 280 mM NaCl, pH7.1). This is solution B.

4. Slowly add Solution A drop wise to Solution B with continuous mixing. Continue adding and mixing until Solution A is depleted.
5. Incubate the resulting solution at room temperature for 30-40 minutes. After ~ 30 minute a fine precipitate will form.
6. Mix the solution and add drop wise to the cells. Swirl the plate to mix in each drop after it is added.
7. Incubate for 16-24 hours at 22-24°C.

**Day 3: Post-transfection**

8. Remove the calcium phosphate solution by using a sterile glass pipette sucking out the top solution in the plate. Add fresh complete DES expression medium (no selection agent blasticidin).
9. Incubate at 22-24°C for 2 days.

**Day 5: Selection**

10. Remove the medium by using a sterile glass pipette sucking out the top solution in the plate. Add fresh complete DES expression medium
containing 25 μg/ml blasticidin. Replace selective medium every 4 to 5 days until resistant cells start growing out (2 weeks).

+2 weeks: Expansion

11. Transfer resistant cells into new plates with medium containing blasticidin and pass cells at a 1:2 dilution (2 ml of fresh media to 1 ml of cells) to remove dead cells.

12. Once stable S2 cell lines have been obtained, the cells can be maintained in medium containing a lower concentration of blasticidin (10 μg/μl).

Table 1. Cell lines used in Chapter 2 and 3.

<table>
<thead>
<tr>
<th>101A1 (Ko-Hsuan, Hung)</th>
<th>102b (Anna Luy)</th>
<th>104 (Anna Luy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmSNAP43FMH</td>
<td>DmSNAP43FMH</td>
<td>DmSNAP43Stop</td>
</tr>
<tr>
<td>DmSNAP50Stop</td>
<td>DmSNAP43VH</td>
<td>DmSNAP50Stop</td>
</tr>
<tr>
<td>DmSNAP190Stop</td>
<td>DmSNAP50Stop</td>
<td>DmSNAP190FMH</td>
</tr>
<tr>
<td></td>
<td>DmSNAP190Stop</td>
<td>DmSNAP190VH</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>202b (Hsien-Tsung Lai)</th>
<th>504 (Mitch Titus)</th>
<th>505 (Mitch Titus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmSNAP43Stop</td>
<td>DmSNAP43Stop</td>
<td>DmSNAP43Stop</td>
</tr>
<tr>
<td>DmSNAP50FMH</td>
<td>DmSNAP50Stop</td>
<td>DmSNAP50Stop</td>
</tr>
<tr>
<td>DmSNAP50VH</td>
<td>DmSNAP190FMH</td>
<td>HFDmSNAP190Stop</td>
</tr>
<tr>
<td>DmSNAP190Stop</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Purification of DmSNAPs from S2 cells by Nickel Chelate Chromatography

MATERIALS
1) 0.1M CuSO₄ (sterile)(MW249.5 CuSO₄•5H₂O): 1.25g/50ml sterile double deionized H₂O (ddiH₂O), filter through corning 0.22 μM CA (cellulose acetate) sterilizing filter system.

2) Leupeptin (Roche Diagnostics Corporation, catalog number: 1017101) stock solution 5mg/5ml in ddi H₂O. Store at -20°C.

3) 10x BCZ(-) buffer For 100ml
   200 mM HEPES 4.77 g  HEPES free acid
   (Fisher Biotech, BP310-1, MW: 238.3)
   50 mM MgCl₂ 5 ml 1 M MgCl₂
   0.1 mM ZnCl₂ 1 ml 10 mM ZnCl₂
   2 mM EDTA 400 μl 0.5 M EDTA
Add dH₂O to 90 ml; adjust pH with KOH to 7.9. Bring the final volume to 100ml. Store at 4°C.

4) 1M MgCl₂ (MW: 203.31; MgCl₂• H₂O): 20.3 g/100mL. Store at 4°C.

5) 4 M KCl (MW: 74.56): 59.6 g/200ml H₂O. Store at room temperature.

6) 10 mM ZnCl₂ (MW: 136.28): 68.14 mg/5 ml H₂O. Store at room temperature

7) 1.0 M DTT (MW: 154.3): 154 mg/ml H₂O. Store at -20°C.
8) 100 mM PMSF (GIBCO BRL, 5521UB, MW: 174.19): 87 mg in 50 ml of 100% ethanol. Store at -20°C.

9) Stock solution A for 5X Native Purification Buffer: 250 mM monobasic sodium phosphate (NaH₂PO₄)(MW: 155.99, NaH₂PO₄•2H₂O) 2.5 M NaCl (MW: 58.5). Prepare by dissolving 7.8 g of monobasic sodium phosphate and 2.9g of NaCl in 200ml of deionized water. Store at 4°C.

10) Stock solution B for 5X Native Purification Buffer: 250 mM dibasic sodium phosphate (Na₂HPO₄)(MW: 141.96, Na₂HPO₄), 2.5M NaCl (MW: 58.5). Prepare by dissolving 7.1 g of monobasic sodium phosphate and 2.9g of NaCl in 200 ml of deionized water. Store at 4°C.

11) 3M imidazole (MW: 68.08, C₃H₄N₂): 10.2 g/50ml deionized water. Store at 4°C.

12) 5X Native Purification Buffer

Put stock solution B in a beaker with a stirring bar stirring. Add stock solution A drop by drop to stock solution B (only need very small amount), until pH is 8.0, this is 5X Native Purification Buffer. Store at 4°C.

13) 1X Native Purification Buffer

To prepare 100 ml of 1X Native Purification Buffer, combine:

- 20 ml of 5X Native Purification Buffer
- Add double deionized H₂O to 95 ml

Adjust pH to 8.0 with NaOH or HCl. Bring volume to 100ml with ddi H₂O. Store at 4°C.
14) Native Binding Buffer

Prepare the Native Binding Buffer with 10mM imidazole to reduce the binding of contaminating proteins.

- 30 ml of 1X Native Purification Buffer
- 100 μl of 3M Imidazole

Adjust pH to 8.0 with NaOH or HCl. Store at 4°C.

15) Native Wash Buffer

To prepare 50 ml of Native Wash Buffer with 20 mM imidazole, combine:

- 50 ml of 1X Native Purification Buffer
- 335 μl of 3M Imidazole

Adjust pH to 8.0 with NaOH or HCl. Store at 4°C.

16) Native elution Buffer BCZ-100 (750 mM Imidazole)

<table>
<thead>
<tr>
<th>For 50ml</th>
<th>1x BCZ</th>
<th>5 ml 10x BCZ (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% glycerol</td>
<td>5 ml 100% glycerol</td>
<td></td>
</tr>
<tr>
<td>100 mM KCl</td>
<td>1.25 ml 4 M KCl</td>
<td></td>
</tr>
<tr>
<td>3 mM DTT</td>
<td>150 μl 1.0 M DTT</td>
<td></td>
</tr>
<tr>
<td>0.5 mM PMSF</td>
<td>250 μl 100 mM PMSF</td>
<td></td>
</tr>
<tr>
<td>750mM Imidazole</td>
<td>12.5 ml 3 M Imidazole</td>
<td></td>
</tr>
</tbody>
</table>

Add deionized water to a final volume of 50ml in a conical 50ml screw cap tube.
Store at 4°C.

