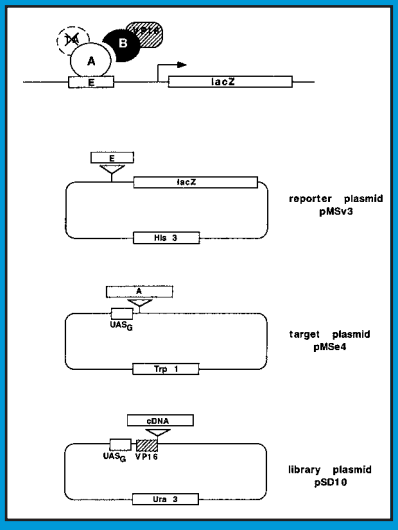
Methods in Molecular Biology™

VOLUME 130

Transcription Factor Protocols

Edited by

Martin J. Tymms



HUMANA PRESS

1

Isolation of Target Gene Promoter/Enhancer Sequences by Whole Genome PCR Method

Dennis K. Watson, Richard Kitching, Calvin Vary, Ismail Kola, and Arun Seth

1. Introduction

Regulation of gene expression is controlled through the combinatorial action of multiple transcription factors, which function to activate or repress transcription via binding to *cis*-regulatory elements. Such regulatory elements are usually present in the promoter sequences located upstream of the transcription initiation sites of a target gene. Identification of functional target gene promoters that are regulated by a specific transcription factor is one of the most critical areas in understanding the molecular mechanisms that control transcription. Furthermore, identification of target gene promoters for normal and oncogenic transcription factors provides insight into the regulation of genes that are involved in control of normal cell growth and differentiation, as well as provide information critical to understanding cancer development.

Methods based on subtractive hybridization and differential display-PCR (polymerase chain reaction) have been described for the identification of genes that are differentially expressed in tissues or cell lines ***(1–6)***. Approaches based upon differential gene expression, however, are likely to identify both indirect as well as direct gene targets for a transcription factor. Whole genome polymerase chain reaction (WGPCR) is a method that identifies direct target gene promoters/enhancer sequences for DNA binding proteins. Briefly, genomic DNA fragments are selected by their binding to a specific transcription factor and amplified by the polymerase chain reaction. Selection and amplification cycles are repeated multiple times, resulting in a pool of DNA fragments that are enriched for the specific transcription factor binding site. Among the variables that affect the experimental outcome are the quality of the transcription

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

factor, the specificity of DNA site, the specificity and affinity of the antibody, the size of the DNA fragment, and the complexity of the genomic DNA. WGPCR was applied to the identification of human DNA sequences that bind to the transcription factors p53 and IIIA (TFIIIA) ***(7,8)***. This chapter provides a detailed description of the whole genome PCR (WGPCR) method, and demonstrates its utility in the identification of Ets transcription factor target gene promoters.

Ets is a family of transcription factors present in species ranging from human to invertebrate, and all family members contain an 85-amino acid region, which has the DNA binding domain, designated the Ets domain ***(9–11)***. *Ets* family gene products bind specific purine rich DNA sequences with a core motif of 5'-GGAA/T-3', and transcriptionally regulate a number of viral and cellular promoters ***(9,12)***. The Ets proteins constitute an important family of transcription factors that control the expression of genes that are involved in various biological processes, including cellular proliferation, differentiation, development, transformation, and apoptosis ***(13–17)***. Ets products have also been implicated in several malignant and genetic disorders. For example, human Ets genes are located at the translocation breakpoints of several leukemias and solid tumors, forming chimeric proteins believed to be responsible for tumorigenesis ***(18–21)***. Recently, the overexpression of Ets2 in transgenic mice has been shown to cause skeletal abnormalities phenotypically similar to those seen in Down's syndrome ***(22)***.

The importance of the Ets family of transcription factors in various biological and pathological processes necessitates the identification of downstream cellular target genes of specific Ets proteins. Although some overlap in the biological function of different Ets proteins may exist, the presence of a family of closely related transcription factors suggests that individual Ets members may have evolved unique roles, manifested through the control and interaction of specific target genes. Previous methods for identification of Ets targets have mainly been based upon the presence of the purine rich GGAA/T core sequences in the promoters/enhancers of various cellular or viral regulatory regions ***(9,12)***. Subsequently, synthetic oligonucleotides containing Ets binding sites (EBS) were used in electrophoretic mobility shift assays (EMSAs), and transactivation assays using different Ets expression constructs together with reporter genes containing the minimum promoter linked to the prospective target genes Ets binding sites ***(23)***. We have recently used the whole genome PCR to identify gene promoters that are direct targets of Ets transcription factors ***(6)***. A diagram of the modified WGPCR strategy we have utilized is shown in **Fig. 1** and is described in detail in the Methods section. In the first step, total genomic DNA is digested with MboI to obtain a pool of DNA fragments that has an average size 250–500 bp. The MboI digested DNA is ligated

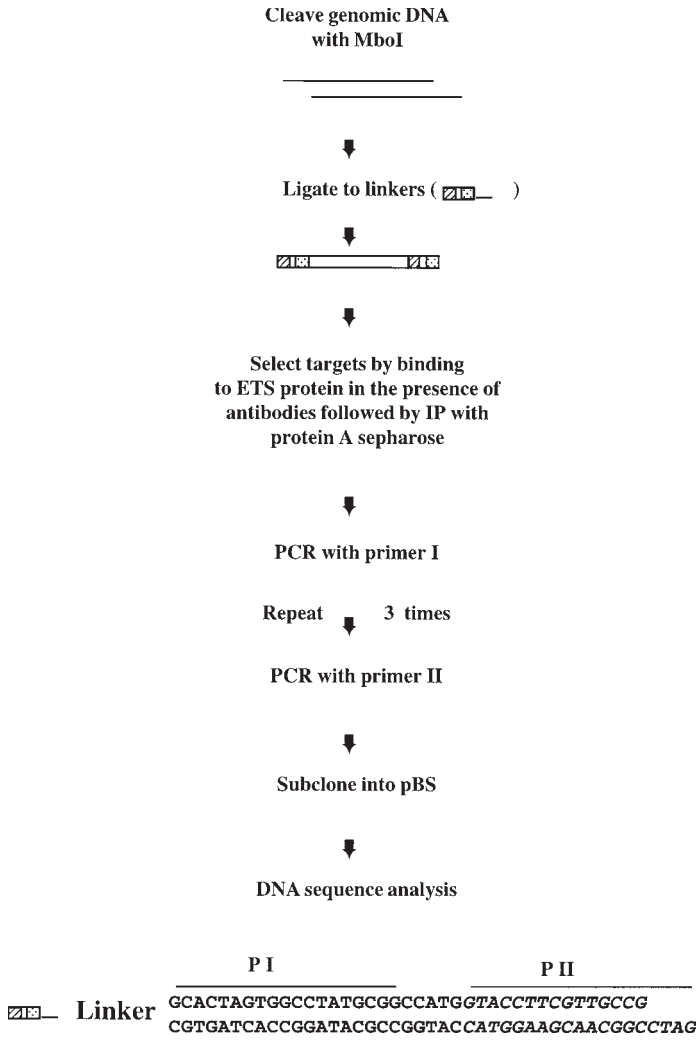


Fig. 1. Diagram of modified WGPCR strategy.

with unphosphorylated linkers to serve as an efficient template for the polymerase chain reaction (PCR). Recombinant Ets1 protein and Ets1 specific monoclonal antibody are incubated with the pool of linker-ligated DNA to allow binding to EBS within genomic DNA fragments. Immunocomplexed Ets1-DNA is recovered by incubation with protein A-sepharose; the Ets1 protein-bound DNA is released and used as template for amplification by PCR. DNA immunoprecipitation and PCR are repeated three times with Primer I; the selected fragments are subjected to a final round of amplification using Primer II, and the resultant bands are fractionated in polyacrylamide gel prior

to cloning into pBS. Of the large number of clones isolated, forty-three clones were examined by DNA sequencing and BLAST analysis; from these, three genomic fragments were found to be derived from the regulatory regions of the human serglycin, preproapolipoprotein *C II* and the *Egr1* genes (**Table 1**) ***(6)***.

We found that the promoter regions of human serglycin, preproapolipoprotein and Egr1 contain consensus Ets binding sites, and are able to bind to Ets proteins by EMSAs. Human serglycin is a proteoglycan that is involved in differentiation of many hematopoietic cells ***(24)***. The Ets binding site in the 5' flanking region (residues –75 to –80) of the human serglycin gene identified by WGPCR is also conserved in the mouse serglycin promoter, suggesting that this site is important for transcriptional regulation. Moreover, elevated expression of the human serglycin gene has been demonstrated in a number of human leukemic cell lines, which have also been shown to express high levels of Ets1 and Fli1 ***(25,26)***. The promoter of the preproapolipoprotein *C-II* gene has an optimal EBS, containing seven residues that are identical to the MSV-LTR; these sequences were originally used to establish Ets1 as a sequence-specific DNA binding protein ***(27)***. Analysis of the Egr1 promoter revealed two SREs (SREI and SREII), each containing CArG box(es) contiguous with Ets binding site(s). Deletion analysis demonstrated that the most 5' EBS and the CArG box of the SREI element are necessary for promoter function, since removal of this SRE resulted in a dramatic loss of promoter activity. The finding that Ets1 binds to and transactivates transcription from the *Egr1*-SREI suggests that *Egr1* is a cellular target of Ets1. Importantly, *Egr1* was also found to be isolated as an Ets1 target gene by RNA differential display cloning, suggesting that it is indeed an Ets transcription factor target gene ***(6)***.

2. Materials

1. Genomic DNA.
2. Restriction and modifying enzymes: Restriction endonuclease MboI, T4 DNA ligase, *Taq* polymerase, and T4 polynucleotide kinase.
3. 5X T4 polynucleotide kinase buffer: 250 m*M* imidazole , pH 6.4, 60 m*M* $MgCl_2$, 5 m*M* 2-mercaptoethanol, 0.35 m*M* ADP.
4. TE: 10 m*M* Tris·Cl , pH 7.4, 1 m*M* EDTA, pH 8.0.
5. 10X T4 DNA ligase buffer: 500 m*M* Tris, 100 mL $MgCl_2$, 10 m*M* DTT, 10 m*M* ATP.
6. Purified recombinant transcription factor protein (*see* **Note 2**).
7. Monoclonal antibody specific to transcription factor protein.
8. Poly-dIdC.
9. Protein A-sepharose.
10. 10X Binding buffer : 200 m*M* Tris pH 7.6, 500 m*M* NaCl, 10 m*M* $MgCl_2$, 2m*M* EDTA, 50% glycerol, 5 m*M* DTT, 0.5 m*M* PMSF.
11. TN buffer: 10 m*M* Tris-HCL pH 7.5,150 m*M* NaCl.

Table 1
ETS Target Genes Identified by Whole Genome PCR

Clone	Strategy	Insert Size (bp)	Sequence Homology	ETS:DNA Binding	RNA Expression NIH3T3	ETS1	ETS2
L510	[a]WG	500	Serglycin	+	nd	nd	nd
L45	[a]WG	500	EGR1	+	–	+	–
L29	[a]WG	500	Preproapolipoprotein CII	+	nd	nd	nd
AE112	[b]DD	240	CArG binding factor (CBF)	nd	–	++	++
AE134	[b]DD	206	PLA2P (rat)	nd	–	–	++
AE117	[b]DD	258	Egr1	+	–	++	–

[a]Whole genome PCR: of 43 clones, three are known and 40 are unknown.
[b]Differential display: from eighty-two cDNA bands, three known and thirteen unknown clones were isolated.
[c]nd – not done.

12. Dissociation buffer: 500 m*M* Tris-HCl pH 9.0, 20 m*M* EDTA, 10 m*M* NaCl, 0.2% SDS.
13. PCR primers: primer I: 5'-GCACTAGTGGCCTATGCGG-3'; primer II: 5'-GTACCTTCGTTGCCGGATC-3'.
14. Oligonucleotide linkers (39/43): 5'-GCACTAGTGGCCTATGCGGCCATG-GTACCTTCGTTGCCG-3' and 5'-GATCCGGCAACGAAGGTACCATGGC-CGCATAGGCCACTAGTGC-3'
15. PCR reaction buffer: 250 µ*M* dNTPs, 10 m*M* Tris-HCl pH 8.3, 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 0.001% w/v gelatin.
16. Radioisotopes: γ-^{32}P ATP (3000 Ci/m*M*) and γ-^{32}P ATP (6000Ci/m*M*).
17. Cloning vector: Dephosphorylated BamHI digested pBluescript SK- plasmid or another suitable plasmid vector.
18. LB/Amp plates: bacto-tryptone: 10 g/L, bacto-yeast extract 5 g/L, NaCl 10 g/L, ampicillin 60 mg/L.
19. X-gal stock solution: 20 mg/mL in dimethylformamide.
20. Isopropylthio-ß-D-galactoside (IPTG) stock solution: 200 mg/mL.
21. 10X Tris-borate (TBE): 108 g Tris base, 55 g boric acid, 40 mL 0.5 *M* EDTA, pH 8.0.
22. 30% acrylamide: 29 g acrylamide, 1 g N,N'-methylenebisacrylamide, H_2O to 100 mL.
23. 4% polyacrylamide gel (100 mL): 13.33 mL 30% acrylamide, 76 mL H_2O, 10 mL 10X TBE, 0.7 mL 10% ammonium persulfate, 35 µL N,N,N',N'-tetramethylethylenediamine (TEMED), pour quickly without bubbles between clean glass plates.
24. 3 *M* sodium acetate, pH 4.8.
25. 100% ethanol.
26. Phenol:chloroform: equal amounts of phenol and chloroform equilibrated with 0.1 *M* Tris·Cl, pH 7.6.
27. Gel fixing solution: 75% H_2O, 15% acetic acid, 10% methanol.
28. Dialysis tubing.
29. DNA sequence analysis software.

3 Methods

3.1. Whole Genome PCR

3.1.1. Preparation of Linker-Ligated Genomic DNA

1. Digest 15 µg of human genomic DNA with MboI to produce DNA fragments with 5' GATC overhangs.
2. Incubate eight units of T4 DNA ligase overnight at 16°C with five µg of purified digested DNA and ten µg of unphosphorylated synthetic double-stranded oligonucleotide linker 39/43 in 1X T4 DNA Ligase buffer (*see* **Note 1**).
3. Separate ligated DNA from free linker by electrophoresis in a 1% agarose gel.
4. Cut out the linker DNA band separately from the smear of linker-ligated genomic DNA.
5. Elute linker-ligated genomic DNA from the gel by electrophoresis in a dialysis tubing bag.

6. Precipitate linker ligated DNA with three volumes of 100% ethanol and 1/10 volume of 3 *M* NaOAc for 1 h on dry ice.
7. Extract with phenol:chloroform and precipitate with three volumes of 100% ethanol and 1/10 volume of 3 *M* NaOAc for one hour on dry ice.
8. Centrifuge, air dry, and resuspend the pellet of linker-ligated DNA in 20 µL sterile water.

3.1.2. Immunoprecipitation of Protein Bound Linker-ligated Genomic DNA Fragments

1. Incubate the following on ice for 20 min in 20 µL binding buffer: 1 µg of linker-ligated DNA, 5 µg recombinant protein, 1 µg of specific monoclonal antibody, 0.5 µg of poly-dIdC to block nonspecific binding (*see* **Notes 2** and **3**).
2. Add 120 µL of protein A-sepharose preincubated with poly-dIdC to block nonspecific binding, and incubate for an additional 20 min, keeping the tube on ice and gently tapping every 5 min to resuspend the beads.
3. Centrifuge 30 s at 14,000*g*, wash once in TN buffer, centrifuge again, and remove supernatant.
4. Resuspend the pellet in 120 µL of dissociation buffer and remove protein A-sepharose beads by microcentrifugation at 14,000*g* for 1 min.
5. Isolate linker-ligated genomic DNA bound to immunoprecipitated protein from the supernatant by phenol:chloroform extraction, followed by ethanol precipitation.
6. Air dry and resuspend immunoprecipitated linker-ligated DNA in 10 µL of sterile water.

3.1.3. PCR Amplification of Linker-ligated Genomic DNA

1. Amplify immunoprecipitated linker-ligated genomic DNA fragments by PCR using 3–5 µ*M* of primer I with 5 U of TAQ polymerase in a final volume of 50 µL using the following parameters: 94°C, 1'; 56°C, 1'; 72°C, 2'; 30 cycles.
2. Repeat **steps 1–6** in **Subheading 3.1.2., step 1** of **Subheading 3.1.3.** three times (*see* **Note 4**).
3. Finally, amplify the transcription factor-bound fragments with a fourth round of PCR using Primer II.

3.1.4. Recovery of Genomic DNA Fragments

1. Digest fourth-round PCR products with BamHI.
2. Label 1 µg of digested DNA in 150 µL of 1X T4 polynucleotide kinase buffer containing 3 U of T4 polynucleotide kinase and 1 m*M* γ-^{32}P ATP (3000 Ci/mmol), for 30 min at 37°C.
3. Separate reaction products by electrophoresis of the labeling mixture in a 4% polyacrylamide gel at 25 V overnight.
4. Visualize products by autoradiography.
5. Cut out bands of interest (approx 0.5 kb) and place in 1.5 mL Eppendorf tubes.
6. Elute DNA with 10:1 TE and ethanol precipitate DNA as described above.

3.1.5. Subcloning of Recovered DNA Fragments

1. Ligate recovered DNA fragments with dephosphorylated BamHI digested pBluescript SK-plasmid in a 2:1 ratio.
2. Transform competent host bacteria.
3. Spread 40 µL of X-gal and 4 µL of IPTG on LB/Amp plates.
4. Plate transformed bacteria and incubate for 18 h at 37°C.
5. Pick white colonies.
6. Extract plasmids with a minipreparation procedure, and determine the DNA sequence of inserts.
7. Identify binding sites in cloned DNA sequences by FASTA analysis (*see* **Notes 6** and **7**).

3.2. Verification of Targets by Electrophoretic Mobility Gel Shift Assay (EMSA)

1. Incubate approximately 5 µg of cell extract with [^{32}P]-labeled oligonucleotide probe (~50,000 cpm) corresponding to the subcloned PCR sequence of interest in the presence or absence of a 20-fold excess of unlabeled competitor DNA for 20 min at 4°C.
2. Electrophorese samples on a 4% acrylamide gel in 0.4X TBE buffer for 1.5 h at 250 V.
3. Soak the gel in fixing solution for 10–15 min, vacuum dry, and autoradiograph overnight.