Expression of DmSNAPc in S2 cells

1. Grow 8 plates (Corning 100X20 mm tissue culture plates) of cells, in selective medium to 80-90% confluency.
2. Induce cells with copper sulfate; add to a final concentration of 0.5mM.
a) Prepare 0.1M sterile CuSO₄.

b) Use 5 µl 0.1M CuSO₄ per 1ml cells in media. (50 µl per 10ml)

c) Drop the 0.1M CuSO₄ solutions evenly into the plate, swirl the plate.

3. Incubate cells for ~24 hours.
   After 24 hours incubation (22-25°C), cells are ready for harvest.

**Harvesting the cells**

1. Harvest the cells by pipetting up and down to resuspend them in the media.

2. Transfer cells to a GSA centrifuge bottle; centrifuge the cells at 3000Xg for 5 minutes (4300 rpm in GSA rotor).

3. Use about 10ml media (supernatant) from step 2 to resuspend cell pellet, then transfer cell pellet to a 15ml screw-cap tube, centrifuge the cells at 1000Xg (2500 rpm) in Sorvall LegendRT for 5 minutes at 4°C.
   It is preferable to lyse the cells immediately and proceed with the purification. If absolutely necessary, the cell pellet can be frozen and stored in liquid nitrogen.

**Preparation of ProBond™ Columns**

Note: Do not use strong reducing agents such as DTT with ProBond • columns. DTT reduces the nickel ions in the resin. In addition, do not use strong chelating agents such as EDTA or EGTA in the loading buffers or wash buffers, as these will strip the nickel from the columns. Be sure to check the pH of your buffers at room temperature before starting.

When preparing a column as described below, make sure that the snap-off cap at the bottom of the column remains intact. To prepare a column:

1. Resuspend the ProBond resin in its bottle by inverting and gently tapping the bottle repeatedly.
2. Pipette or pour 2 ml of the resin suspension into a 10 ml Purification Column (Bio-Rad, catalog number: 731-1`550). This will give 1ml of settled resin. Allow the resin to settle completely by gravity (5-10 minutes) or gently pellet it by low-speed centrifugation (1 minute at 800 X g, ~2200 rpm) in Sorvall LegendRT centrifuge. Gently aspirate the supernatant.

3. Add 6 ml of sterile, distilled water and resuspend the resin by alternately inverting and gently tapping the column.

4. Re-pellet the resin using gravity or centrifugation as described in Step 2, and gently aspirate the supernatant.

5. For purification under Native Conditions, add 6 ml of Native Binding Buffer, working in the cold room.

6. Resuspend the resin by alternately inverting and gently tapping the column.

7. Re-pellet the resin using gravity or centrifugation as described in Step 2, and gently aspirate the supernatant.

8. Repeat Steps 5 through 7. Keep prepared column at 4°C.

Lysis of cells

1. Resuspend the cell pellet in 7 ml of CelLytic™ M cell lysis reagent (Sigma C2978-250 ml) with 70 µl protease inhibitor cocktail (Sigma P8340) and final 10 mM imidazole.

2. Lyse the cells by incubation at 4°C for 15 minutes. (From now on all procedures are carried out in 4°C environment.

3. Centrifuge the lysate at 13,000 X g (~12,100 rpm) in Marathon Micro A desktop centrifuge for 15 minutes to pellet the cellular debris.

4. Transfer the supernatant to a fresh tube.

5. Remove 200 µl of the lysate for future analysis and store in -80°C or liquid nitrogen. Temporally store remaining lysate on ice.
Purification

In 4 °C cold room:

1. Add 6.8 ml of lysate to a prepared purification column.
2. Bind for 2 hours using end-tilt-rolling gentle agitation to keep the resin suspended in the lysate solution (Do not end over end the solution, this will cause unwanted protein denature).
3. Settle the resin by gravity or low speed centrifugation (2 min. at 800Xg, 1833 rpm) in Sorvall LegendRT centrifuge, and carefully aspirate the supernatant. Save 200 µl of the supernatant (flow through from column).
4. Wash Column with 4 ml Native Wash Buffer by resuspending the resin with gentle inverting. Settle the resin by gravity or low speed centrifugation (2 min. at 800Xg) in Sorvall LegendRT centrifuge, and carefully aspirate the supernatant.
5. Repeat Step 4 three more times. Save 200 µl of each washes to monitor non-specific proteins binding (This may depend on cell lines).
6. Clamp the column in a vertical position and snap off the cap on the lower end. Elute the protein with 6 ml, add 1 ml each time, BCZ-100 (750 mM imidazole) elution Buffer. Collect 2 ml fractions into screw-cap tubes and aliquots 200 µl of each elution fractions into 1.5 ml screw-cap tubes (this will hold max. 2 ml). Store the Elution1- Elution3 in liquid nitrogen. Or these elution are subjected to dialysis.

Note: The elution fractions can be directly used for band-shift or super-shift at 2-12 µl per reaction, or the elution fractions can be concentrated and less volume used for band-shifts. For photo-cross-linking experiments, imidazole in the elution fractions will interfere with photo-cross-linking. Therefore, the elution fractions must be dialyzed, in BCZ-100, and concentrated if they are to be used for photo-cross-linking.
Procedure for dialyzing DmSNAPc

1. Select dialysis tubing with molecular weight cutoff (MWCO) 12-14 kDa, 16.0 mm diameter (Spectra/Por Dialysis Membrane), cut the tubing into 15-20cm pieces (depending upon how much protein you want to put in each tubing).
2. Boil the tubing in de-ionized water for half an hour, and then cool down in cold room overnight.
3. In the cold room, use dialysis clamp to close one end of the tubing, pipette the protein into the tubing, use another clamp close the other end.
4. Put the dialysis tubes into a beaker containing cold 1 L BCZ-100 (20 mM HEPES, 5 mM MgCl₂, 10 μM ZnCl₂, 200 μM EDTA, 100 mM KCl, 3 mM DTT, 0.5 mM PMSF and 10% glycerol (by volume)) with a stirring bar continuously stirring. The DTT and PMSF added immediately prior to dialysis. Dialyze against 2 more changes of 1 L BCZ-100 for 2 hours each for total 6 hours.
5. After dialysis, separate into aliquots 150 μl in 1 ml screw-cap tubes, freeze quickly in liquid nitrogen. Avoid frequent freeze/thaw cycle for this will denature protein.

Procedure for concentrate dialyzed DmSNAPc

If necessary, liquid nitrogen stored elution can be concentrated as described below.

1. Select Centricon YM-30 centrifugal devices (with molecular weight cutoff 30 kDa), insert sample reservoir into filtrate vial according to the manufacturer's instruction.
2. Add dialyzed protein to sample reservoir (2 mL maximum volume). Do not touch membrane with pipette tip. Seal the device by covering the sample reservoir with parafilm, use 25G needle to punch some holes on the parafilm to release the vacuum.
3. Place covered device and attached filtrate vial into an SS-34 centrifuge rotor; counterbalance with a similar device.
4. Spin **Centricon YM-30** centrifugal devices at 1000–5000 X g (3000-6500 rpm) in a SS34 rotor. Check the protein level every hour. When the protein left in sample reservoir is around 1/10 of the original volume, stop centrifugation. This may take 4-8 hours depending on the condition of the protein you want to concentrate.

**CAUTION:** Do not exceed centrifugation limits described in “Limitations” section.

5. Remove centrifugal filter assembly from centrifuge; then separate filtrate vial from sample reservoir.

6. Place retentate vial over sample reservoir and invert unit to recover the retentate. Centrifuge at 300–1000 Xg (1600-3000rpm) in a SS34 rotor for two minutes to transfer concentrate into retentate vial.