4. Notes

1. The formation of concatamers during the ligation reaction can be avoided by using unphosphorylated catch linkers.
2. Use monoclonal antibodies and purified recombinant proteins, if possible. We have successfully used nuclear extracts of baculovirus expressed proteins and monoclonal antibodies.
3. Two major variables in WGPCR are the source of the protein and the method of selection of bound complexes. Several reports have demonstrated that it is not necessary to use purified proteins for the successful selection of genomic DNA fragments. For example, nuclear extracts prepared from COS cells transfected with RAR-alpha or RAR-beta expression vectors have been used as the source of protein, resulting in the identification of genomic DNA containing retinoic acid–response elements ***(28)***. Several investigators have used in vitro transcription and translation to obtain sufficient proteins for incubation, including human thyroid-hormone receptor beta, human RXR (a retinoic acid receptor family member) and WT1 (Wilms' tumor) suppressor gene ***(29,30)***. Expression of fusion proteins in bacteria (for example, GST [glutathione-S-transferase]) not only serves as a source of protein, but also can abolish the need for immunoprecipitation (i.e., by GST-pull down) and, thus, eliminate the requirement for specific antibodies. In a study identifying Evi-1 binding sequences, expressed proteins were transferred to nitrocellulose, renatured, and incubated with labeled, linker-ligated genomic DNA. Bound DNA fragments enriched for Evi-1 binding sequences were then

visualized by autoradiography and recovered and amplified. Alternatively, the binding and amplification reactions can be carried out in suspension, incubating the linker-ligated genomic DNA with GST-protein immobilized on GST-sepharose, and after washing, the fusion-protein:DNA complex can be eluted with reduced glutathione ***(31)***.

4. Multiple cycles of binding/release/PCR reduces the amount of nonspecific amplification; we have found that three cycles is a minimum
5. Separate the PCR products by polyacrylamide electrophoresis, elute each band, and clone them separately in order to reveal the complexity of fragment sizes before sequencing.
6. Sequence a short region of each cloned DNA fragment and then use bioinformatic software such as the NCBI BLAST program to identify those that correspond to known promoter/enhancer sequences. Sequence the entire length of clones, which are not present in the databases, and analyze for binding site sequences.
7. After cloning presumptive target gene promoter sequences, a multiplicity of independent approaches can be used to verify that the isolated DNA sequences are derived from the regulatory region of genes that are responsive to the transcription factor. Bioinformatic DNA sequence analysis should reveal the presence of consensus DNA binding site(s) for various factors. The functionality of such elements should be verified by in vitro analysis, including electrophoretic mobility shift assays (EMSA), DNAse I footprinting and transient transfection assays using reporter gene constructs. A "true" target of the original protein will alter the mobility of the binding site DNA fragment in EMSA, will be protected by the protein in DNAse footprinting, and will activate reporter gene transcription in transient transfection assays.

References

1. Burger, A. M., Zhang, X., Papas, T. S., and Seth, A. (1998) Detection of novel genes that are up- (Di12) or down-regulated (T1A12) with disease progression in breast cancer. *European J. Cancer Prevention*, **7(suppl 1),** S29–S35.
2. Burger, A. Zhang X-K, Li, H., Ostrowski J. L., Beatty, B., Venanzoni, M., Papas, T., and Seth, A. (1998a) Down-regulation of T1A12, a novel insulin-like growth factor binding protein related gene, is associated with disease progression in breast carcinomas. *Oncogene*, **16,** 2459–2467.
3. Burger, A., Li, H., Zhang, X-K., Pienkowska, M., Venanzoni, M., Vournakis, J., Papas, T. S. and Seth, A. (1998b) Breast cancer genome anatomy: Correlation of morphological changes in breast carcinomas with expression of the novel gene product Di12. *Oncogene*, **16,** 327–333.
4. Salesiotis, A. N., Wang, C. K., Wang, C. D., Burger, A., Li, H., and Seth, A. (1995) Identification of novel genes from stomach cancer cell lines by differential display. *Cancer Lett.*, **91**, 47–54.
5. Schweinfest, C. W., Graber, M. W., Chapman, J. M., Papas, T. S., Baron, P. L., and Watson, D. K. (1997) CaSm: an Sm-like protein that contributes to the transformed state in cancer cells. *Cancer Res.* **57**, 2961–2965.

6. Robinson, L., Papas, T. S., Kola I., and Seth A. (1997) The ETS target genes: Isolation of novel Ets responsive genes by whole genome PCR and differential display of mRNAs expressed in Ets1 and Ets2 transformed cells. *Proc. Natl. Acad. Sci. USA* **94**, 7170–7175.
7. Kinzler, K. W. and Vogelstein B. (1989) Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. *Nucleic Acids Res*., **17**, 3645–3653.
8. el-Deiry, W. S., Kern, S. E., Pieteupol, J. A., Kinzler, K. W., and Vogelstein, B. (1992) Definition of a consensus binding site for p53. *Nature Genetics* **1**, 45-49.
9. Seth, A., Ascione, R., Fisher, R. J., Mavrothalassitis, G. J., Bhat, N. K., and Papas, T. S. The ets gene family. *Cell Growth Differ* **3**, 327–334.
10. Papas, T. S., Bhat, N. K., Spyropoulos, D. D., Mjaatvedt, A. E., Vournakis, J., Seth, A., and Watson, D. K. (1997) Functional relationships among ETS gene family members. Proc. 18th Sympo. Comparative Res. Leukemia and Related Diseases, *Leukemia* **11** (Supplement 3): 557–566.
11. Karim, F. D., Urness, L. D., Thummel, C. S., Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C., Maki, R. A., Gunther, C. V., Nye, J. A., et al. (1990) The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev*. **4**, 1451–1453.
12. Janknecht, R. and Nordheim, A. (1993) Gene regulation by Ets proteins. *Biochim Biophys Acta* **1155**, 346–356.
13. Kola, I., Brookes, S., Green, A. R., Garber, R., Tymms, M., Papas, T. S., and Seth, A. (1993) The Ets1 transcription factor is widely expressed during murine embryo development and is associated with mesodermal cells involved in morphogenetic processes such as organ formation. *Proc. Natl. Acad. Sci. USA*, **90**, 7588–7592.
14. Maroulakou, I. G., Papas, T. S., and Green, J. E. (1994) Differential expression of ets-1 and ets-2 proto-oncogenes during murine embryogenesis. *Oncogene*, **9**, 1551–1565.
15. Muthusamy, N., Barton, K., and Leiden, J.M. (1995) Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature* **377**, 639–642.
16. Bhat, N. K., Thompson, C. B., Lindsten, T., June, C. H., Fujiwara, S., Koizumi, S., Fisher, R. J., and Papas, T. S. (1990) Reciprocal expression of human ETS1 and ETS2 genes during T-cell activation: regulatory role for the protooncogene ETS1. *Proc. Natl. Acad. Sci. USA* **87**, 3723–3727.
17. Huang C. C, Papas TS, and Bhat N. K. (1997) A variant form of ETS1 induces apoptosis in human colon cancer cells. *Oncogene* **15**, 851–856.
18. Shimizu, K., Ichikawa, H., Tojo, A., Kaneko, Y., Maseki, N., Hayashi, Y., Ohira, M., Asano, S. and Ohki, M. (1993) An ets-related gene, ERG, is rearranged in human myeloid leukemia with t(16;21) chromosomal translocation. *Proc. Natl. Acad. Sci USA* **90**, 10,280–10,284.
19. Golub, T. R., Barker, G. F., Lovett, M. and Gilliland, D. G. (1994) Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5; 12) chromosomal translocation. *Cell* **77**, 307–316.

20. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G. et al. (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162–165.
21. Sorensen, P. H., Lessnick, S. L., Lopez-Terrada, D., Liu, X. F., Triche, T. J., and Denny, D. T. (1994) A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nature Genet.* **6**, 146–141.
22. Sumarsono, S. H., Wilwon, T. J., Tym*m*s, M. J., Venter, D. J., Corrick, C. M., Kola R., Lahoud M. H., Papas, T. S., Seth A., and Kola, I. (1996) Down's syndrome-like skeletal abnormalities in Ets2 transgenic mice. *Nature* **379**, 534–538.
23. Seth, A., Robinson, L., Panayiotakis, A., Thompson, D. M., Hodge, D. R., Zhang, X. K., Watson, D. K., Ozato, K., and Papas, T. S. (1994) The *EndoA* enhancer contains multiple ETS binding site repeats and is regulated by ETS proteins. *Oncogene* **9**, 469–477.
24. Toyama-Sorimachi, N., Sorimachi, H., Tobita, Y., Kitamura, F., Yagita, H., Suzuki, K., and Miyasaka, M. (1995) A novel ligand for CD44 is serglycin, a hematopoietic cell lineage-specific proteoglycan. Possible involvement in lymphoid cell adherence and activation. *J. Biol. Chem*. **270**, 7437–7444.
25. Papas, T. S., Watson, D. K., Sacchi, N., Fujiwara, S., Seth, A. K., Fisher, R. J., Bhat, N. K., Mavrothalassitis, G., Koizumi, S., Jorcyk, C. L., Schweinfest, C. W., Kottaridis, S. D., and Ascione, R. (1990) ETS family of genes in leukemia and Down syndrome. *Am. J. Med. Gene*t. **7**, 251–261.
26. Watson, D. K., Smyth, F. E., Thompson, D. M., Chen, J., Testa, J. R., Papas, T. S., and Seth, A. (1992) The ERGB gene: isolation and characterization of a new member of the family of human ETS transcription factors. *Cell Growth Differ.* **3**, 705–713.
27. Gunther, C. V., Nye, J., Bryner, R., and Greaves, B. (1990) Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes and Dev*. **4**, 667–679.
28. Costa-Giomi, M. P., Gaub, M. P., Chambon, P., Abarzua, P. (1992) Characterization of a retinoic acid responsive element isolated by whole genome PCR. *Nucleic Acids Res.* **20**, 3223–3232.
29. Caubin, J., Iglesias, T., Bernal, J., Munoz, A., Marquez, G., Barbero, J. L., Zaballos, A. (1994) Isolation of genomic DNA fragments corresponding to genes modulated in vivo by a transcription factor. *Nucleic Acids Res*. **22**, 4132–4138.
30 Nakagama, H., Heinrich, G., Pelletier, J., Housman, D. E. (1995) Sequence and structural requirements for high-affinity DNA binding by the WT1 gene product. *Mol Cell Biol*. **15**, 1489–1498.
31. Matsugi, T., Kreider, B. L., Delwel, R., Cleveland, J. L., Askew, D. S., Ihle, J. N. (1995) The Evi-1 zinc finger myeloid transforming protein binds to genomic fragments containing (GATA)n sequences. *Oncogene* **11**, 191–198.

2

In Vivo Footprinting Using UV Light and Ligation-Mediated PCR

Gerd P. Pfeifer and Stella Tommasi

1. Introduction

The analysis of chromatin structure at single-nucleotide resolution (genomic footprinting) has long been considered technically difficult, at least in mammalian cells. Recently, techniques have been developed that give a sufficient specificity and sensitivity to analyze single-copy genes by genomic footprinting ***(1)***. The most sensitive method uses ligation-mediated polymerase chain reaction (LMPCR) to amplify all fragments of a genomic sequence ladder ***(2,3)***. LMPCR is based on the ligation of an oligonucleotide linker onto the 5' end of each DNA molecule that was created by a strand cleavage reaction during the footprinting procedure. This ligation reaction provides a common sequence on all 5' ends allowing exponential PCR to be used for signal amplification. Thus, by taking advantage of the specificity and sensitivity of PCR, one needs only a microgram of mammalian DNA per lane to obtain good quality DNA sequence ladders, with retention of all information relating to DNA methylation, DNA structure, and protein footprints. The general LMPCR procedure is outlined in **Fig. 1**. The first step of the procedure is cleavage of DNA, generating molecules with a 5'-phosphate group. This is achieved, for example, by chemical DNA sequencing (ß-elimination), by cutting with an enzyme such as DNaseI, or by converting ultraviolet (UV) photolesions into strand breaks. Next, primer extension of a gene-specific oligonucleotide (primer 1) generates molecules that have a blunt end on one side. Linkers are ligated to these blunt ends, and then an exponential PCR amplification of the linker-ligated fragments is done using the longer oligonucleotide of the linker (linker-primer) and a second gene-specific primer (primer 2). After 18–20 PCR amplification cycles, the DNA fragments are separated on a sequencing gel, electroblotted onto nylon

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

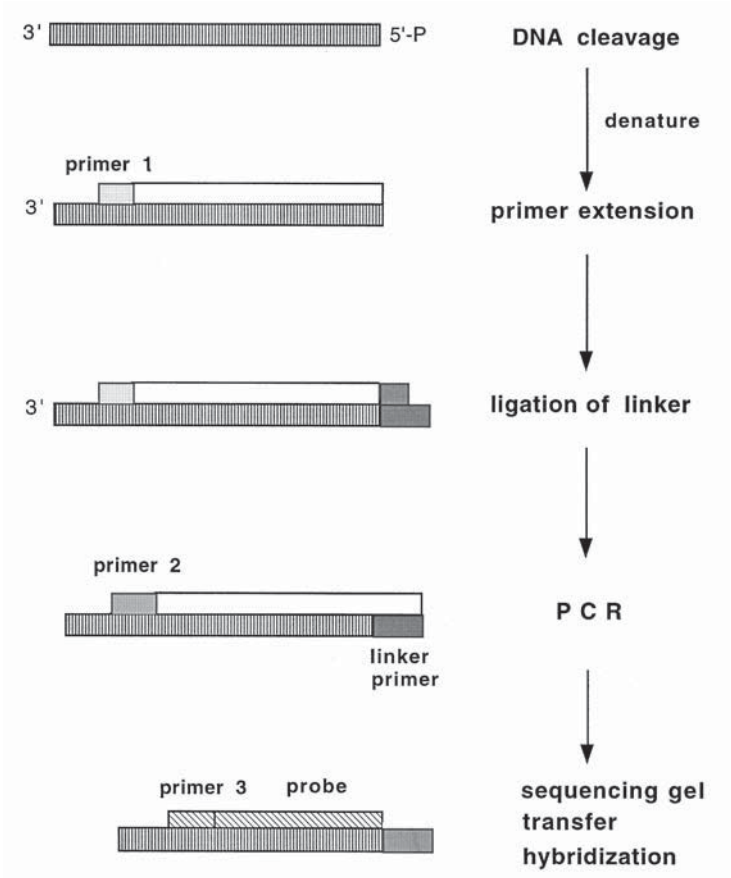


Fig. 1. Outline of the ligation-mediated PCR procedure. The steps include cleavage and denaturation of genomic DNA, annealing and extension of primer 1, ligation of the linker, PCR amplification of gene-specific fragments with primer 2 and the linker-primer, detection of the sequence ladder by gel electrophoresis, electroblotting, and hybridization with a single-stranded probe that does not overlap primers 1 and 2.

membranes, and hybridized with a gene-specific probe to visualize the sequence ladders ***(3)***.

LMPCR is generally suitable for detection of any DNA strand breaks that provide ligatable ends, either directly or indirectly. The method has been used for chemical sequencing of genomic DNA and for determination of DNA cytosine methylation patterns ***(3,4)***. To obtain information about protein binding, in vivo footprinting experiments can be done on intact cells using dimethylsulfate (DMS), a small molecule that penetrates the cell membrane ***(5)***. DMS does not, however, reveal all protein–DNA contacts. By reasoning that more

bulky agents, such as enzymes would give better information on chromatin structure, DNaseI footprinting was adopted for use with LMPCR and very informative DNaseI footprints can be obtained using cells that have been permeabilized with lysolecithin ***(6–8)***. DNaseI footprinting is technically somewhat more difficult and requires cell permeabilization.

LMPCR provides adequate sensitivity to map certain DNA lesions, like those formed after UV irradiation ***(9,10)***. Thus, LMPCR can be used in conjunction with footprinting methods that use ultraviolet (UV) light as a probing agent ***(9–13)***. UV irradiation is clearly one of the less disruptive methods to detect DNA bound proteins in vivo, where intact cells are irradiated for a short period of time. The primary target of 254 nm UV irradiation is DNA, making perturbation of other cellular processes and secondary events that could lead to disturbance of factor binding unlikely. Thus, the results obtained from UV footprinting most probably reflect the true in vivo situation.

The most frequent UV photoproducts formed in DNA after irradiation with 254 nm UV light are cyclobutane pyrimidine dimers (CPDs), which arise by photodimerization and are characterized by two covalent bonds between adjacent pyrimidines, and pyrimidine (6-4) pyrimidone photoproducts [(6-4) photoproducts], which arise through a complex rearrangement mechanism and involve a single covalent bond between positions 6 and 4 of two adjacent pyrimidines (**Fig. 2**). The (6-4) photoproducts are formed at a frequency of approximately 20–30% of that of cyclobutane dimers ***(14)*** but the relative abundance of the two photoproducts depends on the local DNA sequence context. CPDs are found at all dipyrimidines and are most common at 5'-TpT sites, whereas (6-4) photoproducts are most frequently detected at 5'-TpC sequences ***(15)***. Because of the specificity of UV photoproduct formation, UV footprinting will be informative only at sequences that contain dipyrimidines. However, a systematic analysis of known factor binding sites indicates that at least one of the two complementary strands of a transcription factor binding site will almost always contain a dipyrimidine sequence ***(16)***.

Becker and Wang initially introduced the use of ultraviolet radiation as a footprinting agent ***(17)***. We have recently used LMPCR for UV footprinting in mammalian cells after conversion of the UV-induced lesions into ligatable DNA strand breaks ***(9,10)***. The (6-4) photoproducts are converted into DNA strand breaks with 5' phosphate groups by heating in piperidine (**Fig. 3**). Cyclobutane dimers are not cleaved under these conditions, but can be mapped at the DNA sequence level by use of specific endonucleases such as T4 endonuclease V (**Fig. 3**). This enzyme does not recognize (6-4) photoproducts, and, thus, the two types of photoproducts can be mapped separately. T4 endonuclease V cleaves the glycosidic bond of the 5' base within a pyrimidine dimer and also cleaves the sugar phosphate backbone between the two dimerized

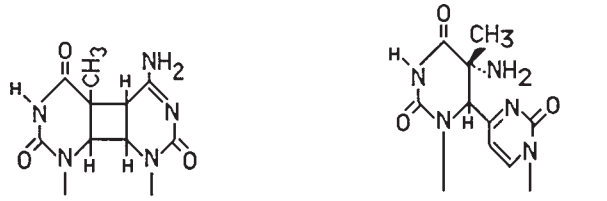


Fig. 2. Structure of a cyclobutane pyrimidine dimer and a (6-4) photoproduct at a 5'-TpC dipyrimidine sequence.

pyrimidines through its AP lyase activity. We found that the resulting fragments could be amplified efficiently by LMPCR only after the cyclobutane ring of the dimerized base was reverted with *Escherichia coli* photolyase to result in a normal base on a 5'-terminal sugar-phosphate ***(10)***.