7. Remove device from centrifuge. Separate retentate vial from concentrator. The concentrated protein is in the retentate vial. Aliquot concentrated protein into 1 ml screw-cap tubes, each tube 150 µl, store in liquid nitrogen.
Appendix D. Protocol for band-shift or super-shift assay

Material:

1. 10x Non-circulation buffer [0.25M Tris, 1.9M Glycine, 10mM EDTA], pH8.3

   **Reagents**

<table>
<thead>
<tr>
<th></th>
<th>MW</th>
<th>1 liter</th>
<th>2 liters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>121.14</td>
<td>30.29g</td>
<td>60.58g</td>
</tr>
<tr>
<td>Glycine</td>
<td>75.07</td>
<td>142.64g</td>
<td>285.28g</td>
</tr>
<tr>
<td>EDTA</td>
<td>372.24</td>
<td>3.72g</td>
<td>7.44g</td>
</tr>
</tbody>
</table>

   ddH2O (add 800ml, then pH to 8.3 with concentrated HCl before bring to the final volume)

2. 40% (30:1) non-denaturing acrylamide stock solution

   **Reagents**

<table>
<thead>
<tr>
<th></th>
<th>100ml</th>
<th>500ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide</td>
<td>38.71g</td>
<td>193.55g</td>
</tr>
<tr>
<td>bis-acrylamide</td>
<td>1.29g</td>
<td>6.45g</td>
</tr>
</tbody>
</table>

   (filter solution)

3. 10% APS:

   Weigh 1.0 gram APS, add ddH2O to 10 ml.

4. Running Buffer

   1x non-circulation buffer (Note: use the same buffer for gel preparation)

5. poly d(I-C) or poly d(A-T) preparation, (from Pharmacia)

   - assume 50 µg/unit
   - use BCZ-100 as diluent
   - target dilution of 3 µg/µl
   - check OD260 reading of resuspended poly d(DNA): -2 µl diluted sample +
   - 158 µl of ddH2O
6. **10x BCZ(-) buffer**

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
<th>Volume</th>
<th>Buffer Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 mM HEPES</td>
<td>4.77 g</td>
<td>90 ml</td>
<td>HEPES free acid</td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>5 ml</td>
<td>1 M MgCl₂</td>
<td></td>
</tr>
<tr>
<td>0.1 mM ZnCl₂</td>
<td>1 ml</td>
<td>10 mM ZnCl₂</td>
<td></td>
</tr>
<tr>
<td>2 mM EDTA</td>
<td>400 µl</td>
<td>0.5 M EDTA</td>
<td></td>
</tr>
</tbody>
</table>

Add deionized H₂O to 90 ml, adjust pH with KOH to 7.9. Bring the final volume to 100 ml. Store at 4°C.

7. 1M MgCl₂ (MW: 203.31, MgCl₂·H₂O): 2.03 g/10ml. Store at 4°C.

8. 4 M KCl (MW: 74.56): 5.96 g/20ml H₂O. Store at room temperature.

9. 10 mM ZnCl₂ (MW: 136.28): 6.814 mg/50 ml H₂O. Store at room temperature.

10. 1.0 M DTT (MW: 154.3): 7.72 g/50 ml H₂O. Store at -20°C.

11. 100 mM PMSF (GIBCO BRL, 5521UB, MW: 174.19): 87 g in 50 ml of 100% ethanol. Store at -20°C.

12. **1x BCZ-100**

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x BCZ</td>
<td>5 ml</td>
<td>10x BCZ(-)</td>
</tr>
<tr>
<td>20% glycerol</td>
<td>10 ml</td>
<td>100% glycerol</td>
</tr>
<tr>
<td>100 mM KCl</td>
<td>1.25 ml</td>
<td>4 M KCl</td>
</tr>
<tr>
<td>3 mM DTT</td>
<td>150 µl</td>
<td>1.0 M DTT</td>
</tr>
<tr>
<td>0.5 mM PMSF</td>
<td>250 µl</td>
<td>100 mM PMSF</td>
</tr>
<tr>
<td>Add H₂O to 50 ml</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annealing synthetic oligonucleotides to make double stranded probe

10X MPS

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Volume</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 µl 1M Tris pH 7.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>125 µl 4M NaCl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 µl 1 M MgCl2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.7 mg DTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 µl of 50 mg/ml BSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>665 µl Water</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 ml total

Store at -20°C for up to 2 months.

This is a 60 µl total volume reaction in a 1.5 ml eppendorff tube or other similar tube.

1. Add 20 µl of 10X MPS buffer (recipe above) to tube.
2. Add water to 60 µl total volume after oligos are added
3. Add 6 µg each of the upper and lower oligos.
4. Final concentration will be 0.2 µg/µl.
5. After oligos are mixed, place in 95°C heat block or water bath for 5 minutes.
6. Move to 65°C water bath for 15 min.
7. Move to 37°C for 15 min.
8. Leave at room temperature for 15 min.
9. Make sure to spin down condensation from sides’ prior to use.

Labeling oligonucleotides with α-dCTPs

U1PSEA/B-22/-68Up and LO

Sequence of the oligonucleotides:

5' GTTCGTTGCAATTCACAACCTGTTTTTAGCTGCTCAGCCATGGAAACC 3'
3' AAGCAACGTTAAGGGTTGACCAAAATCGACGAGTCGGTACCTTTGGG 5'

To a 1.5 ml Eppendorf tube, add the following ingredients:
1. Up to 1 μg of the annealed oligonucleotide with available ends for labeling.
2. 2 μl each of up to 3 unlabeled cold dNTPs, 2 mM each dTTP, dATP, and dGTP, for 6 μl total.
3. 5 μl of 10 X Klenow buffer (supplied with enzyme).
4. 7 μl of labeled nucleotide dCTP.
5. Water to 49 μl total volume.
6. Add 1 μl of Klenow fragment- usually purchased from BRL. Make sure we have some prior to starting.
7. Incubate at room temperature for 30 min.
8. Stop reaction with 2 μl of 0.5 M EDTA.
9. Remove free nucleotides by running through a G-25 Sephadex quick spin column.

**To prepare and use spin column:**

Purify labeled probe from unincorporated nucleotides over a Sephadex G-25 column [Roche (Boehringer Mannheim)].

1. Remove spin column, and 2 collection tubes from refrigerator.
2. Remove top and bottom caps from column after checking that no column media is caught in cap (if it is, shake tube like thermometer to dislodge and continue).
3. Drain column by placing in first collection tube and emptying tube regularly until no more drains from tube.
4. Place the empty first collection tube back on bottom of column, and place entire assembly into 15 ml snap cap tube, spin in RT-6000 at 3200 rpm for 3 min.
5. Remove collection tube from column, and make sure that no liquid remains in bottom of column. If not properly spun, remaining liquid will dilute labeling reaction.

6. Apply the 50 µl. of labeling reaction to center top of column. Liquid will saturate the top few mm of column. If liquid is applied to edge, it will not properly move down the center of the column.

7. Place the second clean collection tube on the bottom of the column, place back into snap cap, and spin in Sorvall LegendRT at 3200 rpm for 3 min.

8. After spin, column should be radioactive,(usually more so than the tube), it should be discarded in radioactive waste.

9. Collection tube should be fitted with provided cap from kit, and a 2 µl aliquot removed and counted.

10. Usually 1 µg of labeled oligo yields around 50 million cpm total.

**Band- or Super-shift Gel**

1) Prepare a 5% non-denaturing polyacrylamide gel for the band-shift or super-shift assay. Prepare a solution of 7.5 mL 40% acrylamide, 6 mL 10x non-circulation buffer, and 46.5 mL dH₂O. Take 55 mL of this solution and add 400 µl of 10% APS and 40 µl of TEMED, mix well and pour a gel. Allow gel to polymerize for approximately 1 hr.

2) Prepare to set up binding reactions. All chemicals, proteins and antibodies must be kept on ice at all times.

3) Make appropriate DNA probe dilution: 50,000 CPM/µl using BCZ-100 as diluent.
4) Prepare probe master solution by adding 1 μl of 0.1 M DTT and 1 μl of diluted probe for every desired lane. Prepare enough for one more reaction than necessary.

5) DmSNAP: Each reaction should contain 10 μl of DmSNAP, according to the previous band-shift titration, add appropriate amount of BCZ-100 to adjust the final volume to 10ul. Dilute enough protein for one more reaction than necessary.

6) To prepare binding reactions, add components in the following order: 2 μl BCZ-100, 1 μl of 2 μg/μl poly d(I-C), 2 μl of master probe mix, 10 μl diluted DmSNAP where applicable. If it's a band-shift assay, allow DNA to bind to DmSNAP for 30 minutes in a water bath in a beaker adjusted to 25°C.

7) If it's a super-shift assay, allow DNA to bind to DmSNAP for 15 minutes at 25°C, then add 1 μl of the appropriate pre-immune serum, antiserum or antibody to the appropriate fractions, then allow to react for 15 minutes at 25°C.

8) Set up polyacrylamide gel by using 1x non-circulation band shift buffer. Allow gel to run without samples at 100 volts for 1/2 hr.

9) Load reacted mixes to gel and flank lanes with 10 μl of non-denaturing dye. Allow gel to run at 100 volts until fast dye is approximately (3/4) through the gel.

10) Expose gel to regular x-ray film with intensify screen at -80°C for 18 hr.
Appendix E: Detailed protocol for western blot

Stock solutions:

1) 40% acrylamide stock solution: (19:1; acrylamide: bis-acrylamide) Store at 4ºC.
   
   For 500 ml
   190 g acrylamide
   10 g bis-acrylamide
   add ddH2O to 500 ml, filter the solution.

2) Lower Buffer: Store at 4ºC.
   
   1.5 M Tris-HCl (pH 8.0)
   0.4% SDS

3) Upper Buffer: Store at 4ºC.
   
   0.5 M Tris-HCl (pH 6.8)
   0.4% SDS

4) 2 x Sample Buffer Solution: Store at -20ºC.
   
   0.2 mM Tris-HCl (pH 6.8)
   40% glycerol
   4% SDS
   0.56 M b-mercaptoethanol
   2 µg/ml of bromophenol Blue dye

5) 1x Electrophoresis Buffer: Store at RT.
   
   6.0 g Tris-base
   28.5 g of glycine
   1 g SDS
   Add ddH2O to bring the final volume to 1 liter.

6) Tris buffered saline (TBS): Add 6.05 gm Tris base (50 mM) and 29.22 gm NaCl (500mM) to 800 ml ddi water, adjust pH to 7.5 with HCl, adjust to 1 liter with ddi water.

7) TBS-Tween® 20 (TBST): Dilute 1 ml of Tween® 20 in 1 liter of TBS.

8) 1% (w/v) blocking solution: Add 1 g of BSA into 100 ml of TBST.
Gel recipes:

4.5 % stacking gel on a 10% PAA resolving gel:

<table>
<thead>
<tr>
<th></th>
<th>1 gel</th>
<th>2 gels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolving gel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30% PAA</td>
<td>7 mL</td>
<td>14 mL</td>
</tr>
<tr>
<td>1.5 M Tris pH 8</td>
<td>5.25ml</td>
<td>10.5 ml</td>
</tr>
<tr>
<td>ddi H₂O</td>
<td>8.5 ml</td>
<td>17 ml</td>
</tr>
<tr>
<td>10% SDS</td>
<td></td>
<td>420 µl</td>
</tr>
<tr>
<td>10% APS</td>
<td></td>
<td>400 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td></td>
<td>40 µl</td>
</tr>
</tbody>
</table>

Stacking gel

<table>
<thead>
<tr>
<th></th>
<th>1 gel</th>
<th>2 gels</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% PAA</td>
<td>1.3ml</td>
<td>2.6 ml</td>
</tr>
<tr>
<td>0.5 M Tris pH 6.8</td>
<td>2.5ml</td>
<td>5 ml</td>
</tr>
<tr>
<td>ddi H₂O</td>
<td>6 ml</td>
<td>12 ml</td>
</tr>
<tr>
<td>10% SDS</td>
<td></td>
<td>200 µl</td>
</tr>
<tr>
<td>10% APS</td>
<td></td>
<td>200 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td></td>
<td>20 µl</td>
</tr>
</tbody>
</table>

1. Layer stacking gel with t-amyl alcohol until polymerized. Pour off and rinse twice with ddi water. Blot water with filter paper. Pour stacking gel and let polymerize.

2. Run until bromophenol Blue marker is at the bottom of the gel.

3. Perform electrophoretic transfer of the proteins to PVDF membrane as follows:

4. Equilibrate gel in transfer buffer (Tris, Glycine, and Methanol (MeOH)) for 15 minutes.

5. Soak the PVDF membranes first in methanol (PVDF is so hydrophobic that it will float in transfer buffer, but once soaked in MeOH, it can then be equilibrated in transfer buffer. Soak gel pads and filter paper in transfer buffer. Make gel sandwiches and transfer for 1.5 hr. at 100 V.
6. Wash membranes for 5 min in TBS (with rocking).

7. Block 2 hours (usually 1 hour to 2 hours is okay) in 1% BSA blocking solution (1 g BSA in 100 mL fresh, refrigerated TBST).

8. Pour off the blocking solution and wash membranes in TBST twice for 5 minutes each wash.

9. Prepare a dilution of antibody (1:3000) - 1 µl in 3 mL BSA blocking solution.

10. Incubate membranes two hours in primary antibody solution.

11. Wash membranes three times in TBST, 5 minutes each wash.

12. Prepare a dilution of secondary antibody (1:5000) – 4 µl Anti-Rabbit IgG(Fc) AP conjugate (1mg/ml) in 20 mL BSA blocking solution.

13. Incubate membranes 1 hour in secondary antibody solution.

14. Wash membranes twice in TBST, 5 minutes each wash.

15. Wash in TBS (to remove Tween-20).

16. Pour on ~12 mL AP substrate (Promega Western Blue) and watch for color development.

17. Incubated between 1 and 10 minutes or until bands became distinct.

18. Wash membranes in water for at least 3 minutes to stop color development.
Appendix F: Site-specific protein-DNA photo-cross-linking of DmSNAPc/PSEA using Azidophenacyl Bromide as Cross-linking agent

All procedures involving azidophenacyl bromide, or its derivatives, are carried out in dim light, just enough light to see the label on the tubes and to be able to perform the experiments. All samples containing azidophenacyl bromide or its derivatives are stored in containers wrapped with aluminum foil and stored at -20°C or -80°C.