When promoter regions of several genes were analyzed by UV irradiation and the photoproduct spectrum of irradiated purified DNA was compared with the photoproduct spectrum after irradiation of cells, some striking differences were observed ***(10,11)***. At sequences that contain binding sites of transcription factors, photoproduct formation can either be suppressed or enhanced. Some strong (up to 30-fold) enhancements of photoproduct formation can be observed at specific dipyrimidines within certain transcription factor binding sites such as the CCAAT box in several genes and the serum response factor binding site in the human *FOS* gene ***(11)***. The mechanisms leading to the formation of photofootprints are not precisely known. Most likely, structural alterations in the DNA double helix induced upon factor binding, such as DNA unwinding or bending, will favor or disfavor the formation of UV photoproducts at specific sequences ***(11)***.

Ultraviolet light has the potential to detect a wide variety of sequence-specific protein–DNA contacts. The distribution of CPDs and (6-4) photoproducts has been examined along the promoter sequences of several mammalian genes, including *PGK1*, *JUN*, *PCNA*, *FOS*, and *Xist* ***(10–13)***. A comparison of the UV photofootprinting data obtained from these studies with data from experiments using other probing strategies to analyze the same genes showed that UV light has the potential to reveal all protein–DNA interactions provided that there is a dipyrimidine sequence on either DNA strand within a factor binding site and that both photoproducts are analyzed. **Fig. 4** shows an example

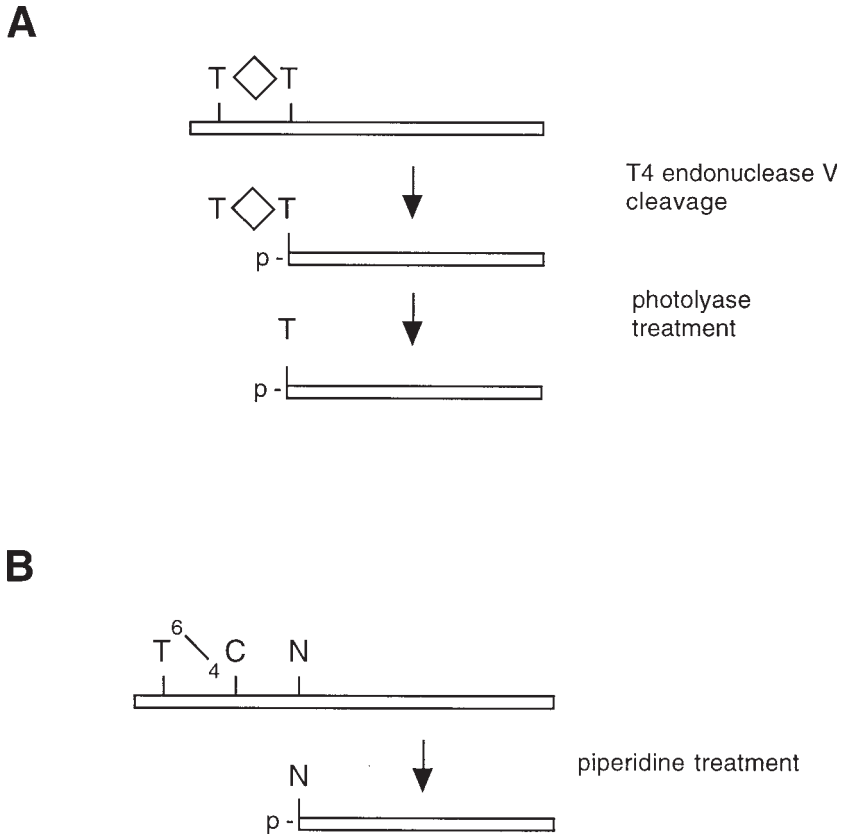


Fig. 3. Detection of cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts in mammalian genes. **(A)** CPDs are converted into DNA strand breaks with a 5' phosphate group by cleavage with T4 endonuclease V and by photolyase treatment to create ligatable ends. The resulting DNA break positions can be detected by ligation-mediated PCR. **(B)** Detection of (6-4) photoproducts in mammalian genes. The (6-4) photoproducts are converted into DNA strand breaks with 5' phosphate groups by alkaline cleavage. Break positions are then detected by ligation-mediated PCR. Note that the amplification product derived from a (6-4) photoproduct is one nucleotide shorter than the product derived from a CPD at the same dipyrimidine sequence. Only one strand of the DNA duplex is shown.

of a UV footprinting experiment in which the promoter of the human thymidine kinase gene was investigated. The ability to detect a large variety of different factors should make UV footprinting a generally useful method for in vivo studies of protein–DNA interactions at promoters.

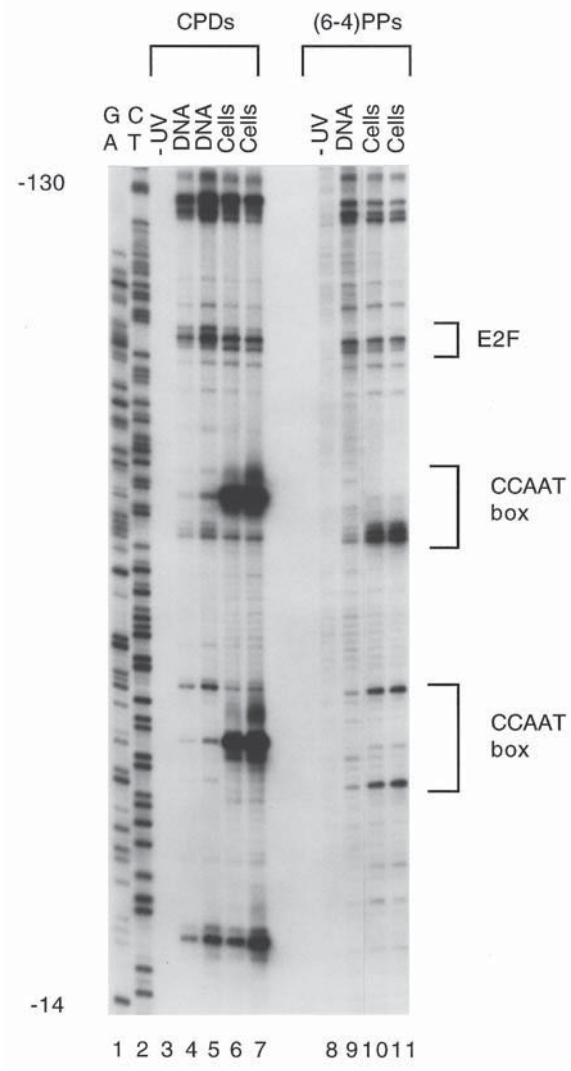


Fig. 4. An example of UV footprinting obtained by ligation-mediated PCR. The region analyzed contains sequences of the promoter of the human thymidine kinase gene. Footprints are seen at an E2F site and at the two inverted CCAAT boxes (indicated by brackets) in irradiated human fibroblasts (Cells). Purified, UV-irradiated genomic DNA is shown in the control lanes (DNA). Cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts [(6-4)PPs] were analyzed separately. Samples were UV-irradiated at a dose of 1000 J/m^2 (lanes 4, 6, 9, and 10) or 2000 J/m^2 (lanes 5, 7, and 11).

2. Materials

1. UV light source. Light sources emitting 254 nm light are available in most laboratories (germicidal lamps; inverted transilluminators from which the lids have been removed; commercially available devices that crosslink DNA to nylon membranes).
2. UVX radiometer (Ultraviolet Products, San Gabriel, CA).
3. Buffer A: 0.3 *M* sucrose, 60 m*M* potassium chloride, 15 m*M* sodium chloride, 60 m*M* Tris-HCl, pH 8.0, 0.5 m*M* spermidine, 0.15 m*M* spermine, 2 m*M* EDTA.
4. Nonidet P40.
5. Buffer B: 150 m*M* NaCl, 5 m*M* EDTA, pH 8.0.
6. Buffer C: 20 m*M* Tris-HCl, pH 8.0, 20 m*M* NaCl 20 m*M* EDTA, 1% sodium dodecyl sulfate.
7. Proteinase K.
8. DNase-free RNAase A.
9. Phenol: Equilibrate with 0.1 *M* Tris-Cl, pH 8.
10. Chloroform.
11. Ethanol.
12. 3 *M* sodium acetate, pH 5.2.
13. TE buffer: 10 m*M* Tris-HCl, pH 7.6, 1 m*M* EDTA.
14. 1 *M* Piperidine (Fluka): Prepare fresh.
15. 10X T4 endonuclease V buffer: 500 m*M* Tris-HCl, pH 7.6, 500 m*M* NaCl, 10 m*M* EDTA, 10 m*M* dithiothreitol, 1 mg/mL bovine serum albumin.
16. T4 endonuclease V: This enzyme was kindly provided by Dr. R.S. Lloyd, Vanderbilt University; it is also commercially available from Epicentre Technologies (Madison, WI), or from Texagen (Plano, TX).
17. *E. coli* photolyase: This enzyme was kindly provided by Dr. Aziz Sancar (University of North Carolina at Chapel Hill). It is commercially available from Pharmingen, San Diego, CA.
18. Two 360 nm black lights (Sylvania 15W F15T8).
19. Agarose.
20. 50 m*M* NaCl, 4 m*M* EDTA.
21. Running buffer: 30 m*M* NaOH, 2 m*M* EDTA.
22. Loading dye: 50% glycerol, 1 *M* NaOH, 0.05% bromocresol green.
23. 0.1 *M* Tris-Cl, pH 7.5.
24. 1 µg/mL Ethidium bromide.
25. Oligonucleotide primers for primer extension: The primers used as primer 1 (Sequenase primers) are 15- to 20-mers with a calculated *Tm* of 48 to 56°C (*see* **Note 1**). Prepare primers as stock solutions of 50 pmol/µL in water or TE buffer and keep at –20°C.
26. 5X Sequenase buffer: 250 m*M* NaCl, 200 m*M* Tris-Cl, pH 7.7.
27. Mg-DTT-dNTP mix: 20 m*M* $MgCl_2$, 20 m*M* DTT, 0.25 m*M* of each dNTP.
28. Sequenase 2.0 (USB): 13 U/µL.
29. 300 m*M* Tris-HCl, pH 7.7.
30. 2 *M* Tris-HCl, pH 7.7.

31. Linker: Prepare double-stranded linker by annealing a 25-mer (5'-GCGGTGACCCGGGAGATCTGAATTC, 20 pmol/µL) to an 11-mer (5'-GAATTCAGATC, 20 pmol/µL) in 250 m*M* Tris-Cl, pH 7.7, by heating to 95°C for 3 min and gradually cooling to 4°C over a time period of 3 h. Linkers can be stored at –20°C for at least 3 mo. They are thawed and kept on ice.
32. Ligation mix: 13.33 m*M* $MgCl_2$, 30 m*M* DTT, 1.66 m*M* ATP, 83 µg/mL BSA, 3 U/reaction T4 DNA ligase (Promega), and 100 pmol linker/reaction (= 5 µL linker).
33. *E. coli* tRNA.
34. 2X Taq polymerase mix: 20 m*M* Tris-HCl, pH 8.9 , 80 m*M* NaCl, 0.02 % gelatin, 4 m*M* $MgCl_2$, and dNTPs at 0.4 m*M* each.
35. Oligonucleotide primers for PCR: The primers used in the amplification step (primer 2) are 20- to 30-mers with a calculated *Tm* between 60 and 68°C (*see* **Note 2**). Ten picomoles of the gene-specific primer (primer 2) and 10 pmol of the 25-mer linker-primer (5'-GCGGTGACCCGGGAGATCTGAATTC) are used per reaction along with 3 U Taq polymerase, and these components can be included in the 2X Taq polymerase mix.
36. Taq polymerase.
37. Mineral oil.
38. 400 m*M* EDTA, pH 7.7.
39. Formamide loading buffer: 94% formamide, 2 m*M* EDTA, pH 7.7, 0.05% xylene cyanol, 0.05% bromophenol blue.
40. 1 *M* TBE: 1 *M* Tris, 0.83 *M* boric acid, 10 m*M* EDTA, pH 8.3.
41. Whatman 3MM and Whatman 17 paper (Maidstone, UK).
42. Gene Screen nylon membranes (New England Nuclear, Boston, MA).
43. Electroblotting apparatus (Owl Scientific, Cambridge, MA).
44. An appropriate plasmid or PCR product containing the sequences of interest.
45. Oligonucleotide primer to make the hybridization probe. This primer is used together with the cloned template and Taq polymerase to make single-stranded hybridization probes (*see* **Note 3**).
46. ^{32}P-dCTP(3000 Ci/mmol).
47. 7.5 *M* ammonium acetate.
48. Hybridization buffer: 0.25 *M* sodium phosphate, pH 7.2, 1 m*M* EDTA, 7 % SDS, 1% BSA.
49. Washing buffer: 20 m*M* sodium phosphate, pH 7.2 , 1 m*M* EDTA, 1 % SDS.
50. Kodak XAR-5 film.

3. Methods

3.1. Irradiation of Cells and DNA

A 254-nm UV light source is required. A transilluminator used for cross-linking DNA to nylon membranes from which the upper lid and filter have been removed or any commercially available UV crosslinker can be used. It is also possible to use a UVB irradiation source (*see* **Note 4**). UV doses are measured with a UV radiometer.

3.1.1. Cells

1. Prepare approximately 5×10^6 to 2×10^7 cells for a typical experiment. Wash cells that grow as monolayers in Petri dishes with PBS. Wash suspension culture cells in PBS and resuspend.
2. Irradiate cells with a UV dose typically between 500 to 2000 J/m^2 of 254-nm light (*see* **Note 5**).

3.1.2. In Vitro Treatment

1. Prepare purified genomic DNA from the same cell type as used for in vivo crosslinking.
2. Pipet small 5 µL droplets of DNA at a concentration of 0.2–0.5 µg/mL in water or in TE buffer onto Parafilm.
3. Irradiate DNA. Because of a shielding effect, it is usually necessary to use a UV dose that is twice as high for purified DNA as for cells in order to achieve the same frequency of UV damage.

3.2. DNA Isolation

1. Lyse the cells after UV irradiation by adding 10 mL of buffer A containing 0.5% Nonidet P40. This step will release nuclei and remove most of the cytoplasmic RNA. Transfer the suspension to a 50 mL tube. Incubate on ice for 5 min.
2. Centrifuge at 1000*g* for 5 min at 4°C.
3. Wash the nuclear pellet once with 15 mL of buffer A.
4. Resuspend nuclei thoroughly in 2–5 mL of buffer B, add one volume of buffer C, containing 600 µg/mL of proteinase K (added just before use). Incubate for 2 h at 37°C.
5. Add DNase-free RNAase A to a final concentration of 100 µg/mL. Incubate for 1 h at 37°C (*see* **Note 6**).
6. Extract with one volume of buffer-saturated phenol. Then, extract with 0.5 vol of phenol and 0.5 vol of chloroform. Repeat this step until the aqueous phase is clear and no interface remains. Finally, extract with 1 vol of chloroform.
7. Add 0.1 vol of 3 *M* sodium acetate, pH 5.2, and precipitate the DNA with 2.5 vol of ethanol at room temperature.
8. Centrifuge at 2000*g* for 1 min (*see* **Note 7**). Wash the pellet with 75% ethanol and airdry briefly.
9. Dissolve the DNA in TE buffer to a concentration of approximately 0.2 µg/µL. Keep at 4 °C overnight. The DNA should be dissolved well before T4 endonuclease V cleavage.

3.3. Cleavage of DNA at Sites of UV Photodamage

3.3.1 (6-4) Photoproducts

To obtain DNA fragments with a 5' phosphate group at the positions of (6-4) photoproducts, DNA is heated in 1 *M* piperidine. This will destroy the

photolesion and create strand breaks, while the sugar residue at the 3'-base of the (6-4) photoproduct is destroyed.

1. Dissolve 10–50 µg of UV-irradiated DNA in 100 µL of 1 *M* piperidine.
2. Heat the DNA at 90°C for 30 min in a heat block (use lid locks to prevent tubes from popping). Cool samples briefly on ice after heating.
3. Add 10 µL of 3 *M* sodium acetate, pH 5.2, and 2.5 vol of ethanol. Put on dry ice for 20 min.
4. Spin at 14,000*g* in an Eppendorf centrifuge for 15 min.
5. Wash twice with 1 mL of 75% ethanol.
6. Remove traces of remaining piperidine by drying the sample overnight in a vacuum concentrator. Dissolve DNA in TE buffer to a concentration of approx 0.5–1 µg/µL.
7. Determine the frequency of (6-4) photoproducts by separating 1 µg of the DNA on a 1.5% alkaline agarose gel (3.4).

3.3.2. Cyclobutane Pyrimidine Dimers

DNA is first incubated with T4 endonuclease V and then with *E. coli* photolyase (*see* **Fig. 3**) to create fragments with 5'-phosphate groups and ligatable ends.

1. Mix UV-irradiated DNA (≈10 µg in 50 µL) with 10 µL of 10X T4 endonuclease V buffer and a saturating amount of T4 endonuclease V in a final volume of 100 µL. Saturating amounts of T4 endonuclease V activity can be determined by incubating UV-irradiated (2000 J/m^2) genomic DNA with various enzyme dilutions and separating the cleavage products on alkaline agarose gels (*see* **Subheading 3.4.**). Add dithiotreitol to a final concentration of 10 m*M*. Incubate at 37°C for 1 h.
2. Add 5 µg of *E. coli* photolyase under yellow light.
3. Irradiate the samples in 1.5 mL tubes from two 360-nm black lights (Sylvania 15 W F15T8) for 1 h at room temperature at a distance of 3 cm.
4. Extract once with phenol–chloroform and once with chloroform.
5. Precipitate the DNA by adding 0.1 vol of 3 *M* sodium acetate, pH 5.2, and 2.5 vol of ethanol. Leave on dry ice for 20 min. Centrifuge samples for 15 min at 14,000*g* at 4°C.
6. Wash pellets with 1 mL of 75% ethanol and air-dry.
7. Dissolve DNA in TE buffer to a concentration of about 0.5 to 1 µg/µL.
8. Determine the frequency of CPDs by running 1-2 µg of the samples on a 1.5% alkaline agarose gel.