I. INCORPORATION OF AZIDOPHENACYL BROMIDE INTO OLIGONUCLEOTIDE

1. Preparation of azidophenacyl bromide (APB).

Weigh out 10 mg APB (Sigma CAT# A5407, 500 mg, Mw = 240.1) in dim light and dissolve in 1 ml of chloroform (42 mM). Pipette 100 µl aliquots (1 mg or 4.2 µmole) in 1.5 ml microcentrifuge tubes, and dry each aliquot in Speed-Vacuum. Store at -20°C. Immediately before use, dissolve one aliquot in 220 µl of methanol, and use 55 µl (250 µg, 1 µmole) for each of the incorporation reactions in step 3.

2. Phosphorothioate substituted oligo.

20-28 base oligonucleotides containing phosphorothioate 5' to the third nucleotide are commercially synthesized from Integrated DNA Technologies, Inc. Gel purify oligonucleotides on 20% polyacrylamide gels (29:1 acrylamide/bis acrylamide) containing 8 M urea, 1X TBE. Divide the purified oligo into 103 µg (~12.5 nmole) aliquots in 1.5 ml microcentrifuge tubes, dry each aliquot in heated Speed-Vacuum, and store at -20°C.
For determining the concentration of the oligonucleotides, use the equation: \( C \text{(pmole/µl)} = A_{260} \times 100/(1.5 N_A + 0.71 N_C + 1.20 N_G + 0.84 N_T) \) or \( C \text{(µg/µl)} = A_{260} \times 10^{-4} \times M_w/(1.5 N_A + 0.71 N_C + 1.20 N_G + 0.84 N_T) \) or \( C \text{(µg/µl)} = A_{260} \times 30/1000 \) for approximate calculation. \( C \) is the concentration of the oligonucleotides. \( A_{260} \) is the absorbance at 260 nm. \( N \) is the number of residues of base A, G, C, or T. \( M_w \) is the molecular weight of oligonucleotides. For the detail, see "Current Protocols in Molecular Biology", A.3D.1-3. Or \( A_{260} \times 3.3 \times \text{Dilution Factor (unit = pmole/µl)} \). The IDT Company also ships the oligonucleotide specification sheet with oligos.

3. Incorporation of cross-linking agent. (Remember to carry out all steps in dim light.)
   a. Dissolve 103 µg of dried phosphorothioate substituted oligo (~25-mer) in 1.5ml tube with 50 µl of ddH2O
   b. Add 5 µl 1 M potassium phosphate buffer (1M K₂HPO₄ 61.5 ml + 1M KH₂PO₄ 38.5ml, pH 7.0).
   c. Add 55 µl of APB in methanol from step 1 (250 µg, 1 µmole).
   d. Incubate 3 hrs at 37°C.

   Longer incubation times may cause side products of shorter oligonucleotides and may not improve incorporation yield. The incorporation yield, which ranges from 80~90%, can be estimated using a 15% denaturing polyacrylamide gel. The APB-incorporated oligo migrates slower than the phosphorothioate-substituted oligo.

**Note:** Four criteria for good denaturing condition for gel electrophoresis of oligos:
   a. Formamide containing LB (1-2X LB, 95% formamide, 18 mM EDTA, 0.025% SDS, Xylene Cyanol, bromophenol Blue).
   b. Urea (8M) containing gel.
   c. Warm up samples before running at 65°C.
   d. Run gel in warm temp (by high voltage).
4. Precipitation of oligo.

To the reaction above,

a. Add 11 µl of 3 M sodium acetate and 400 µl of ethanol, sit in ethanol-dry ice bath for 30 min, and spin 15 min in microcentrifuge.

b. Pipette out the solution. Re-dissolve pellet in 100 µl 0.3 M sodium acetate. Then add 300 µl ethanol, sit in ethanol-dry ice bath for 30 min then spin 15 min in microcentrifuge.

c. Wash the pellet using 70% ethanol

d. Re-dissolve in 25 µl of ddH₂O. Divide the solution into 12 µl aliquots and store in light-tight containers at -80°C. It is not necessary to dry the aliquots.

5. Determination of concentration.

Determine concentration on one aliquot by OD₂₆₀. Mix 2 µl of an aliquot from step 4d in 198 µl ddH₂O to make 1/100X dilution. Make further dilutions as needed to get an accurate OD₂₆₀ reading (between 0.3 and 0.8). Then make a 50 µl stock of the incorporated oligo solution with the final concentration 5 pmole/µl in ddH₂O. (see equations above, step 2)

II. PHOSPHORYLATION, ANNEALING, EXTENSION AND LIGATION

Phosphorylation

1. T4 polynucleotide kinase. (NEB. CAT# 201L, 2500 units, 10 U/µl)

2. Adenosine-5'-triphosphate [γ-³²P] crude (~7000 Ci/mmole: ICN CAT# 35020, 22.9 pmole/µl, 166 µCi/µl). Use 166 µCi/reaction.

or

Adenosine-5'-triphosphate [γ-³²P] (~6000 Ci/mmole; 222 TBq/mmol, 10 µCi/µl 0.05 ml of 50 mM Tricine (pH7.6), green. Perkin-Elmer CAT# BLU502Z250UC, BLU502Z500UC or BLU502Z001MC). Use 150 µCi/reaction.
3. Prepare:

- Eppendorff opener to open all tubes
- 12% denaturing PAGE for checking extension/ligation efficiency.
- 10% SDS PAGE
- 8% Native PAGE (Large) for purification
- dNTP (65°C thaw for 10 sec, vortex, spin down, keep on ice)

Take 3X2 = 6 μl APB-oligos to 1 set of eppendorff & 1 set of screw cap microtube.

<table>
<thead>
<tr>
<th>10X phosphorylation buffer</th>
<th>Stock</th>
<th>Amount (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 mM Tris-HCl (pH 7.6)</td>
<td>1 M</td>
<td>500 μl</td>
</tr>
<tr>
<td>100 MgCl2</td>
<td>2 M</td>
<td>50 μl</td>
</tr>
<tr>
<td>15 mM β-mercaptoethanol</td>
<td>1 M</td>
<td>15 μl</td>
</tr>
<tr>
<td>H2O</td>
<td></td>
<td>435 μl</td>
</tr>
</tbody>
</table>

4. Phosphorylation reaction. *(Shielded at all time)*

Component (mix well) | Rx (μl)
---------------------|---------
ddH2O                | 19 or 5 |
10X phosphorylation buffer [yes, use 2 μl] | 2 |
T4 polynucleotide kinase (BioLabs #0201S, 10U/μl) | 1 |
• γ-32P-ATP
  5 mCi Bottle; 166 μCi; 23.7 pmole or 5 |
  250, 500 or 1,000 μCi Bottle; 150μCi; 25 pmole | 15 |

Vortex gently, spin down, then react at 37°C for 60 min in the dark.

Turn on 65°C water bath in hood in dark room for annealing. (Turn on many hours ahead of time)
Annealing

1. "10X" annealing buffer: 400 mM Tris-HCl (pH 7.9), 500 mM NaCl, 100 mM MgCl₂.

2. A ~82 bases oligo is used as a template for annealing the upstream primer (Described below) and the oligo derivatives prepared from the previous step. Purify the template oligo on an 8% polyacrylamide gel (29:1 acrylamide/bisacrylamide) containing 8 M urea and 1 X TBE, it migrates around the slow dye. Dilute an aliquot of the purified template oligo to a concentration of 1 pmole/µl.