3.4. Estimation of Cleavage Frequency by Alkaline Agarose Gels

The size of the fragments obtained after cleavage of UV-irradiated DNA is determined on an alkaline 1.5% agarose gel (*see* **Note 8**).

1. Prepare a 1.5 % alkaline agarose gel by suspending agarose in 50 m*M* NaCl, 4 m*M* EDTA and microwaving.
2. After the gel solidifies, soak it in running buffer for at least 2 h.
3. Dilute the DNA sample with 1 vol of loading dye. Incubate for 15 min at room temperature.
4. Load the samples and run the gel at 40 V for 3–4 h.
5. Neutralize the gel by soaking in 500 mL of 1 *M* Tris-Cl, pH 7.6, 1.5 *M* NaCl.
6. Stain with ethidium bromide (1 µg/mL).
7. Destain in water.

3.5. Ligation-Mediated PCR

1. Mix in a siliconized 1.5 mL tube: 0.5–2 µg of cleaved DNA, 0.6 pmol of primer 1, and 3 µL of 5X Sequenase buffer in a final volume of 15 µL.
2. Incubate at 95°C for 3 min, then at 45°C for 30 min.
3. Cool on ice; spin 5 s.
4. Add 7.5 µL cold, freshly prepared Mg-DTT-dNTP mix.
5. Add 1.5 µL Sequenase, diluted 1:4 in cold 10 m*M* Tris, pH 7.7.
6. Incubate at 48°C, 15 min, then cool on ice.
7. Add 6 µL 300 m*M* Tris-Cl, pH 7.7.
8. Incubate at 67°C, 15 min (heat inactivation of Sequenase).
9. Cool on ice; spin 5 s.
10. Add 45 µL of freshly prepared ligation mix.
11. Incubate overnight at 18°C.
12. Incubate 10 min at 70°C (heat inactivation of ligase).
13. Add 8.4 µL 3 *M* sodium acetate, pH 5.2, 10 µg *E. coli* tRNA, and 220 µL ethanol.
14. Put samples on dry ice for 20 min.
15. Centrifuge 15 min at 4°C in an Eppendorf centrifuge.
16. Wash pellets with 950 µL 75% ethanol.
17. Remove ethanol residues in a SpeedVac.
18. Dissolve pellets in 50 µL H_20 and transfer to 0.5 mL siliconized tubes.
19. Add 50 µL freshly prepared 2x Taq polymerase mix containing the primers and the enzyme and mix by pipeting.
20. Cover samples with 50 µL mineral oil and spin briefly.
21. Cycle 18–20 times at 95°C, 1 min, 60–66°C, 2 min, and 76°C, 3 min.
22. Add 1 unit of fresh Taq polymerase per sample together with 10 µL reaction buffer. Incubate 10 min at 74°C. This step is to extend completely all DNA fragments and add an extra nucleotide through Taq polymerase's terminal transferase activity (*see* **Note 9**).
23. Add sodium acetate to 300 m*M* and EDTA to 10 m*M* to stop reactions and 10 µg tRNA.
24. Extract with 70 µL of phenol and 120 µL chloroform (premixed).
25. Add 2.5 vol of ethanol and put on dry ice for 20 min.
26. Centrifuge samples 15 min in an Eppendorf centrifuge at 4°C.
27. Wash pellets in 1 mL 75% ethanol.
28. Dry pellets in SpeedVac.

3.6. Preparation of Probe

To prepare labeled single-stranded probes 200 to 300 nts. in length, use repeated primer extension by Taq polymerase with a single primer (primer 3) on a double-stranded template DNA ***(19)***. This can be either plasmid DNA restriction-cut approx 200–300 nts 3' to the binding site of primer 3 or a PCR product made from genomic DNA that represents the target area of interest.

1. Mix 50 ng of the restriction-cut plasmid DNA (or 10 ng of the gel-purified PCR product) with primer 3 (20 pmol), 100 µCi of [^{32}P]dCTP, 10 µM of the other three dNTPs, 10 m*M* Tris-Cl, pH 8.9, 40 m*M* NaCl, 0.01% gelatin, 2 m*M* $MgCl_2$, and 3 U of Taq polymerase in a volume of 100 µL.
2. Perform 35 cycles at 95°C (1 min), 60–66°C (1 min), and 75°C (2 min).
3. Recover the probe by phenol/chloroform extraction, addition of ammonium acetate to a concentration of 0.7 M, ethanol precipitation at room temperature, and centrifugation.

3.7. Sequencing Gel Analysis of Reaction Products

1. Prepare an 8% polyacrylamide gel containing 7 *M* urea and 0.1 *M* TBE, 0.4 mm thick and 60 cm long. To allow identification of the sequence position of all UV-specific bands, include Maxam–Gilbert sequencing standards prepared from genomic DNA as previously described ***(18)***.
2. Dissolve pellets in 1.5 µL of water and add 3 µL formamide loading buffer.
3. Heat samples to 95°C for 2 min prior to loading.
4. Load only one half of the sample or less using a very thin flat tip.
5. Run the gel until the xylene cyanol marker reaches the bottom. Fragments below the xylene cyanol dye do not hybridize significantly.
6. After the run, transfer the gel (i.e., the bottom 40 cm of it) to Whatman 3MM paper and cover with Saran Wrap.
7. Electroblotting of the gel piece can be performed with a simple homemade apparatus ***(18)*** or with a transfer box available from Owl Scientific (*see* **Note 10**). The electroblotting procedure is performed at a current of 1.6 A. After 30 min, the nylon membrane is removed and the DNA side is marked. A high ampere power supply is required for the transfer.
8. After electroblotting, dry the membrane briefly at room temperature, then crosslink the DNA by UV irradiation. UV irradiation is performed in a commercially available crosslinker or by mounting six 254-nm germicidal UV tubes (15 W) into an inverted transilluminator from which the upper lid has been removed. With this device, the distance between membrane and UV bulbs is 20 cm; the UV irradiation time is 30 s.
9. Soak the nylon membranes in 50 m*M* TBE and roll them into the cylinders and transfer to 250 mL plastic or glass hybridization oven cylinders so that the membranes stick completely to the walls of the cylinders without air pockets.
10. Prehybridize the membrane with 15 mL hybridization buffer for 10 min at 62°C.

11. Dilute the labeled probe into 5 mL hybridization buffer and hybridize for 18 h at 62°C.
12. Following hybridization, wash each nylon membrane with 2 L of washing buffer at 60°C. Perform several washing steps in a dish at room temperature with prewarmed buffer. After washing, dry the membranes briefly at room temperature, wrap in Saran wrap and expose to Kodak XAR-5 films. If the procedure has been done without error, a result can be seen after 0.5–8 h of exposure with intensifying screens at –80°C. Nylon membranes can be used for rehybridization if several sets of primers have been included in the primer extension and amplification reactions ***(3)***. Probes can be stripped from the nylon membranes by soaking them in 0.2 *M* NaOH for 30 min at 45°C.

4. Notes

1. Calculation of the *Tm* is done with a computer program ***(20)***. Primers do not need to be gel-purified, if the oligonucleotide synthesis quality is sufficiently good (less than 5% of *n*-1 material on analytical polyacrylamide gels). If a specific target area is to be analyzed (e.g., the binding site of a known transcription factor), primer 1 should be located approximately 100 nts upstream of this target.
2. Primer 2 is designed to extend 3' to primer 1. Primer 2 can overlap several bases with primer 1, but we have also had good results with a second primer that overlapped only one or two bases with the first.
3. The primer that is used to make the single-stranded probe (primer 3) should be on the same strand just 3' to the amplification primer (primer 2) and should have a *Tm* of 60 to 68°C. It should not overlap more than 8–10 bases with primer 2.
4. UV-B light emitted from sunlamps sufficiently penetrates the plastic material of Petri dishes and can be administered from the bottom of the dish without the need to remove the cell culture medium before irradiation. When UVB is used, a dose of at least 20,000–80,000 J/m^2 as measured with a 310-nm sensor is required to produce equivalent amounts of DNA lesions compared to 500 to 2000 J/m^2 of UVC.
5. At these UV doses, the average photoproduct frequencies are one cyclobutane dimer every 200–300 nucleotides and one (6-4) photoproduct every 500–1500 nucleotides.
6. Although RNase is probably active only for a very short time in the proteinase *K* solution, this step seems to aid in removal of traces of RNA.
7. If the initial number of cells was very low, the DNA may need to be pelleted at 10,000*g* for 10 min.
8. The average fragment size, as determined from the smear on the alkaline agarose gels, is not particularly critical but should be somewhere between 200 and 2000 nts. However, it is important that the average fragments sizes are similar for samples that need to be compared directly, e.g., the in vitro and in vivo UV-irradiated samples. The approximate amount of DNA used in the LMPCR reactions can be estimated from the relative amount of DNA visible on these gels. This estimation is important and allows one to obtain similar band intensities on

the sequencing gel in all lanes without having to rerun the sequencing gel to achieve equal loading.

9. If this step is omitted, double bands may occur.
10. The advantages of the hybridization approach over the endlabeling technique ***(2)*** have been discussed previously ***(18)***.

Acknowledgments

This work of G. P. Pfeifer was supported by National Institute of Environmental Health Sciences under Grant ES06070.

References

1. Saluz, H. P. and Wiebauer, K. (eds.) (1995) *DNA and Nucleoprotein Structure In Vivo,* R.G. Landes Co., Austin, TX.
2. Mueller, P. R. and Wold, B. (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246,** 780–786.
3. Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B., and Riggs, A. D. (1989) Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246,** 810–813.
4. Pfeifer, G. P., Tanguay, R. L., Steigerwald, S. D., and Riggs, A. D. (1990) In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.* **4,** 1277–1287.
5. Ephrussi, A., Church, G. M., Tonegawa, S., and Gilbert, W. (1985) B lineage specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* **227,** 134–140.
6. Pfeifer, G. P. and Riggs, A. D. (1991) Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNaseI and ligation-mediated PCR. *Genes Dev.* **5,** 1102–1113.
7. Rozek, D. and Pfeifer, G. P. (1993) In vivo protein-DNA interactions at the c-jun promoter: preformed complexes mediate the UV response. *Mol. Cell. Biol.* **13,** 5490–5499.
8. Chin, P. L., Momand, J., and Pfeifer, G. P. (1997) In vivo evidence for binding of p53 to consensus binding sites in the *p21* and *GADD45* genes in response to ionizing radiation. *Oncogene* **15,** 87–99.
9. Pfeifer, G. P., Drouin, R., Riggs, A. D., and Holmquist, G. P. (1991) In vivo mapping of a DNA adduct at nucleotide resolution: detection of pyrimidine (6-4) pyrimidone photoproducts by ligation-mediated polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **88,** 1374–1378.
10. Pfeifer, G. P., Drouin, R., Riggs, A. D., and Holmquist, G. P. (1992) Binding of transcription factors creates hot spots for UV photoproducts in vivo. *Mol. Cell. Biol.* **12,** 1798–1804.
11. Tornaletti, S. and Pfeifer, G. P. (1995) UV-light as a footprinting agent: modulation of UV-induced DNA damage by transcription factors bound at the promoters of three human genes. *J. Mol. Biol.* **249,** 714–728.

12. Pfeifer, G. P. and Tornaletti, S. (1997) Footprinting with UV irradiation and LMPCR. *Methods* **11,** 189–196.
13. Komura, J.-I., Sheardown, S. A., Brockdorff, N., Singer-Sam, J., and Riggs, A. D. (1997) In vivo ultraviolet and dimethyl sulfate footprinting of the 5' region of the expressed and silent Xist alleles. *J. Biol. Chem.* **272,** 10,975–10,980.
14. Mitchell, D. L., Brash, D. E., and Nairn, R. S. (1990) Rapid repair kinetics of pyrimidine (6-4) pyrimidone photoproducts in human cells are due to excision repair rather than conformational change. *Nucleic Acids Res.* **18,** 963–971.
15. Pfeifer, G. P. (1997) Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment. *Photochem. Photobiol.* **65,** 270–283.
16. Faisst, S. and Meyer, S. (1992) Compilation of vertebrate-encoded transcription factors. *Nucleic Acids Res.* **20,** 3–26.
17. Becker, M. M. and Wang, J. C. (1984) Use of light for footprinting DNA in vivo. *Nature* **309,** 682–687.
18. Pfeifer, G. P. and Riggs, A. D. (1993) Genomic sequencing, in *Methods in Molecular Biology, DNA Sequencing Protocols, vol. 23* (Griffin, H. and Griffin, A., eds.), Humana Press, Totowa, NJ, pp. 169–181.
19. Törmänen, V. T. and Pfeifer, G. P. (1992) Mapping of UV photoproducts within *ras* protooncogenes in UV-irradiated cells: correlation with mutations in human skin cancer. *Oncogene* **7,** 1729–1736.
20. Rychlik, W. and Rhoads, R. E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. *Nucleic Acids Res.* **17,** 8543–8551.

3

Identification of DNaseI Hypersensitive Sites Within Nuclei

Peter N. Cockerill

1. Introduction

The vast bulk of the genome of eukaryotic cells exists as a highly regular array of nucleosomes. A single nucleosome comprises 146 bp of DNA wrapped around a histone core particle. Arrays of nucleosomes usually exist as condensed chromatin fibers with a nucleosome repeat length of about 200 bp. The DNA within condensed chromatin is for the most part relatively inaccessible to proteins, such as nucleases. However, about 1% of the genome exists as discrete regions of decondensed chromatin termed DNaseI hypersensitive (DH) sites **(1)** which provide greatly increased access to factors that interact with DNA. DH sites represent local perturbations of the nucleosome array and probably arise via the reorganization of single nucleosomes within the chromatin fiber. Wherease DH sites exist as open regions of chromatin in nuclei, they are not typically hypersensitive to DNaseI within protein-free DNA.

DH sites are powerful indicators for the existence of specific structural or regulatory elements such as enhancers, locus control regions, silencers, and boundary elements **(1)**. DH sites have frequently provided the first evidence of the existence of distant regulatory elements located several kilobases upstream or downstream of genes, or in some cases within introns. The pattern of DH sites in the vicinity of a gene is often a reflection of the activity of that gene. Hence, some DH sites have tissue-specific or inducible patterns of distribution. These types of DH sites probably arise via interactions with specific transcription factors or developmental regulators. Other DH sites can represent specific structural features, such as centromeres.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

Not all DH sites have clearly defined functions or factor-binding sites. Some DH sites appear to arise as a result of sequence-specific positioning of nucleosomes, which leaves a natural gap in the nucleosomal array. However, both naturally occurring and factor-inducible DH sites can fulfill the role of providing increased access to regulatory molecules within the nucleus. A comprehensive discussion of the properties of DH sites is available in the review by Gross and Garrard (**1**).

Although the actual structure of DH sites in vivo is essentially unknown, in the past they have been assumed to represent nucleosome-free regions of DNA. This assumption has stemmed largely from the observation that DH sites usually occupy about 200 bp of DNA, which represents the average nucleosome repeat length. There is, however, the potential for essentially any change in the folding of the chromatin fiber or the organization of the nucleosome to result in the appearance of a DH site. Furthermore, it is now known that nucleosomes remain associated with the DNA within at least some DH sites. It is best, therefore, to regard DH sites as regions where the local organization of one or more nucleosomes differs from the surrounding nucleosomal array.

To understand the nature of DH sites it is important to be aware that DH sites are not single sites of DNaseI cleavage, but broad regions of enhanced DNaseI sensitivity. After Southern-blot hybridization analysis, DH sites usually appear as discrete bands migrating faster than the parent restriction enzyme fragment encompassing the specific probe. Although some DH sites appear as a single diffuse band 150–250 bp across, many DH sites are more complex and include either protected regions or regions of enhanced DNaseI cleavage (**1**,**2**). Hence, when more than one DH region is detected within the space of 200–300 bp, it should be treated as a single DH site. In instances where transcription factors induce the formation of a DH site, it is common to encounter a short protected region covering the factor binding site within the DH site.

DH sites are, by definition, identified on the basis of their hypersensitivity to digestion by DNaseI within a chromatin environment. Consequently, assays for DH sites usually employ isolated nuclei as a substrate. Purified DNA cannot be used to detect DH sites as they only exist in the context of chromatin. Whereas DNaseI remains the agent most commonly used to detect DH sites, they can also be detected by employing any of a wide variety of DNA-modifying reagents (**1**). In this chapter, I will only address the use of DNaseI.

As most of the DNA in a chromatin fiber is susceptible to prolonged digestion by DNaseI, DH-site analyses are performed under very limited digestion conditions. Ideally, DNaseI should cut on average only once within the restriction enzyme fragment chosen for analysis, and a proportion of the parent fragment should remain intact. In practice, the enzyme may cut several times within

the DH site itself. The hypersensitive DNaseI cleavage sites are then identified by Southern blot hybridization analysis using the indirect end-labeling method. In this procedure, the probe detects a heterogeneous mixture of digestion products. In regions where DH sites exist, these products are more concentrated and cluster to form obvious bands on Southern blots. These approaches were pioneered by Wu and others **(1)**, and I employ a modification of the method described by Siebenlist et al. **(3)**.

The detection method is illustrated in **Fig. 1**, and the general approach can be summarized as follows. Nuclei are recovered from either homogenized or detergent-lysed cells in a low-salt buffer. Nuclei are digested for just a few minutes at room temperature with a range of 1–10 µg/mL of DNaseI, and then lysed in SDS to stop the digestion. The optimum range of DNaseI has to be determined empirically, and the viscosity of the nuclear lysates provides the best indication of the extent of DNaseI digestion. The DNA is then purified, digested with restriction enzymes, and analyzed by Southern-blot hybridization. The indirect end-labeling detection method (**Fig. 1**) employs a short hybridization probe that abuts one end of the restriction enzyme fragment under investigation and reveals specific DNaseI cleavage sites that occur at discrete distances from that end. DH sites can be identified at distances up to about 15–20 kb from the probe using this approach, but are best mapped using restriction enzymes that cut about 2–10 kb from the region of interest. To determine the locations of the boundaries of a DH site, it should be mapped from both directions as much of the DNA within the DH site itself may be digested away. In most cases, probes should be made from clones of genomic DNA rather than cDNA, as cDNA clones are likely to span intron/exon boundaries. In instances where genomic clones are unavailable, it is still sometimes possible to map DH sites in the genomic locus through careful selection of the restriction enzymes and cDNA probe fragments employed as probes.