3. Upstream primer: the 20 mer upstream primer is complementary to the last 20 bases of the 3' end of the template. Purify it on 20% polyacrylamide (29:1, acryl/bis) containing 8 M urea and 1X TBE. Dilute an aliquot of the purified oligo to 5-pmole/µl concentration.

4. Annealing

Prepare two sets of annealing reactions. One is used as a control in the next step without adding ligase.

Below for 1 reaction:

<table>
<thead>
<tr>
<th>Sample</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>³²P-Labeled APB-oligo</td>
<td>25 (from phosphorylation step)</td>
</tr>
<tr>
<td>&quot;10X&quot; annealing buffer [yes, use 2 µl]</td>
<td>2</td>
</tr>
<tr>
<td>Upstream primer (5 pmole/µl; 20 pmole)</td>
<td>4</td>
</tr>
<tr>
<td>H₂O (sterile)</td>
<td>2</td>
</tr>
<tr>
<td>Template (1 pmole/µl)</td>
<td>2</td>
</tr>
</tbody>
</table>

35 µl

Heat to 65°C for 2 min; cool down slowly to 35°C over 1-1.5 hrs in 400-600 ml 65°C water in 1-liter glass beaker. (Leave beaker of water at RT not in water bath)
**Extension and ligation**

1. T4 DNA polymerase. (NEB CAT# 203L-750 units; 3U/μl)

2. T4 DNA ligase. (NEB CAT# 202L 100,000 units, 400 U/μl).
   (Note that NEB and Invitrogen define the ligase units very differently)

3. To the above annealed sample, add:

<table>
<thead>
<tr>
<th>Rx</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X BSA (10mg/ml)</td>
</tr>
<tr>
<td>10 mM dNTP</td>
</tr>
<tr>
<td>100 mM ATP</td>
</tr>
<tr>
<td>H₂O</td>
</tr>
<tr>
<td>T4 DNA polymerase (Last) 3U/μl</td>
</tr>
<tr>
<td>T4 DNA ligase (Last) 400U/μl</td>
</tr>
<tr>
<td>Mixture from annealing</td>
</tr>
</tbody>
</table>

   Total 43 μl

   - Keep components on ice before use & return to ice immediately after use
   - Spin > Gently shake > Spin
   - React at 37°C for 30-60 min. Extending the reaction time to 2 hours may improve the yield.
   - Return components to freezer.

4. Check the extension and ligation on a 12% denaturing PAGE.

   Clean & pre-run the gel for 30 min at 450V 200Watt (Gel should be warm but not hot too touch). Mix 1 μl reaction mixture with 5 μl denaturing loading buffer. Heat at 90°C for 5 min. Then load on a 12% polyacrylamide (29:1 acrylamide/bisacrylamide) containing 8M urea, 1X TBE gel. Also load ³²P-labeled 82-mer template as control on the same gel. Run for 45 minutes until fast dye migrates 2/3 of the length of the gel. Expose the gel for 1–2 hour. Compare the radioactive band of the reaction with the control, which does not contain T4 ligase.
or from previous labeled probes. On this denaturing gel, the band that migrates at approximately the same position as the 82-mer template is the ligated product. The non-ligated products run at the same positions as those in the non-ligated control sample.

III. PURIFICATION OF THE DERIVATIZED DOUBLE-STRANDED DNA

**IMPORTANT**: Every step must be performed behind the shield including gel electrophoresis. And each step must be monitored with Geiger counter.

Mix the remainder of the ligated sample (and non-ligated control) (~ 43 µl) with 10 µl of 60% glycerol (final glycerol: 10-15%). Load both the ligated and control samples on a 40x29 cm 8% polyacrylamide (29:1 acryl/bis)-1XTBE native gel with a 1 mm gel thickness and 1.2 cm well width. A 10 µl sample of 1X loading buffer (5X loading buffer: 0.3125 M Tris-HCl [pH 6.8], 50% glycerol, 0.125% bromophenol blue, 0.2% xylene cyanol) is loaded in a separate well as a migration control. Electrophorese at 500-600 volts at RT for 2.5 hrs. **IMPORTANT**: Keep the gel temperature below 45°C.

Stop electrophoresis when the fast dye has migrated 8 inches from the bottom of the sample well or slow dye has migrated 8 cm from the bottom of the sample well. Transfer the running buffer in the upper reservoir carefully into the specified radioactive waste container. Before taking the gel plates off the gel apparatus, use Kimwipes to clean all the residual buffer or solution on the plates and dispose Kimwipes in radioactive waste. Then lay the plates behind the shield. Separate the two glass plates with the help of a spatula, leaving the gel on the bottom plate. First cut the gel right between the fast dye and slow dye then dispose the bottom part of the gel that contains most of the undesired radio-labeled products and free 32P-nucleotides. This will reduce your potential radiation exposure in the following steps. Autoradiograph the upper part of the gel for one min. Determine
the desired band on the gel (Fully extended and ligated product) as described in the note below.

**Note:** In order to separate the ligated and nicked fragments of the same size, it is necessary to allow the long electrophoresis time. However, it is important to retain the radioactive nucleotides in the gel that will make the cleanup work much easier and greatly reduce contamination during the gel slicing procedure. The non-ligated control sample from step 4 should be included in order to differentiate the ligated double-stranded fragment from the nicked (non-ligated) one. The ligated fragment migrates faster than the nicked fragment of the same size (these difference may not be obvious if electrophoresis is stopped too early).

**Elution.**

Slice out the band containing the ligated fragment. Cut the slice into smaller pieces. Transfer them into a 2 ml microcentrifuge tube (AxyGen, Inc. SCT-200-C-S, 331-55-061, Cat. # 22-251). Add 150 µl of 0.1 X SSC (15 mM NaCl; 1.5 mM sodium citrate; 0.22 µM filtered) and put the tube in a 24-well beta rack (VWR CAT# 60985-428). Wrap the container with two-layer aluminum foil and place the rack in the incubator shaker. Elute the DNA at 200 rpm at 37°C for 4-5 hr, 300 rpm at 30°C for 4-5 hrs or 200-300 rpm at 25°C for overnight. Pipette the eluent into 1.5 ml microcentrifuge tube. Take 2 µl to measure the radioactivity by Beckman-Coulter LS 6500 Multi-Purpose Scintillation Counter for automatic counting. The counts from different samples vary from 150,000-200,000 cpm/µl depending upon the ligation yield. If a lower yield is expected, less volume of eluting buffer can be used (such as 100 µl).

**CAUTION:** Extreme care must be taken when working with radioactivity in dim light. Check the operating area after each step, especially after the purification and elution steps. Avoid opening tubes with fingers. Use tube opener at all times.
IV. PHOTO-CROSS-LINKING of DmSNAP

1. HEMG-100 buffer: 25 mM Heps, K⁺ (PH 7.6), 12.5 mM MgCl₂, 0.1 mM EDTA, 10% (v/v) glycerol, 100 mM KCl.

2. Nonspecific DNA carrier: prepare a solution that contains both 1 µg/µl poly (dI-dC)-poly (dI-dC) and 1 µg/µl poly (dG-dC)-poly (dG-dC) in TE buffer.