An example of DH-site detection is shown in **Fig. 2B** where a 1.0 kb probe is used to detect a cluster of DH sites located 1.5, 2.2, and 3.5 kb from one end of a 4.5 kb BamHI fragment of genomic DNA. Note that the proximal and distal DH sites appear to different extents depending on the degree of DNaseI digestion. For this reason it is usually necessary to analyze a series of DNaseI digestions in the first instance. Also note that the mobility of all of the fragments in a mixture is greater when the average fragment length is shorter (compare lanes 1 and 2 of **Figs. 2A** and **2B**). For both of the reasons mentioned above, it is preferable to compare different samples of nuclei that have been DNaseI-digested to the same extent. **Fig. 3** illustrates the principle of mapping DH sites from both directions and employing restriction enzyme sites as internal markers (*see* **Subheading 3.7.** for discussion).

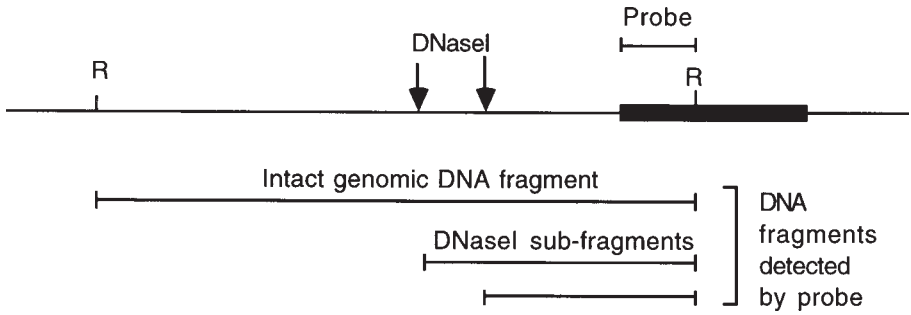


Fig. 1. Detection of DH sites. Limited DNaseI digestions are performed such that only one region of DNA is cut in the vicinity of the gene (indicated by the box) in any one copy of the gene. If DNaseI cuts at more than one site, then a mixture of DNA fragments will be generated. The DNA is purified and cut with a restriction enzyme (R) carefully selected to allow examination of a specific region, usually upstream of the gene. The DNA is then separated by electrophoresis, transferred to a membrane, and hybridized to a probe that recognises just one end of the restriction enzyme fragment. When DH sites are present, one or more specific subfragments of DNA will appear in addition to the intact parent restriction enzyme fragment of genomic DNA.

2. Materials

2.1. Stock Solutions Stored at Room Temperature

1. TE: 10 m*M* Tris, 1 m*M* EDTA, pH 7.4.
2. 100 m*M* $CaCl_2$.
3. 100% ethanol (equal or greater than 99%).
4. 75% (v/v) ethanol: Prepared by mixing one part 50 m*M* sodium acetate to three parts 100% ethanol.
5. Chloroform.
6. Southern-blot denaturation buffer: 1.5 *M* NaCl, 0.5 *M* NaOH.
7. Southern-blot neutralization buffer: 1.5 *M* NaCl, 1 *M* Tris pH 7.0.
8. 10X SSC: 1.5 *M* NaCl, 150 m*M* sodium citrate, pH 7.0.
9. RapidHyb hybridization buffer (Amersham, Buckinghamshire, UK).
10. 10% (w/v) sodium dodecyl sulfate (SDS).
11. Nuclei lysis buffer: 300 m*M* sodium acetate, 5 m*M* EDTA, pH 7.4, 0.5 % SDS.
12. 3 *M* sodium acetate.
13. 10X TNAE gel electrophoresis buffer: 500 m*M* Tris, 200 m*M* sodium acetate, 20 m*M* EDTA, adjusted to pH 8.0 with acetic acid.
14. 0.5 *M* sodium phosphate: Prepare from sodium dihydrogen phosphate adjusted to pH 7.0 with sodium hydroxide.
15. 0.2 *M* sodium pyrophosphate: Prepare from disodium pyrophosphate adjusted to pH 7.0 with sodium hydroxide.

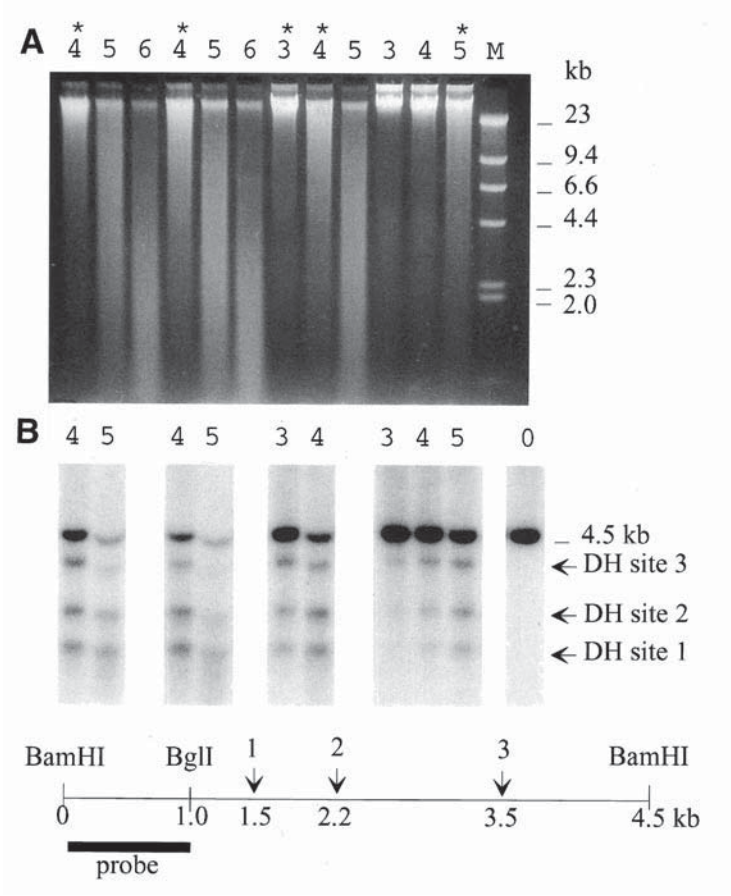


Fig. 2. Analysis of DNaseI-digested nuclei prior to DNA purification (A) and following Southern-blot hybridization (B). The concentrations of DNaseI used to digest nuclei are indicated above each lane (ug/mL) and those samples that subsequently yielded optimal detection of DH sites are asterisked. Shown in panel A are four series of three sequential DNaseI digested samples of Jurkat T cell nuclei. Lanes contain 15 µL of nuclear lysate or 1 µg of HindIII-digested lambda DNA (M). Lane 1 shows the ideal pattern of DNaseI digestion, which subsequently gave the best detection of DH sites. Lanes 3, 6, and 9 show DNA that was too extensively digested and was therefore eliminated prior to DNA purification. Lane 10 shows underdigested DNA. Panel B shows a subsequent Southern blot of the same DNaseI-digested samples following digestion of the purified DNA with BamHI. The indirect end-labeling technique was used to detect DH sites within a 4.5 kb BamHI fragment of the human interleukin-3 locus. Also included in panel B is a sample of intact genomic DNA digested with BamHI (0). The probe is a 1.0 kb BamHI/BglI fragment of DNA, and the three DH sites are arrowed.

16. EGTA: 50 m*M* ethylene glycol tetraacetic acid adjusted to pH 7.4 with NaOH (*see* **Note 1**).
17. Restriction enzyme stop buffer (electrophoresis loading buffer): 20% Ficoll (Sigma type 400), 1% SDS, 0.05% bromophenol blue.
18. 10 mg/mL Ethidium bromide.

2.2. Stock Solutions Stored at 4°C

1. 10% (v/v) NP40 (Nonidet P40, Sigma, St. Louis, MO).
2. 2.2 *M* Sucrose: autoclave and maintain sterile to avoid fungal growth.
3. 10X cell lysis buffer and DNaseI digestion buffer: 600 m*M* KCl, 150 m*M* NaCl, 50 m*M* $MgCl_2$, 1 m*M* EGTA, 100 m*M* Tris, pH 7.4.
4. 10X hypotonic lysis buffer: 100 m*M* NaCl, 50 m*M* $MgCl_2$, 1 m*M* EGTA, 100 m*M* Tris, pH 7.4.
5. Phenol: Liquefied phenol prepared by extracting crystalline phenol twice with an equal volume of 100 m*M* Tris, pH 8.0, 1 m*M* EDTA, and adding hydroxyquinilone to 0.1% as an antioxidant. Phenol can be stored at 4°C for up to 1 yr if kept dark.
6. 10 mg/mL herring DNA: carrier DNA is prepared by dissolving either herring or salmon testes DNA in TE to 10 mg/mL by shaking overnight. The DNA is then sonicated to shear it to fragments several hundred base pairs in length.
7. PBS: 137 m*M* NaCl, 2.68 m*M* KCl, 1.47 m*M* KH_2PO_4, 8.1 m*M* Na_2HPO_4.

2.3. Stock Solutions Stored at –20°C

1. 2 mg/mL DNaseI (*see* **Note 2**): Prepare 500 µL aliquots of DNaseI in 50% glycerol by dissolving 5 mg (~10,000 U) Worthington DPFF bovine pancreatic DNaseI in 1.25 mL 10 m*M* NaCl, 3 m*M* $MgCl_2$, 1 m*M* $CaCl_2$, 10 m*M* Tris, pH 7.4. followed by the addition of 1.25 mL glycerol.
2. 50 m*M* PMSF: Phenyl methyl sulfonyl fluoride (PMSF) (Sigma) dissolved in isopropanol (*see* **Note 3**).
3. 10 mg/mL BSA: 10 mg/mL purified bovine serum albumin (Sigma # A-7638) dissolved in water.
4. 10 mg/mL proteinase K: Dissolve 100 mg of Proteinase K (Merck, Darmstadt, Germany) in 5 mL 10 m*M* $CaCl_2$, 20 m*M* Tris, pH 7.4. Add 5 mL glycerol and store in 1 mL aliquots.
5. 5 mg/mL leupeptin (Sigma).
6. 10 mg/mL aprotinin (Sigma).

2.4. Solutions Prepared on Their Day of Use

1. Cell lysis buffer: 60 m*M* KCl, 15 m*M* NaCl, 5 m*M* $MgCl_2$, 10 m*M* Tris, pH 7.4, 300 m*M* sucrose, 0.1 m*M* EGTA, 0.5 m*M* PMSF (added fresh), 5 µg/mL leupeptin, 10 µg/mL aprotinin, 0.1% (v/v) NP40 (*see* **Note 4**).
2. Hypotonic lysis buffer: 10 m*M* NaCl, 5 m*M* $MgCl_2$, 10 m*M* Tris, pH 7.4, 0.1 m*M* EGTA, 0.5 m*M* PMSF (added fresh), 5 µg/mL leupeptin, 10 µg/mL aprotinin (*see* **Notes 4** and **5**).

3. Nuclei digestion buffer: 60 m*M* KCl, 15 m*M* NaCl, 5 m*M* $MgCl_2$, 10 m*M* Tris, pH 7.4, 300 m*M* sucrose, 0.1 m*M* EGTA.
4. Proteinase K/nuclei lysis buffer: nuclei lysis buffer containing 0.1 mg/mL proteinase K (made up just before use).
5. Low stringency hybridization wash buffer: 2X SSC, 0.1% SDS, 25 m*M* sodium phosphate, pH 7.0.
6. High stringency hybridization wash buffer: 0.1X SSC, 0.1% SDS, 1 m*M* sodium pyrophoshate, pH 7.0.

2.5. Special Reagents and Equipment

1. Refrigerated benchtop centrifuge with swingout rotor to hold 10 or 50 mL tubes and ability to spin at 3000 rpm (*see* **Note 6**).
2. Heated water bath set at 55°C to heat 10 mL tubes.
3. Shaking heated water set at 65 or 67°C for membrane hybridization and washing.
4. Platform shaker for washing membranes at room temperature.
5. Horizontal gel electrophoresis apparatus with 15 × 15 cm gel tray.
6. Variable voltage electrophoresis power supply (e.g., 200 V).
7. Nylon membrane (e.g., Amersham HybondN).
8. α-[^{32}P]-dATP (or other nucleotide as appropriate).
9. Random oligonucleotide-primed DNA labeling kit (e.g., Geneworks, Gigaprime, Adelaide, Australia).
10. Polythene bags and plastic bag sealer for hybridising and washing membranes.
11. X-ray film and cassettes or alternative imaging system.
12. Dounce homogenizer for use with hypotonic lysis buffer only.
13. Agarose.
14. UV light transilluminator and camera.
15. India ink.
16. Whatman 3MM paper.
17. Screw capped 1.5 mL centrifuge tubes.
18. Boiling water jug.

3. Methods

3.1. Isolation of Nuclei (*see* Notes 6–8)

Nuclei isolations and DNaseI digestions should be completed within the shortest time possible, which should amount to no more than 2 h. It is ideal, therefore, if DNaseI assays are performed on no more than two or three lots of cells at a time. It is possible, however, to handle up to six cultures at a time following this protocol. Nuclei isolation should be done in an ice bucket in the cold room and using a precooled centrifuge (*see* **Note 6**). Great care must be taken to handle nuclei gently and to pre-cool all items used in this procedure.

1. Grow approximately (2–6) × 10^7 cells for each series of DNaseI digestions (*see* **Note 7**). This will yield approx 2–6 mL of nuclei at a concentration of 0.4 mg/mL nucleic acid, which is sufficient for three to six DNaseI digests. If fewer than 10^7

cells are available, then only one or two DNaseI digestions per sample are practical, and it becomes more difficult to achieve an optimal DNaseI digestion.
2. Spin down cells at 1500 rpm for 5 min in 50 mL screw-cap plastic tubes (Falcon).
3. Resuspend the cells in a total of 40 mL ice cold PBS and spin down at 1500 rpm for 5 min. Pour off the supernatant and drain the tubes well.
4. Resuspend the cell pellet in 5 mL of cell lysis buffer. Make into a uniform suspension and reduce the amount of cell debris associated with the nuclei by squeezing the lysed cells five times out of a 10 mL pipet held hard on the bottom of the 50 mL plastic centrifuge tube (*see* **Notes 8** and **9**).
5. Make the cell lysate up to 30 mL in cell lysis buffer, spin down the nuclei at 1500 rpm for 5 min, and allow the centrifuge to stop without the brake when the speed reaches 300 rpm. Draw off the supernatant with a pipet, taking care to remove bubbles and to not disturb the nuclear pellet.
6. To remove the final traces of supernatant, respin the tube for 1 min at 1500 rpm and draw off the remaining liquid (*see* **Note 10**).
7. Resuspend the nuclear pellet in 1 mL of nuclei digestion buffer for each amount of 2 x 10^7 cells present in the starting material. This should yield a suspension of approximately 0.4–0.8 mg/mL total nucleic acid as determined by absorbance in **Subheading 3.2.**

3.2 Quantitation of the Nuclei

The extent of DNaseI digestion that occurs in a reaction is determined by the concentration of the nuclei as well as the concentration of DNaseI. It is recommended, therefore, to employ uniform concentrations of nuclei in each series of DNaseI digestions to obtain optimal results. The following procedure estimates the total nucleic acid concentration in a suspension of nuclei. This value is typically 1.6-fold greater than the final yield of purified DNA.

1. Remove 10 µL nuclei, add 1 µL 10 m*M* $CaCl_2$ and ~0.5 µL 2 mg/mL DNaseI and stand at room temperature ~1 min.
2. Add 500 µL 1 *M* NaOH and vortex.
3. Determine the total nucleic acid concentration by measuring the absorbance at 260 nm. Total nucleic acid concentration is estimated by assuming that 1 mg/mL nucleic acid in 1 *M* NaOH has an absorbance of 27 at 260 nm. A suspension of nuclei at a concentration of 0.4 mg/mL nucleic acid will give an absorbance of 0.21 when quantitated in this way.

3.3. DNaseI Digestions

The aim is to perform a series of DNaseI digestions designed to include a point at which DNaseI cuts on average only once every 10–30 kb (*see* **Note 11**). Each preparation of nuclei requires a different range of DNaseI concentrations to achieve optimal detection of DH sites and the ideal window is often very narrow (*see* **Note 12**). As the optimum usually occurs between 1 and 10 µg/mL DNaseI (2–20 U/mL), it is often best to commence with just one or

two digests in the range of 3–5 µg/mL DNaseI, and to vary the subsequent digestions according to the viscosity of the lysed nuclei (*see* **Notes 13** and **14**).

1. Adjust the concentration of nuclei to 0.4 mg/mL nuclei acid by addition of nuclei digestion buffer (*see* **Note 15**).
2. Prepare fresh stocks of 10X DNaseI by diluting DNaseI in nuclei digestion buffer containing 1 m*M* $CaCl_2$ and 0.1 mg/mL BSA (*see* **Note 16**).
3. Add 5 µL 100 m*M* $CaCl_2$ to a series of up to six 10-mL polypropylene screw-capped tubes (*see* **Note 17**).
4. Resuspend the suspension of nuclei and transfer 500 µL of nuclei to the first 10 mL tube containing $CaCl_2$. Gently mix by tapping and transfer to a bucket of water adjusted to 22°C to equilibrate for 3 min.
5. At 30 s intervals, prepare up to five additional 500 µL aliquots of nuclei equilibrating at 22°C.
6. After three minutes equilibration at 22°C add 50 µL of 10X DNaseI solution to the first tube, mix gently by tapping (do not vortex), and return to 22°C. Incubate for a further 3 minutes, re-mixing once by tapping after 1 or 2 min if time permits.
7. At 30 s intervals, add DNaseI to the remaining tubes.
8. After 3 min terminate the DNaseI digestion by adding 3.5 mL proteinase K/nuclei lysis buffer to the first tube, recap the tube, and mix by inversion.
9. At 30 s intervals, add proteinase K/nuclei lysis buffer to the remaining tubes.
10. Add one aliquot of 500 µL undigested nuclei directly to 3.5 mL of proteinase K/nuclei lysis buffer to provide an intact genomic DNA control sample (*see* **Note 18**).
11. Invert each tube several times and transfer tubes to a 55°C water bath for 5 min to allow for complete dispersal of nuclei. Return the tubes to the 22°C water bath for 2 min to cool the tubes. Check the viscosity of the digested DNA by inverting and swirling the tube several times and observing the motion of air bubbles moving within the DNA solution (*see* **Note 13**).
12. Repeat **steps 2** to **11** as many times as is required (*see* **note 19**).