3. DNA probes: The oligo derivatives is prepared and purified as described above. Use 100,000 cpm for each reaction.

4. Source of DmSNAP. Drosophila SNF is subjected to DEAE cellulose and heparin agarose chromatography as previously described [Su et al., Europ. J. Biochem. 248, 231-237 (1997)]. The HA300 fraction, enriched in DmSNAPc activity, is dialyzed against two changes of HEMG-100 buffer over 5 hrs and concentrated 5-8 times using ultra free-15 Centrifugal Filter Devices (Millipore CAT#UFV2BGC10) to a final protein concentration of approximately 2.8 mg/ml. Thaw rapidly and immediately place on ice.

If DmSNAPc is used that has been prepared using nickel-chelate chromatography, the DmSNAPc must be dialyzed to remove imidazole. If DmSNAPc is used that has been prepared by FLAG affinity chromatography, the DmSNAPc must be dialyzed to remove 3X FLAG peptide. Imidazole and excess FLAG protein will interfere with the photo-cross-linking. In each case, the dialysis should be against 3 changes of HEMG-100.
5. Formation of DNA- DmSNAPc.

In a 1.5 ml microcentrifuge tube, mix the following solutions on ice.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Volume</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$O</td>
<td>(4-N) μl</td>
<td></td>
</tr>
<tr>
<td>HEMG-100</td>
<td>1 μl</td>
<td></td>
</tr>
<tr>
<td>Nonspecific DNA carrier</td>
<td>1 μl</td>
<td></td>
</tr>
<tr>
<td>DmSNAPc containing samples</td>
<td>4 μl</td>
<td></td>
</tr>
<tr>
<td>DNA probe</td>
<td>N μl (100,000 cpm)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 μl</td>
<td></td>
</tr>
</tbody>
</table>

Incubate in 25°C water bath for 30 min. The final concentration of reaction is 12.5 mM Hepes [pH=7.6], 50 mM KCl, 6.25 mM MgCl$_2$, 0.05 mM EDTA, 5% (v/v) glycerol.

If a specific competitor is used, incubate the specific competitor first with the reaction mix that includes nonspecific DNA carrier, HEMG-100, and DmSNAPc fraction on ice for 5 min before the addition of DNA probe.

6. UV-irradiation.

Open the reaction tubes and place them in an 80-tube-place rack. Then place the holder in a Spectro-linker$^\text{TM}$ XL-1500 UV Cross-linker (Spectronic Corporation) equipped with light tubes emitting 313 nm UV light. Irradiate for 10 min at room temperature with the top of the tubes about 5 cm from the bulbs.

7. DNase I digestion. (digest ds/ss DNA)

Before this step heat samples to 65°C for 3 minutes then cool down on ice for 1 minute. (This step was not in the original protocol and is optional)

To 10 μl of the reaction from the above step, add 0.5 μl of 0.2 M CaCl$_2$ (final ~10 mM), 0.5 μl of 4 mM PMSF (final ~0.2 mM), 0.5 μl of 100 μg/ml Leupeptin (final ~5 μg/ml) and 0.5 μl of 200 μg/ml Aprotinin (final ~10 μg/ml) respectively.
Then add 2 µl DNase I (Promega CAT#M6101, 1U/µl) with a final volume of 14 µl then incubate at 25°C water bath for 20-30 min.

8. S1 nuclease digestion. (digest ssDNA)

Add 0.7 µl of 10% SDS (final ~0.5%) to the above DNase I digestion reaction mix. Heat at 65°C for 3 min.

Then at room temperature add 0.69 µl of 1.75 M HOAc, (final ~73 mM), 0.57 µl of 30 mM ZnSO4 (final ~1 mM), 0.57 µl of 4 mM PMSF (final ~0.2 mM), and 1 µl (400 units) S1 nuclease (Invitrogen, CAT#: 18001-016, 400-1,500 U/µl) with the final volume of 17 µl. React at 37°C for 20-30 min. Caution: DO NOT use expired or < 400 U/µl S1 nuclease for they may give fuzzy bands.

9. SDS-PAGE.

Adjust pH to 6.8 by adding 1 µl of 1.5 M Tris-HCl (pH 8.8). Add 7.7 µl of 10 M urea. Finally add 6 µl of "5X" loading dye (0.3125 M Tris-HCl (pH 6.8), 50% glycerol, 10% SDS, 25% β-mercaptoethanol and 0.125% bromophenol blue). Heat at 90-100°C for 2 min.

Electrophorese on a SDS-10%-polyacrylamide gel. Run at 35 mA for stacking gel and 50 mA for running gel per gel. Dry the gel (~1.5 Hr) and autoradiograph overnight with fast film (Kodak Cat# 8264985). The % polyacrylamide of the gel can be adjusted to maximize separation of the bands of interest.
Appendix G: Hydroxylamine incubation after DNA digestion

- This experiment requires 10 reactions-worth of original photo-cross-linking experiment. You can prepare on 10-fold reaction initially, but these need to be aliquoted into 10 separate tubes just before UV irradiation. By doing this we can assure the UV’s penetration of each reaction. After UV irradiation, pool all 10 reactions back together into 1 tube. (Work with proper protection against $^{32}$P and corrosive hydroxylamine)

- Save 2 reactions-worth for untreated control. For each of these two reactions, adjust pH to 6.8 by adding 1 μl of 1.5 M Tris-HCl (pH 8.8). Add 7.7 μl of 10 M urea. Heat at 65°C for 3 min. Keep the samples at -20°C until the hydroxylamine-digested samples are ready to load on the gel.

Prepare Hydroxylamine 1.8 M Solution (wear gloves and eye protection)

Prepare immediately before use. Place 50 g hydroxylamine hydrochloride (NH$_2$OH.HCl) MW: 69.49 (Fisher, H330-500, 500 gm) in a 500-ml beaker. Add 200 ml of 5 M NaOH (final pH should be ~7.2. One time I did the pH was 11). If salt does not dissolve, continue adding NaOH until pH is 7 to 8. Add 10 g Na$_2$CO$_3$ and adjust pH to 9 with 12 M (concentrated) HCl (temperature will raise to ~40°C). Add double deionized-water to final 400 ml and readjust pH to 9 if needed. Transfer all 400 ml hydroxylamine to a 2L PYREX flask (wider opening is better for dialysis tubing-clamp passage). For more detail see Current Protocol in Protein Science, Chapter 11. Turn on the incubator and keep the 2L PYREX flask at 45°C.
Hydroxylamine digestion (all following works should be done behind beta-shield and follow radioactive material safety rules)

1. Dialyze samples against 1.8 M, pH 9, Hydroxylamine solution for 5 hours at 45°C (Innova 4335 shaker, 100 rpm, in flask) in 12-14 kDa MW cut off tubing (Spectrum, Spectra/Por 2 molecularporous membrane #132676, change this tubing if the desired protein fragments is smaller than 14 kDa). This might be change to use 10 kDa MW cut off Slide-A-Lyzer dialysis cassette (Pierce Prod# 66380). No need to do any changes of dialysis. The time of dialysis can be adjusted to maximize nearly complete digestion but minimizing non-specific breakdown products.

2. Remove the hydroxylamine salts. In the same dialysis bag, dialyze the samples at room temperature (or 25°C) twice for 2 hours each time against 500 ml of 1/10 X incomplete SDS gel loading buffer:
   6 mM Tris, pH 6.8
   250 mM urea
   0.2 % SDS
   0.5 % β-mercaptoethanol.

3. Following dialysis, measure and record the volume of the sample, transferring it to a microcentrifuge tube.

4. Concentrate the samples. Dry samples down to dryness under vacuum in the Speed-Vacuum Concentrator for 1.5 Hr approximately. (The water and the β-mercaptoethanol should evaporate away; the Tris, urea, and SDS should remain as solids.)

5. Readjusting digested sample’s volume. Take up the sample in 2/25 (4/5 x 1/10) volume (step 3) of water (compared to the volume measured following dialysis) and dissolve completely. (For example, if the volume following dialysis was 200 µl, take up the dried sample in 16 µl water).
6. Add \( \frac{1}{4} \) volume of 5X Completion Loading Buffer:

(For example, if the sample was taken up in 16 \( \mu \)l water, add 4 \( \mu \)l Completion Loading Buffer.).

5X Completion Loading Buffer:

- 50 % glycerol
- 25\% \( \beta \)-mercaptoethanol
- 0.125\% bromophenol blue
- 25 \% water.

7. Radioactivity. Count the radioactivity of samples before loading the gel (Count the entire sample, not an aliquot.). The photo-cross-linking efficiency of protein to derivatives DNA is about 5%, so these radioactivity readings can verify the lose \% of samples during experiment. And provide the approximately the time required for good auto-radiography (5000 cpm/digested sample can give good image overnight with fast film).

8. SDS-PAGE. Heat all samples, both treated and untreated, at 65 °C for 3 min. Electrophoreses on a 7-15% SDS-polyacrylamide gel. Run at 35 mA for stacking gel and 50 mA for running gel per gel. Vacuum Heat-Dry the gel for about one hour at 70 °C and autoradiograph overnight or longer.
Appendix H: Detailed protocol for dumbbell probe preparation

Reagents

- "10X" annealing buffer: 400 mM Tris-HCl (pH7.9), 500 mM NaCl, 100 mM MgCl₂.
- Two 59 bases oligo (below) are used as a template or non-template for annealing. Purify the template oligo on an 8% polyacrylamide gel (29:1 acrylamide/bisacrylamide) containing 8 M urea and 1 X TBE, it migrates around the slow dye. Dilute an aliquot of the purified template oligo to a concentration of 5 pmole/μl.
- Non-template (Up) strand:
  5′-TTGCAATTCCCAACTGGTTTTAGCTGCTCAGCCATGGAAACCTGGCTACTTTCTAGCCA-3′
Template (Low) strand:
  3′-GCTTGGCTTTTCCCAAGCAACGTTAAGGGTGACCAAATCGACGAGTCGGTACCTTTGG-5′

Annealed as below (Up: bold fonts, Low: regular fonts)

\[
\begin{align*}
\text{Up:} & \quad TCGGTTCGTTGCAATTCCCAACTGGTTTTAGCTGCTCAGCCATGGAAACCTGGCTACTTTCTAGCCA \\
\text{Low:} & \quad TCCCAAGCAACGTTAAGGGTGACCAAATCGACGAGTCGGTACCTTTGGACCGATCT
\end{align*}
\]

Phosphorylation

1. T4 polynucleotide kinase. (NEB. CAT# 201L, 2500 units, 10 U/μl)

2. Adenosine-5′-triphosphate \([\gamma^{32P}]\) (6000 Ci/mmmole): Perkin-Elmer CAT# BLU502Z250UC, BLU502Z500UC or BLU502Z001MC (10 μCi/μl).
3. Prepare:

10X phosphorylation buffer (1ml):

<table>
<thead>
<tr>
<th>Component</th>
<th>Stock</th>
<th>Amount (μl)</th>
<th>Concentration in 10 X stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl (pH 7.6)</td>
<td>1 M</td>
<td>500 μl</td>
<td>500 mM</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>2 M</td>
<td>50 μl</td>
<td>100 mM</td>
</tr>
<tr>
<td>β-mercaptoethanol</td>
<td>1 M</td>
<td>15 μl</td>
<td>15 mM</td>
</tr>
<tr>
<td>ddH₂O</td>
<td></td>
<td>435 μl</td>
<td></td>
</tr>
</tbody>
</table>

4. Phosphorylation reaction. *(Shielded at all time)*

Make Up or Low strand phosphorylation separately. Shown below is for ONE strand only.

<table>
<thead>
<tr>
<th>Component (mix well)</th>
<th>Rx (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>8</td>
</tr>
<tr>
<td>T4 polynucleotide kinase buffer</td>
<td>2</td>
</tr>
<tr>
<td>T4 polynucleotide kinase (BioLabs #0201S, 10U/μl)</td>
<td>2</td>
</tr>
<tr>
<td>γ-³²P-ATP (6000 Ci/mmole, 10 μCi/μl)</td>
<td>4</td>
</tr>
<tr>
<td>Dumbbell probe (5 pmole/μl; Up or Low strand)</td>
<td>4</td>
</tr>
</tbody>
</table>

Vortex & Spin down then react at 37°C for 60 minute.

Turn on 65°C water bath well ahead of time for annealing step.

**Annealing**

Mix two sets of labeling reactions. One is template and the other one is non-template.
Below for 1 reaction:

<table>
<thead>
<tr>
<th>Sample</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>32P-Labeled Up strand mixture from phosphorylation</td>
<td>20</td>
</tr>
<tr>
<td>32P-Labeled Low strand mixture from phosphorylation</td>
<td>20</td>
</tr>
<tr>
<td>&quot;10X&quot; annealing buffer</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42.3 µl</strong></td>
</tr>
</tbody>
</table>

Heat to 65°C for 3 min; cool down slowly to 35°C over 1-1.5 hrs in 400-600 ml 65°C water in 1-liter glass beaker.

**Ligation**

1. To the above annealed sample, add: Rx (µl)

<table>
<thead>
<tr>
<th>Component</th>
<th>Value (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA (10mg/ml)</td>
<td>0.8</td>
</tr>
<tr>
<td>100 mM ATP</td>
<td>1.6</td>
</tr>
<tr>
<td>T4 DNA ligase (Last) 1U/µl</td>
<td>1</td>
</tr>
<tr>
<td>Mixture from annealing</td>
<td>42.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>45.7 µl</strong></td>
</tr>
</tbody>
</table>

- In the control reaction replace T4 DNA ligase with water. Keep components on ice before use & return to ice immediately after use
- Spin Vortex Spin
- React at 37°C for 1Hr
- Return components to freezer.

4. Prepare two Sephadex G-25 column (Roche Applied Science Cat# 11273949001) by centrifugate them twice in 1,100 x g to remove excess buffer.
5. After 1 or more hour ligation reaction, extract DNA with 50 μl phenol/chloroform (1:1) vortex for 30 seconds to remove proteins.

6. Carefully recover supernatant to one Sephadex G-25 column. (DO NOT allow tip contact the beads or to sidewalls, place sample in the center)

7. Centrifuge the Sephadex G-25 column for 2 minutes at 1,100 xg.

8. Collect filtered sample and transfer to **second** Sephadex G-25 column.

9. Centrifuge the Sephadex G-25 column for 2 minutes at 1,100 xg.

10. Collect and measure the total volume of $^{32}$P-labeled dumbbell probe. Calculate the cpm and dilute an aliquot $^{32}$P-labeled dumbbell probe to 50,000 cpm/μl. Keep in -20°C with beta blocker to avoid non-necessary radio-exposure.