3.4. Screening of DNA Digests

When DNaseI analyses include several lots of cells, it is not feasible to include every DNaseI-digested sample in the same Southern blot. It is desirable, therefore, to screen out under- or over-digested DNA samples before proceeding to DNA purification (*see* **Note 20**). This can be done while the nuclear lysates are digesting in proteinase K. Samples should be taken after the first hour of proteinase K digestion at 55°C (after **step 1** or **2** in **Subheading 3.5.**).

1. Prepare a 0.7% agarose gel containing 0.5 µg/mL ethidium bromide and TNAE buffer (ideally 15 × 15 cm with 20 5-mm wide wells).
2. Load 15 µL of each nuclear lysate plus 3 µL 20% Ficoll, 0.05% bromophenol blue, and electrophorese overnight (e.g., 30 V for 15 h).
3. Select samples for further analysis that exhibit a distribution of DNA fragments between about 5 and 30 kb in length (*see* **Note 21**).

An example of this screening procedure is shown in **Fig. 2A**, and a subsequent Southern-blot analysis of DH sites within a BamHI fragment of DNA is shown in **Fig. 2B**. The optimal DNaseI-digested samples are marked with an asterisk. An ideal DNaseI digestion is seen in lane 1, where 4 µg/mL DNaseI leaves most of the DNA greater than 20 kb in length, with a tail of DNA trailing below the main band (*see* **Fig. 2A**). This sample subsequently allowed optimal detection of all three DH sites as well as the parent BamHI fragment (*see* **Fig. 2B**). In another digest with 3 µg/mL DNaseI, this tail of digested DNA is very faint (*see* **Fig.2A**, lane 10), the parent BamHI fragment remains largely intact and the three DH sites are weak or undetectable (*see* **Fig. 2B**). In two other digests with 5 µg/mL DNaseI where most of the DNA is less than about 20 kb in length (*see* **Fig. 2A**, lanes 2 and 5), the parent BamHI fragment is nearly all degraded and the distal DH sites have become very faint or undetectable (*see* **Fig. 2B**).

3.5. Purification of DNA

1. Incubate the tubes of lysed nuclei at 55°C for 1 h.
2. Incubate samples at 37°C overnight and store at room temperature until ready for further purification.
3. Select samples suitable for further analysis on the basis of the preliminary screen (*see* **Subheading 3.4.**).
4. Add additional proteinase K to 0.1 mg/mL.
5. Incubate at 55°C for 1 h and cool to room temperature.
6. Phenol extract with 4 mL phenol by rotating on a wheel for 1 h (*see* **Note 22**).
7. Centrifuge for 15 min at 3000 rpm.
8. Remove the supernatant containing the DNA by Pasteur pipet (*see* **Note 23**).
9. Repeat the extractions with 1:1 phenol/chloroform and with chloroform, removing no more than 3.3 mL of supernatant from the final extraction.
10. Ethanol precipitate the DNA by adding 2 vol of ethanol and gently rocking or inverting the tube until the DNA fibers have come out of solution.
11. Pellet the DNA by centrifuging at 3000 rpm for 10 min.
12. Wash the pellet by resuspending once in 75% ethanol and once in 100% ethanol and centrifuging for 5 min at 2000 rpm.
13. After taking off the final supernatant, respin the tubes for 1 min, remove all final traces of ethanol and allow the DNA pellets to dry.
14. Resuspend the DNA pellets in a 0.25 vol of TE relative to the original DNaseI digest (i.e., 125 µL TE for a 500 µL aliquot of nuclei) to give a solution of ~1.0 mg/mL DNA. Allow the DNA to redissolve by standing at room temperature overnight.
15. Transfer the DNA to 1.5 mL Eppendorf tubes while also checking that the DNA is evenly dissolved by extruding from a pipet tip once or twice. Undigested DNA will be very viscous, and you will need to cut the ends of tips when pipetting.
16. Store DNA at 4°C (*see* **Note 24**).
17. Determine the DNA concentration by diluting 10 µL of DNA in 500 µL TE. As 1 mg/mL DNA in TE gives an absorbance of 20 at 260 nm, a 1 mg/mL DNA sample will give an absorbance of 0.4 when diluted and measured in this way.

3.6. Southern-Blot DNA Transfer

1. Digest 10 µg of each DNA in 15 µL for 2–3 h with 30 U of restriction enzyme (*see* **Notes 25** and **26**).
2. Stop the digestion by adding 3 µL 1 % SDS, 20% ficoll, 0.05% bromophenol blue.
3. Load onto a 0.6–0.7-cm-thick 15 cm × 15 cm 0.8% agarose gel containing TNAE buffer and 0.5 µg/mL ethidium bromide (*see* **Note 27**). Include one or two size markers covering the range of 1 to 20 kb (e.g., 1 µg of HindIII-digested lambda DNA).
4. Apply 25–30 V for about 15 h (400–450 V × h), so that the bromophenol blue has migrated 70% of the distance to the end of the gel.
5. Take a photograph of the gel on a UV light transilluminator.
6. Mark the DNA size marker bands with a syringe needle dipped in India ink (a slurry of photocopy toner will also do).
7. Wash the gel twice for 15 min in 0.5 *M* NaOH, 1.5 *M* NaCl to denature the DNA.
8. Wash the gel twice for 20 min in 1 *M* Tris, pH 7, 1.5 *M* NaCl to neutralize the gel.
9. Transfer the DNA to a membrane as for a conventional Southern blot. For best results use a nylon membrane such as HybondN (Amersham). Label a 15 cm × 15 cm sheet of HybondN membrane in pencil before wetting and keep track of which surface is in contact with the gel during the transfer. Wet the membrane by floating it onto water, and wash it in 10X SSC. Transfer the DNA from the gel to the membrane in 10X SSC overnight using a wick of three sheets of Whatman 3MM paper immersed in 500 mL of 10X SSC under the gel and with six sheets of Whatman 3MM paper and a 3-cm stack of blotting paper above the gel.
10. After the transfer is complete, carefully remove the paper but not the membrane. Stab a needle through the membrane at points where DNA markers lie to calibrate the membrane.
11. Wash the membrane in 2X SSC, blot well, and air dry.
12. Wrap the membrane in plastic film (Gladwrap or Saran Wrap).
13. Fix the DNA to the membrane by placing the DNA side down on a medium wavelength UV light (~305 nm) transilluminator for 30 s (*see* **Note 28**).

3.7. DNA Hybridization

As outlined in **Fig. 1**, the aim is to hybridize the DNA on the membrane with a DNA probe that abuts one end of the restriction enzyme fragment being examined. The DNA probe should be 0.5–1 kb in length and must not overlap any DH sites. I have obtained the best hybridization results using Amersham RapidHyb buffer as the DNA hybridization solution. This gives rapid and enhanced hybridization relative to conventional aqueous buffer systems (*see* **Note 29**).

1. For a 15 cm × 15 cm nylon membrane, prehybridize for 3–6 h at 67°C in 20 mL RapidHyb buffer containing 0.25 mg/mL boiled sonicated salmon or herring DNA (from 10 mg/mL stock) in a heat-sealed plastic bag immersed in a shaking water bath.

2. Label 40 ng of the DNA fragment to be used as a probe with ^{32}P by oligonucleotide priming (e.g., by using Geneworks gigaprime kit).
3. Purify the probe on a Sephadex G50 spin column to separate free ^{32}P (*see* **Note 30**).
4. Denature the probe together with 5 mg sonicated herring DNA by standing in a screw-capped 1.5 mL tube (not a flip-top) in a jug of boiling water for 5 min.
5. Cool on ice, dilute with 0.5 mL RapidHyb buffer, and add to the solution already in the bag, to give final concentrations of 2 ng/mL probe and 0.5 mg/mL carrier DNA.
6. Hybridize at 67°C for 2 h in a shaking water bath.
7. Remove the membrane and wash it twice for 10 min at room temperature in 500 mL 2X SSC, 25 m*M* sodium phosphate pH 7, 0.1% SDS.
8. Wash the membrane twice for 15 min at 65°C in 250 mL prewarmed 0.1X SSC, 1 m*M* sodium pyrophosphate pH 7, 0.1% SDS in a heat-sealed plastic bag immersed in a shaking water bath.
9. Blot the membrane partly dry and wrap it in plastic film (Gladwrap or Saran Wrap).
10. Detect the DH sites either by autoradiography with X-ray film at –70°C for several days with one intensifying screen, or overnight with a Phosphorimager.

By following the above procedures, DH sites usually appear as a single band, and their position can be mapped to within approx 100–200 bp of their exact location in the genome. Once DH sites have been identified, they can be mapped at higher resolution by using restriction enzyme sites located much closer to the DH site (within 0.5 to 1.0 kb). This requires the use of higher concentration agarose gels (1.2–2%) and allows mapping of the boundaries of the DH site to within about 50 bp. Fine mapping of DH sites is greatly assisted by including genomic DNA that is partially digested at additional restriction enzyme sites that lie close to the DH site and serve as internal markers (*see* **Note 31**).

An example of high-resolution DH site mapping is shown in **Fig. 3** where BalI, PstI, ApaI, and BamHI serve as internal markers to determine the location of a DH site relative to BglI sites located upstream and downstream of the DH site. This analysis employs probes of 500 and 600 bp in length to examine a DH site located 600 and 800 bp from the upstream and downstream BglI sites. At this level of analysis the DH site spans about 250 bp and encompasses two protected regions flanking a central region of enhanced DNaseI digestion. This example also indicates that different sets of DNaseI cleavage sites within the DH site are detected preferentially depending on whether the probe is located upstream or downstream of the DH site. Also included in **Fig. 3** is a control DNaseI digest of purified DNA which confirms that the DH site is a feature of chromatin structure rather than a DNA sequence that is intrinsically hypersensitive to DNaseI.

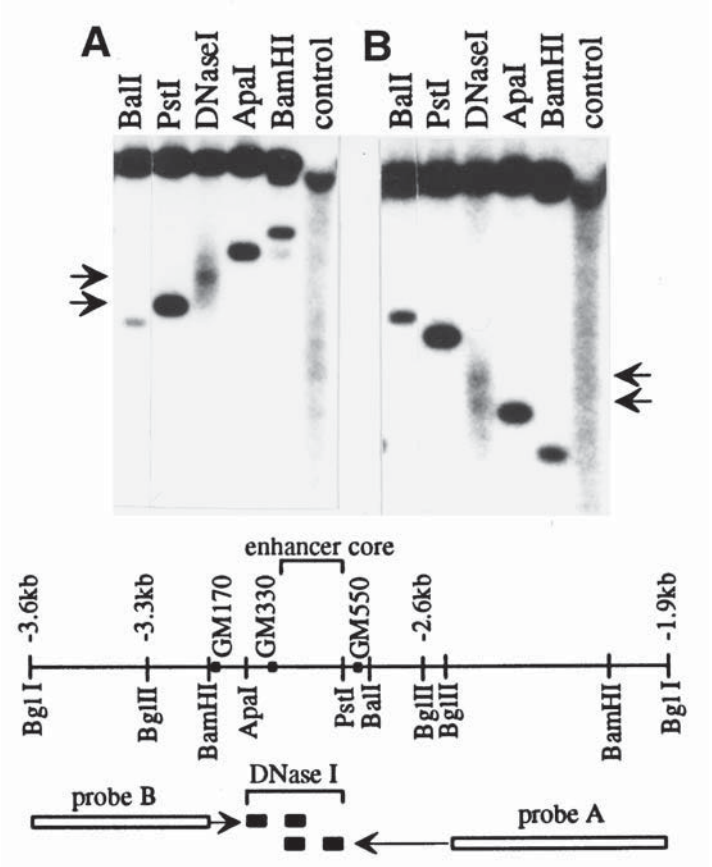


Fig. 3. Fine mapping of an inducible DNaseI hypersensitive site within the human GM-CSF enhancer *(2)*. Nuclei isolated from activated Jurkat cells were digested with 5 μg/mL DNaseI. The DH site was mapped from downstream (A) and upstream (B) by Southern-blot hybridization analysis with probes A and B, respectively, utilising BglI sites located 1.9 (A) and 3.6 (B) kb upstream of the GM-CSF gene. All DNA samples were digested to completion with BglI. The marker lanes utilise intact purified Jurkat cell DNA partially digested with BalI, PstI, ApaI, or BamHI. The control lane contains a DNaseI digest of intact purified DNA. The same filter was used for both panel A and panel B. Areas of enhanced DNase I cleavage within the DH site are indicated by arrows on panels A and B and are displayed as solid boxes on the map below. The location of the GM-CSF enhancer core is bracketed. (This figure has been reproduced from ***[2]*** with the permission of the American Society for Microbiology.)

4. Notes

1. The free-acid form of EGTA is insoluble in water. It must be dissolved by adding NaOH while stirring until the desired pH is achieved.

2. The traditional source of DNaseI for these assays is Worthington bovine pancreas DNaseI that has a specific activity of approximately 2 U/μg. As DNaseI is dependent upon divalent cations, it should not be prepared in buffers containing chelating agents such as EDTA or EGTA. DNaseI digestions should also include both Ca^{2+} and Mg^{2+} to achieve efficient double-stranded cleavage of DNA. Aliquots of DNase I can be used repeatedly and stored for in excess of one year without significant loss of activity. DNaseI is, however, easily denatured if mistreated.
3. PMSF is highly toxic and great care must be taken during handling. As crystallization of PMSF in isopropanol occurs below 4°C, the solution should be warmed and the crystals redissolved before use. PMSF is an unstable serine protease inhibitor and must be added to aqueous solutions just before use. As PMSF is sparingly soluble in water, aqueous solutions must be mixed vigorously immediately upon addition of the stock solution or precipitation of PMSF will occur.
4. The Ca^{2+} chelator EGTA and the protease inhibitors PMSF, aprotinin and leupeptin are likely to maintain the integrity of the nuclei and nuclear proteins, but are not essential for the detection of DH sites.
5. Hypotonic lysis buffer is only required if NP40 lysis buffer is found to be unsuitable.
6. This protocol describes the isolation of cells grown in suspension culture and utilizes a Beckman GS-6R refrigerated benchtop centrifuge with a swing-out rotor. All steps utilize polypropylene screw-cap centrifuge tubes. Speeds of 1500, 2000, and 3000 rpm correspond to maximum *g* forces of 470, 830, and 1900.
7. Nuclei are most easily prepared from cells cultured in suspension but can also be recovered from adherent cell cultures and from most easily homogenized animal tissues. When isolating nuclei from animal tissues it may be necessary to filter the cell homogenates through a material such as Miracloth (Calbiochem, La Jolla, CA) and to sediment the nuclei though nuclei isolation buffer containing 1.8 *M* sucrose by spinning at approximately 10,000*g* for 20 to 30 min.

   When employing cells grown in suspension culture, 120 mL of cells at a concentration of 5×10^5 cells/mL represents the ideal amount of starting material for a series of four to six DNaseI digestions. I usually use between 80 and 200 mL of cells, and reserve 30 mL of cells for parallel RNA analyses. When more than three cultures are handled at a time, it is most practical to start with 80 mL of cells and to limit the digestion series to three or four different concentrations of DNaseI, which is the number of digests usually required to give the optimal appearance of DH sites.
8. Nuclei should be prepared in buffers containing Mg^{2+}. Great care must be taken to avoid damaging the nuclei or extensive clumping of the nuclei will occur. This is best achieved by performing all steps on ice or at 4°C, treating the nuclei gently, and avoiding unnecessary or excessive homogenization or centrifugation steps. Nuclei are less prone to clumping if cells are lysed by hypertonic shock rather than lysis in nonionic detergent. Clumping usually results in the formation of a single rubbery aggregate of nuclei, which is not suitable for DNaseI analysis. Clumping usually occurs as a result of centrifugation at too high a speed, per-

forming too many centrifugation steps, allowing the nuclei to warm up, or from overly vigorous homogenization or resuspension of the nuclei. Also be aware that nuclei like to stick to surfaces such as glass, and are easily lost on the walls of pipets. Pasteur pipets should not be used for pipeting nuclei.

9. Do not employ any form of homogenizer at this stage or the nuclei may become damaged.
10. This step is optional but it may remove some cytoplasmic inhibitors of DNaseI, and reduces the residual quantities of protease inhihibitors that may inhibit the proteinase K in the nuclei lysis buffer.
11. Although DH sites appear as sites of infrequent DNaseI cleavage, DNaseI may actually cleave several times within the DH site itself. This results in the DH site having a slightly different appearance and apparent location depending upon which side it is mapped from. Where more than one DH site is present, the distal DH sites will be lost in the analysis as the DNaseI digestion proceeds to greater extents. An example of this can be seen in **Fig. 2** where the distal DH site 3 is the most prominent site at lower DNaseI concentrations and the proximal DH site 1 is the most prominent site in the more extensive DNaseI digestions.
12. Preparations of nuclei contain inhibitors of DNaseI such as actin. These inhibitors need to be titrated out before significant DNaseI digestion commences. This effect can give rise to a very narrow window of optimum DNaseI digestion (typically 3–5 μg/mL DNaseI). Different concentrations of DNaseI may even be required when parallel cultures of the same cells are used. These phenomena are apparent in **Fig. 2A**, which employs DNaseI titrations of four separate preparations of nuclei derived from the same cells. Each of these preparations has a slightly different optimum DNaseI concentration, and marked differences in DNaseI digestion are evident with increases of just 1 μg/mL of DNaseI within each of the series.
13. DNA viscosity gives an early indication of the extent of DNaseI digestion. Undigested nuclei yield highly viscous lysates. The ideal digest is one where the lysed nuclei offer some resistance to swirling when the tube is shaken and that allows air bubbles to rise slowly. If the DNA is under- or over-digested, air bubbles will either become trapped or rise rapidly when the SDS lysate is shaken. It is important to achieve a range of states of viscosity to ensure that suitable digests are obtained. If a program of many consecutive DH site analyses is planned, it is well worth the effort of generating a series of control lysates that can be saved as a reference point to estimate the extent of DNaseI digestion. This series of controls is most easily obtained if the first series of DNaseI titrations is performed in duplicate. One sample of each can then be saved as a control lysate and one sample processed further to detect DH sites. Note that if clumps of damaged nuclei are present in the suspension of nuclei, then the final lysate may remain viscous under conditions where most of the DNA is in fact digested.
14. It is preferable to perform DNaseI digestions sequentially so that the concentration of DNaseI used in a series of reactions can be increased or decreased as required. For best results it is recommended to increase or decrease the DNaseI

concentration in increments of 1 µg/mL within the range where ideal levels of DNaseI cutting are indicated on the basis of the resulting DNA viscosity. A typical protocol might include a series of four to six digests that are conveniently performed in groups of two. If DNaseI digestions are performed for 3 min, this protocol accommodates the analysis of three preparations of nuclei at a time, yielding a total of 12–18 DNA samples. In instances where the DNaseI is tested one concentration at a time, three digestions are usually sufficient to obtain good results. Useful concentrations of DNaseI to test are 1, 2, 3, 4, 5, 6, 8, 10, and 12 µg/mL DNaseI, where 1 µg corresponds to 2 U of DNaseI.

15. Nuclei should be stored on ice in the original 50 mL tube used for nuclei isolations so that the suspension is easily remixed by swirling each time an aliquot of nuclei is required.
16. As a starting point prepare 10X DNaseI stocks of either 40 µg/mL for one digest; 30 and 50 µg/mL for two digests; or 20, 40, and 60 µg/mL for three digests.
17. The protocol described here is designed for 500 µL aliquots of nuclei that will yield ~125 µg of purified DNA. The method works well for between 200 and 800 µL of nuclei, which requires a 0.01 vol of 100 m*M* $CaCl_2$ and a 0.1 vol of 10X DNaseI. The procedure is limited to six digests at a time as only six samples can be accommodated within 3 min equilibration and DNaseI digestion periods. Six digests take 8.5 min, allowing 30 s for each individual procedure. For speed of handling, leave the 10 mL tubes uncapped during the DNaseI digestion procedure.
18. It is not usually necessary to obtain a separate nondigested control sample from each series of digestions if the same species of DNA is involved in each series. If the number of cells available is limiting it may be preferable to purify the control DNA from a different source. It is important, however, to have one intact genomic DNA control sample included in each Southern blot analysis. In instances where allelic differences arise, it is necessary to have a control DNA sample from each individual source of cells or tissues.
19. If the initial digests appear suitable, then adjust the concentration of DNaseI used in subsequent digestions in increments of 1 µg/mL above, below and in between the initial concentrations used. If the DNA appears essentially undigested, then increase the concentration of DNaseI in increments of 2–4 µg/mL DNaseI until the ideal range is found. If the DNA is over-digested by 3–5 µg/mL, then try 1 and 2 µg/mL DNaseI.
20. DNA purification represents the most time-consuming part of the procedure, and the number of DNA samples purified can be decreased by screening the unpurified DNA electophoretically beforehand. I typically purify just two or three DNA samples from each series of six DNaseI digestions. As DNA samples stored in SDS/EDTA lysis buffer are very stable, they can also be accumulated over a period of weeks and purified in batches. Undigested samples should not be included in this screening procedure as the DNA chains are too large for electrophoretic analysis.
21. Poorly digested samples migrate as a band with an apparent size greater than 30 kb. In optimally digested samples, most of this high-molecular-weight DNA remains

at approximately 30 kb or greater, and a faint smear of DNA will trail below the main band. The appearance of this DNA will vary between digests, and reasonable results can be obtained in some cases where the smear is barely visible and in other cases where it trails down to about 2 kb. If the average size of the digested DNA is below 10 kb, then the digests are unsuitable for further analysis.

22. During phenol and chloroform extractions, small quantities of phenol and chloroform occasionally leak from the tubes. For safety reasons it is recommended that the tubes be sealed in a plastic bag while extracting on the rotating wheel.
23. Some of the DNA samples, especially the undigested samples, will be very viscous. To remove the supernatant cleanly hold the pipet at the surface of the supernatant and very slowly draw up the DNA from as close to the surface as possible. To avoid dragging up strings of DNA ensnared with debris, keep the pipet in motion rotating around the surface. Leave 5–10 % of the supernatant behind so as to avoid debris.
24. DNA samples can be stored at 4°C indefinitely. Over long periods, however, samples tend to evaporate and condense on the lids of the tubes. To reduce condensation, store the samples in a box with a foam insert lining the lid.
25. Choose a restriction enzyme that allows you to examine a segment of DNA that is at least several kilobases in length. Only one end of this fragment needs to have been mapped accurately and cloned. The restriction enzyme site needs to be at least 1 kb away from any DH sites being mapped. To locate reliably all the DH sites in a large locus, probes should be selected at intervals of approximately 10 kb. The utilization of two overlapping sets of restriction enzyme fragments will allow detection of DH sites located close to any one restriction enzyme site that might otherwise be overlooked.
26. Be sure to include at least one non-DNaseI-digested DNA sample as a control. As an additional control it may be advisable to also include a DNaseI digest of purified DNA to control for DNA sequences that are hypersensitive to DNaseI for reasons unrelated to chromatin structure. A DNaseI control can be prepared by digesting 0.1 mg/mL purified genomic DNA for 3 min at 22°C (room temperature) with approximately 0.01–0.025 µg/mL DNaseI in 1 m*M* $CaCl_2$, 10 m*M* MgAcetate, 50 m*M* KAcetate, 10 m*M* Tris-Acetate, pH 7.5 (or similar restriction enzyme buffer supplemented with 1 m*M* $CaCl_2$). It is recommended to perform a series of control DNaseI digestions of genomic DNA and to pool the appropriate digested DNA samples to give an even range of DNA sizes from about 1 to 15 kb.
27. For a 15 × 15 cm gel dissolve 1.2 g agarose in 135 mL water by boiling for about 2 min in a preweighed 500 mL flask. Cool the dissolved agarose to about 60–75°C, weigh the flask, and make the weight of the solution up to 135 g by adding more water. Add 15 mL of 10X TNAE electrophoresis buffer and 7.5 µL 10 mg/mL ethidium bromide and pour immediately into a tray with a 1-mm-thick comb that will form about 20 5-mm wide wells.
28. I have determined the DNA fixation conditions by calibrating a medium wavelength UV transilluminator (Model TM-20 UVP Inc., San Gabriel, CA). Other UV light systems may require a different fixation time for optimum hybridiza-

tion results. A common error is to expose the filter to UV for too long (e.g., 5 min), which reduces DNA hybridization efficiency. If a UV lightbox is not available, then fix by baking in a vacuum oven at 80° for 1 h. An alternative approach is to use alkaline transfer to a postively charged nylon membrane such as Zeta-Probe (BioRad, Hercules, CA) or HybondN+ (Amersham), in which case subsequent fixation is unnecessary. To perform an alkaline transfer, denature the DNA by washing the gel in 0.4 *M* NaOH and transfer directly using 0.4 *M* NaOH, as the transfer solution instead of 10X SSC.

29. If RapidHyb buffer is not available then use a Denhardt's reagent buffer system. This alternative method will require prehybridization for 6 h at 68°C in 15 mL of 10x Denhardt's reagent, 0.25 mg/mL boiled sonicated herring or salmon DNA, 0.5 % SDS, 5X SSC, 50 m*M* sodium phosphate, pH 7. Hybridization is then performed overnight at 68°C in 5 mL 5X Denhardt's reagent, 0.1 mg/mL boiled sonicated herring or salmon DNA, 0.5% SDS, 5X SSC, 50 m*M* sodium phosphate, pH 7 with 5 ng/mL of [^{32}P]-labeled probe. Denhardt's reagent is prepared as a 100X stock as 2% BSA, 2% Ficoll, and 2% polyvinylpyrollidone.
30. A spin column can be prepared by punching a hole in the top and bottom of a 1.5 mL Eppendorf tube, packing the bottom with glass wool, and filling the tube with a thick slurry of Sephadex G-50 in TE buffer. After excess buffer has drained out the tube is then seated over a second Eppendorf tube inside a larger diameter (e.g., 10 mL) centrifuge tube and spun for 2 min at 2000 rpm to expel most of the remaining buffer. To purify the labeled probe, load the labeling reaction mix onto the surface of the Sephadex and overlay with 50 µL of TE. Spin for 2 min at 2000 rpm and use the eluate in the collection tube as the probe.
31. Preparation of genomic DNA size markers requires complete digestion of 10 µg DNA with the same restriction enzyme used to digest the DNaseI-digested samples, followed by digestion for 30 min with 5 U of a second restriction enzyme that cuts close to the DH site.

References

1. Gross, D. S. and Garrard, W. T. (1988) Nuclease hypersensitive sites in chromatin. *Ann. Rev. Biochem.* **57,** 159–197.
2. Cockerill, P. N., Bert, A. G., Jenkins, F., Ryan, G. R., Shannon, M. F., and Vadas, M.A . (1995) Human granulocyte-macrophage colony-stimulating factor enhancer function is associated with cooperative interactions between AP-1 and NFATp/c. *Mol. Cell. Biol.* **15,** 2071–2079.
3. Siebenlist, U., Hennighausenen, L., Battey, J., and Leder, P. (1984) Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. *Cell* **37,** 381–391.

4

Analysis of In Vivo Methylation

Hans-Henrik M. Dahl and Wendy M. Hutchison

1. Introduction

The presence of methylated cytosines (as 5-methylcytosine, 5-MeC) in eukaryotic DNA was established nearly 50 years ago. Nevertheless, the function of methylated nucleotides in DNA has not yet been fully established. They have been proposed to play a role in regulation of gene expression, in genome imprinting, in X chromosome inactivation, in DNA repair mechanisms, tumorigenesis, and aging. A number of methods are available to detect 5-MeC in DNA and most are based on the use of methylation-sensitive restriction enzymes or genomic sequencing protocols. Although technically simple, the use of methylation-sensitive restriction enzymes has the disadvantage that only 5-MeCs that are part of the recognition sequences can be detected and analyzed. In addition, hemimethylated DNA cannot normally be detected. One genomic sequencing protocol ***(1)*** is based on the chemical sequencing method developed by Maxam and Gilbert ***(2)***, but takes advantage of the different reactivity of hydrazine toward C and 5-MeC. Recently, a simple and efficient genomic sequencing technology has been developed ***(3)***. Unlike other methods, this novel approach gives a positive display of 5-MeC residues in the DNA. It is based on the observation that sodium bisulfite, followed by alkaline treatment, converts cytosine residues in single-stranded DNA to uracil under conditions where 5-MeC is unreactive ***(4)***. This deamination process is outlined in **Fig. 1**. Following bisulfite/alkali treatment, the DNA is PCR amplified using strand-specific primers. Polymerase chain reaction (PCR) amplified products are electrophoresed on an agarose gel, and bands of the expected size are isolated. The methylation pattern is then determined by dideoxynucleotide sequencing using the PCR primers or specific internal primers. If the PCR

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

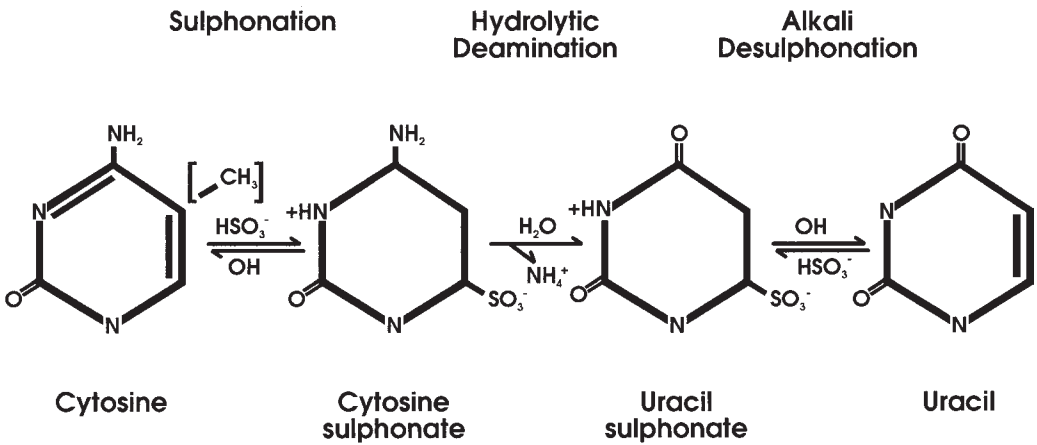


Fig. 1. The deamination process.

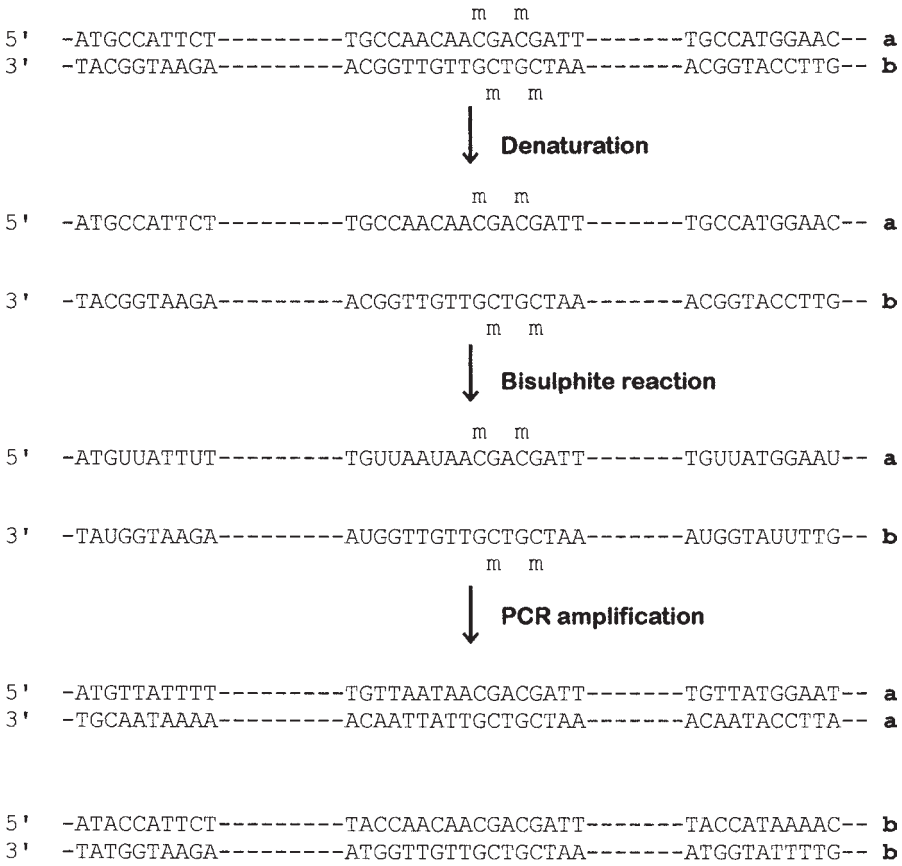


Fig. 2. The bifulfite genomic sequencing technique.

amplified product is sequenced directly, the result will reflect the overall distribution of 5-MeC in the strand. Alternatively, the PCR fragments can be cloned into M13 vectors and individual clones analyzed. These results will reflect the methylation status of single DNA molecules. The bisulfite genomic sequencing technique can be carried out on small quantities of genomic DNA. We normally use 2–10 µg, but there is no reason why smaller amounts of starting DNA cannot be used (*see* **Note 1** and ***ref. 5***).

The method is illustrated in **Fig. 2**. Genomic DNA is isolated from cells or tissues, fragmented, denatured, and treated with bisulfite and alkali prior to PCR amplification.

Special attention should be paid when designing the amplification primers, as the sequence of the bisulfite-treated DNA will be different from that of the untreated DNA (*see* **Note 2**). Furthermore, the two bisulfite treated DNA

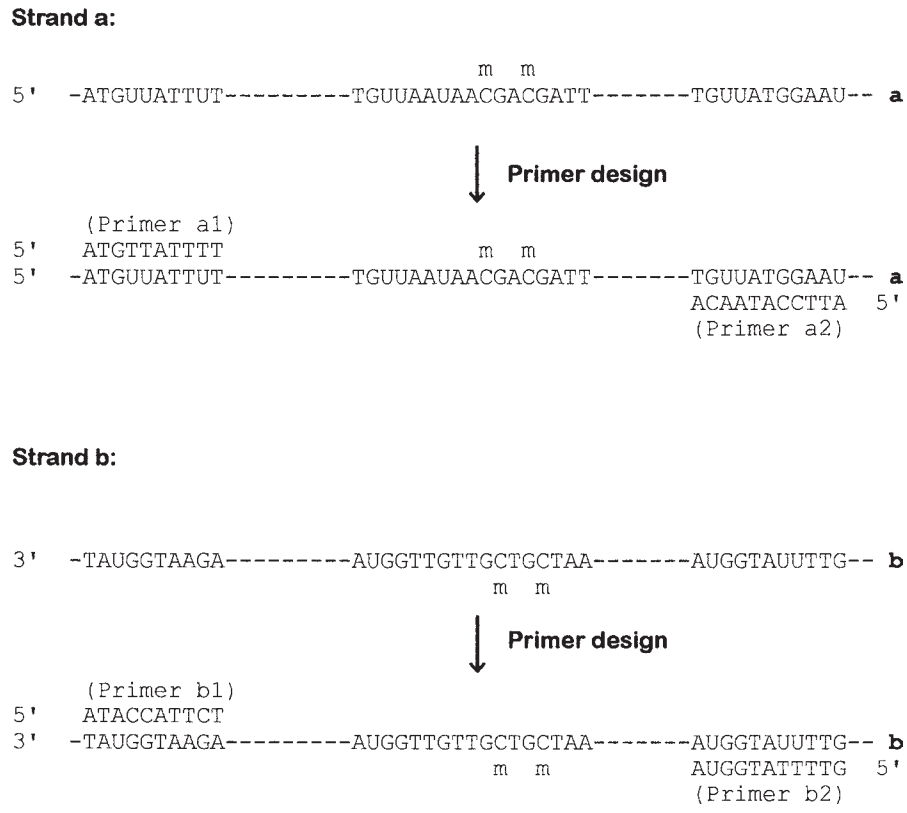


Fig. 3. Design of strand-specific oligonucleotide primers.

strands are no longer complementary, and amplification of the DNA, therefore, necessitates the use of strand-specific oligonucleotide primers (**Fig. 3**). As a consequence, each of the two DNA strands are amplified and analyzed individually. **Figure 4** illustrates the in vivo methylation analysis of a fragment, subcloned into M13.

Since the initial description of the bisulfite genomic sequencing technique, several modifications for detection of 5-MeC have been published. The starting material in these modifications is the bisulfite-treated DNA, which is then used in methylation-specific PCR ***(6)***, restriction enzyme analysis of amplified bisulfite-treated DNA ***(7,8)*** or in methylation-sensitive single nucleotide primer extension experiments ***(9)***.

2. Materials

High-quality deionized water is used for making up solutions. All chemicals for general use are AnalaR grade.

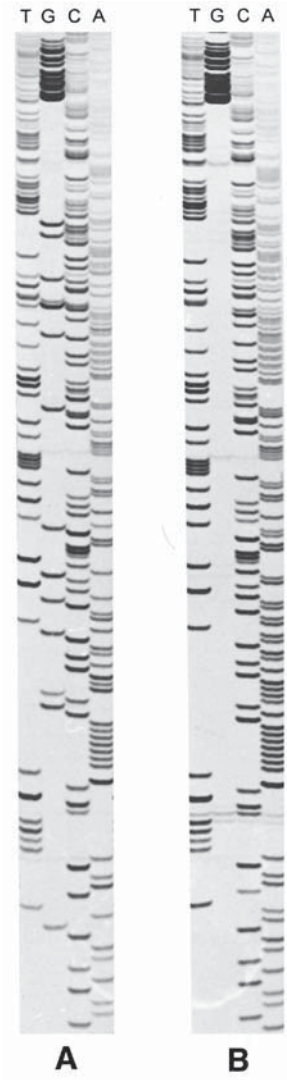


Fig. 4. DNA sequence analysis of a fragments subcloned into an M13 sequencing vector. In panel A, the cytosine residues that are part of CpG dinucleotides are methylated (presence of bands in the G track). In panel B, the cytosine residues that are part of CpG dinucleotides are not methylated (absence of bands in the G track). The cluster of bands in the G track at the top of both panels is because of the reading of sequences from the M13 vector.

2.1. Stock Solutions Stored at Room Temperature

1. Water: Autoclave.
2. 70% Ethanol.

3. 100% Ethanol.
4. 1 *M* Tris-HCl: Adjust the pH with conc. HCl.
5. 0.5 *M* EDTA: Adjust the pH of the acid form of EDTA with 10 *M* NaOH to 7.6 and autoclave.
6. 2 *M* NaOAc, pH 5.6: Filter sterilize.
7. 0.5 *M* NaOAc, pH 5.2: Filter sterilize.
8. 2 *M* NaOH: Make fresh.
9. 3 *M* NaOH: Make fresh.
10. 5 *M* NH_4OAc, pH 7.0.
11. DNA buffer: 10 m*M* Tris, pH 7.5, and 0.1 m*M* EDTA.
12. 10X TBE: 0.89 *M* Tris-base, 0.89 *M* boric acid, and 20 m*M* EDTA.
13. Gel fixing solution: 10% acetic acid, 10% methanol.

2.2. Stock Solutions Stored at 4°C

1. Phenol: Saturated with 0.1 *M* Tris-HCl, pH 8.0.
2. Phenol:chloroform:isoamylalcohol (25:24:1).
3. 10 m*M* hydroquinone (Quinol, BDH Australia). Dissolve 0.55 g in 50 mL water to make a 0.1 *M* stock solution. Store in light-tight container and make the 10 m*M* working solution fresh.
4. 3.1 *M* sodium bisulfite: Make fresh. Dissolve 6.45 g sodium bisulfite in approx 18 mL water. Adjust the pH to 5.0 with 10 *M* NaOH and make volume up to 20 mL.
5. 8% acrylamide-urea gel mix: 8% acrylamide:bisacrylamide (19:1), 7 *M* urea, 0.5 × TBE.
6. Dialysis tubing: Spectra/Por Industries (Los Angeles, CA) molecular porous membrane tubing no. 1 (molecular weight cutoff: 6–8000).
7. Dialysis buffer: 5 m*M* NaOAc; pH 5.2, 0.5 m*M* hydroquinone. Make 4 L using 40 mL 0.5 *M* NaOAc; pH 5.2 and 20 mL 0.1 *M* hydroquinone and cool to 4°C.
8. l. 5 m*M* NaOAc, pH 5.2.

2.3. Stock Solutions Stored at –20°C

1. Restriction enzyme buffers (Boehringer Mannheim, GmbH, Germany).
2. Restriction enzymes (Boehringer Mannheim).
3. 70% EtOH.
4. 100% EtOH.
5. Stop solution: 95% Formamide, 20 m*M* EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol.
6. Taq DNA polymerase (Boehringer Mannheim): 5 U/mL.
7. 10X dNTP mix: 2 m*M* of each of dATP, dCTP, dGTP and TTP.
8. 10X PCR buffer: 670 m*M* Tris-HCl pH 8.8; 166 m*M* $(NH_4)SO_4$; 4.5% Triton X-100; 2 mg/mL gelatin.
9. 25 m*M* $MgCl_2$.

2.4. Reagents/Special Equipment

1. 50°C oven.

2. 37°C heating block or oven.
3. NuSieve Agarose.
4. Gel running tank for agarose gel electrophoresis.
5. Power supply for agarose gel electrophoresis.
6. Gel running tank for acrylamide gel electrophoresis.
7. Power supply for DNA sequencing gel electrophoresis.
8. QIAEX II Gel extraction kit (Qiagen, Australia).
9. Sequenase version 2 DNA sequencing kit (USB).
10. [α-^{33}P]-dATP.
11. X-ray film: Kodak Biomax MR.
12. TEMED.
13. Ammonium persulphate.

3. Methods

3.1. Restriction Enzyme Digestion of Genomic DNA (*see* Note 3)

1. Digest 2–10 µg genomic DNA in 100 µL with a restriction enzyme that does not cut in the region that is to be analyzed (*see* **Note 3**); we often use EcoRI.
2. Add EDTA to a final concentration of 5 m*M* and extract once with an equal volume of phenol:chloroform:isoamylalcohol.
3. Precipitate the DNA with 10 µL 3 *M* NaOAc, pH 5.2, and 250 µL 100% EtOH at –70°C for 1 h.
4. Recover the DNA by centrifugation in a microcentrifuge (12,000*g*, 15 min). Wash pellet carefully with 70% EtOH, dry briefly under vacuum. Resuspend DNA in 100 µL H_2O.

3.2. Alkaline Denaturation and Neutralization (*see* Note 4)

1. Add 11 µL fresh 2 *M* NaOH to the 100 µL restriction enzyme digested genomic DNA. Vortex briefly.
2. Leave at room temperature for 5 min, then add 44 µL of 5 *M* NH_4OAc, pH 7.0. Vortex briefly and leave at room temperature for 5 min.
3. Add 450 µL of 100% EtOH (-20ºC). Vortex briefly, freeze on dry ice for 5 min, and pellet DNA by centrifugation in a microcentrifuge for 20 min (*see* **Note 4**).
4. Remove supernatant, rinse pellet with 70% EtOH (–20°C), and centrifuge for 20 min..
5. Remove supernatant and dry pellet. Resuspend DNA in 100 µL DNA buffer.

3.3. Bisulfite Reaction (*see* Notes 5 and 6)

1. To 100 µL denatured DNA add 60 µL of 10 m*M* hydroquinone. Vortex briefly and add 1040 µL of 3.1 *M* sodium bisulfite, pH 5.0. Mix well. (*see* **Note 5**).
2. Incubate at 50°C for 40 h (*see* **Note 6**).

3.4. Dialysis (*see* Note 7)

1. Make up the dialysis buffers and cool to 4°C (for example, by leaving them in a cold room overnight).
2. Place the bisulfite-treated DNA (approx 1200 µL) in a dialysis bag.

3. Dialyze against 2 L of dialysis buffer at 4°C for > 1 h, replace the dialysis buffer, and repeat dialysis for > 1 h.
4. Dialyze against 2 L of 5 m*M* NaOAc, pH 5.2, at 4ºC for > 1 h, replace the dialysis buffer, and repeat the dialysis for > 1 h.
5. Dialyse against 2 L H_2O for > 1 h.
6. Transfer the DNA to a 5-mL tube and lyophilize sample.

3.5. Alkali Treatment/Deamination

1. Resuspend pellet in 100 µL DNA buffer.
2. Add 11 µL fresh 3*M* NaOH. Vortex briefly and leave 5 min at room temperature.
3. Add 66 µL of 5 *M* NH_4OAc, pH 7.0. Vortex briefly.
4. Add 450 µL of 100% EtOH (-20°C). Freeze on dry ice for 5 min and pellet DNA by centrifugation in a microcentrifuge for 20 min.
5. Remove supernatant, rinse pellet with 70% EtOH (–20°C) and centrifuge for 20 min.
6. Remove supernatant and dry pellet under vacuum. Resuspend DNA in 100 µL H_2O.

3.6. PCR Amplification (*see* Notes 8–9)

1. Use 1–2 µL of the DNA solution in a 100-µL PCR amplification.
2. In addition to the DNA, each PCR contains 500 ng of each of the two PCR primers (*see* **Notes 8** and **9**), 200 µ*M* each of the four dNTPs (dATP, dCTP, dGTP and TTP), 1X PCR buffer, and 1–2 U Taq polymerase.
3. The PCR reaction conditions are usually 40 cycles consisting of 2 min at 94°C, 2 min at 56°C and 2 min at 74°C (*see* **Note 8**).
4. Ten microliters of the PCR amplified DNA is mixed with 2 µL gel loading buffer and run on a 3% NuSieve agarose gel. DNA markers are also included on the gel in order to identify the DNA fragment of the expected size. If the expected product is not visible or the yield is very low, a nested PCR amplification is recommended (*see* **Note 8**).
5. The remainder (90 µL) of the PCR reaction is precipitated with 9 µL 2 *M* NaOAc, pH 5.6, and 230 µL 100% EtOH. The pellet is resuspended in 40 µL water and 8 µL gel loading buffer, and run on a preparative 3% NuSieve agarose gel (10 mm × 1 mm wells). The desired DNA fragment is excised from the gel and purified using a QIAEX II Gel extraction kit.

3.7. DNA Sequencing (*see* Note 10)

1. Approximately 500 ng–1 mg of the isolated PCR fragment is sequenced using a Sequenase version 2 DNA sequencing kit (USB) and [α-^{33}P]-dATP. The sequencing primers are either the same as those used in the PCR reaction or can be internal primers made to correspond to a section of the DNA fragment unlikely to contain 5-MeC groups (*see* **Note 2**). If the PCR product has been subcloned into an M13 vector, M13 sequencing primers can be used.
2. Make an 8% denaturing acrylamide-urea gel (0.4 mm thick with 20 × 50 cm glass plates by mixing 50 mL 8% acrylamide-urea gel mix with 25 µL TEMED and 400 µL fresh 10% ammonium persulfate. Allow the gel to set for at least 2 h.

3. Preheat the gel to approx 50°C. Running buffer is 0.5 × TBE. The gel is run at 55 W.
4. Denature samples at 94°C for 3–5 min. and immediately load 2.5 µL on the gel.
5. Run the gel at 55 W until the bromophenol blue has run off the gel (*see* **Note 11**).
6. Fix the gel in gel fixing solution (10% methanol, 10% acetic acid) for 15 min, transfer gel onto blotting paper and dry the gel in an oven or use a gel drier.
7. Expose the gel directly to X-ray film at room temperature for 1–4 d.

4. Notes

1. We normally use 2–10 µg DNA as starting material. However, the procedure can be scaled down and applied to less genomic DNA. When less genomic DNA is used, we add 2 µg *Escherichia coli* tRNA as carrier. Several publications have described the successful analysis of small amounts of DNA ***(10–12)***.
2. Careful consideration must be given to primer design. Primers with a length of approx 30 nucleotides seem to work well. As a consequence of the bisulfite reaction, the sequence of the two strands will change and, in addition, they will no longer be complementary. It is helpful to draw the sequences around the area selected for priming for each of the two strands (*see* **Fig. 3**). The C residues can then be changed to As, the complementary strands deduced, and primers designed. An additional complication is that 5-MeCs are not converted to uracil residues. Because 5-MeCs normally are part of a CpG dinucleotide sequence, we avoid making primers to a sequence containing a CpG dinucleotide, as we do not know the status of the C residue.
3. The best results are obtained if the DNA is fragmented before denaturation and bisulfite treatment. This can be done by use of a suitable restriction enzyme or by shearing the genomic DNA.
4. It is important that the DNA is single-stranded during bisulfite treatment. An improved protocol, therefore, recommends that the DNA is alkaline-denatured directly prior to bisulfite treatment, that the initial DNA concentration is decreased (to 2 µg/100 µL), and the bisulfite treatment is at 0°C for 24 h ***(13)***. Alternatively, the bisulfite reaction temperature can be raised to 95°C ***(14)***. Another protocol recommends embedding cells or DNA in low-melting-point agarose beads ***(12)***. DNA isolation, restriction enzyme digestion, denaturation, bisulfite treatment and PCR amplification can be done in the beads, with minimal loss of DNA.
5. The bisulfite concentration can range from 3.0 to 4.0 *M* (pH 4.8–5.8), but a higher concentration might be more advantageous ***(13)***. The bisulfite solution is sensitive to light and must therefore be kept in the dark. Suppliers of bisulfite (sodium metabisulfite) include Merck (Australia) and Sigma Chemicals (St. Louis, MO).
6. We found that a 40-h incubation at 50°C works well. However, other studies have recommended incubation times of 8–16 h ***(10)***.
7. The dialysis step can be replaced by other methods that allow a change of buffer and concentration of the DNA. These methods include purification using glass

milk (Gene-clean II) ***(15)***, Spinclean columns (Progen Industries, Australia) ***(16)*** or Promega Magic DNA Clean-Up system columns ***(10)***.

8. It is not unusual to experience problems with the PCR, resulting in no fragment or a multitude of fragments being amplified. Careful optimization of the PCR conditions, especially of the annealing temperature and $MgCl_2$ concentration, may be necessary. The use of nested primers has also proven very useful in generating sufficient quantities of a pure DNA fragment ***(10,17)***.
9. Incorporation of restriction enzyme cleavage sites into the PCR primers simplifies subcloning of the bisulfite-treated fragment.
10. Manual dideoxynucleotide sequencing using the Sequenase kit is only one of several methods available for DNA sequence analysis. Other protocols, commercial kits, and methods, including automated sequencing using fluorescent labels ***(18)***, give good results.
11. For reading longer DNA fragments, run a 6% acrylamide:urea gel for a longer time (until xylene cyanol has run off the gel).
12. The original protocol recommended the use of control plasmids ***(3)***. In most cases, we do not include control plasmids in our experiments. This is because internal controls are a consequence of the bisulfite method: all cytosines that are not part of a CpG (or in rare cases CpNpG ***[19]***) should be converted to uracils. Working with control plasmids also increases the risk of contamination of the PCR reactions.

References

1. Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B., and Riggs, A. D. (1989) Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246**, 810–813.
2. Maxam, A. M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
3. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **89**, 1827–1831.
4. Shapiro, R., Braverman, B., Louis, J. B., and Servis, R. E. (1973) Nucleic acid reactivity and conformation. II. Reaction of cytosine and uracil with sodium bisulfite. *J. Biol. Chem.* **248**, 4060–4064.
5. Grigg, G. W. (1996) Sequencing 5-methylcytosine residues by the bisulfite method. *DNA Seq.* **6**, 189–198.
6. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D., and Baylin, S. B. (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. USA* **93**, 9821–9826.
7. Sadri, R. and Hornsby, P. J. (1996) Rapid analysis of DNA methylation using new restriction enzyme sites created by bisulfite modification. *Nucleic Acids Res.* **24**, 5058–5059.
8. Xiong, Z. and Laird, P. W. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* **25**, 2532–2534.

9. Gonzalgo, M. L. and Jones, P. A. (1997) Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* **25**, 2529–2531.
10. Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997.
11. Raizis, A. M., Schmitt, F., and Jost, J. P. (1995) A bisulfite method of 5-methylcytosine mapping that minimizes template degradation. *Anal. Biochem.* **226**, 161–166.
12. Olek, A., Oswald, J., and Walter, J. (1996) A modified and improved method for bisulfite based cytosine methylation analysis. *Nucleic Acids Res.* **24**, 5064–5066.
13. Feil, R., Charlton, J., Bird, A. P., Walter, J., and Reik, W. (1994) Methylation analysis on individual chromosomes: improved protocol for bisulfite genomic sequencing. *Nucleic Acids Res.* **22**, 695–696.
14. Tasheva, E. S. and Roufa, D. J. (1994) Densely methylated DNA islands in mammalian chromosomal replication origins. *Mol Cell Biol.* **14**, 5636–5644.
15. Reeben, M. and Prydz, H. (1994) An improved method for detection of 5-methylcytosine by PCR-based genomic sequencing. *BioTechniques* **16**, 416–417.
16. Garrick, D., Sutherland, H., Robertson, G., and Whitelaw, E. (1996) Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. *Nucleic Acids Res.* **24**, 4902–4909.
17. Iannello, R. C., Young, J., Sumarsono, S., Tymms, M. J., Dahl, H. H., Gould, J., Hedger, M., and Kola, I. (1997) Regulation of Pdha-2 expression is mediated by proximal promoter sequences and CpG methylation. *Mol. Cell Biol.* **17**, 612–619.
18. Myohanen, S., Livingston, D. C., Wahlfors, J., and Juhani, J. (1995) Automated fluorescent genomic sequencing as applied to the methylation analysis of the human ornithine decarboxylase gene. *DNA Seq.* **5**, 1–8.
19. Clark, S. J., Harrison, J., and Frommer, M. (1995) CpNpG methylation in mammalian cells. *Nat. Genet.* **10**, 20–27.

5

Detection of Transcription Factor Partners with a Yeast One Hybrid Screen

Michael Sieweke

1. Introduction

Transcription factors usually engage in multiple regulatory interactions with other proteins in their normal promoter or enhancer context. Conventional yeast two hybrid screens, in which a putative protein interaction domain is fused to a heterologous DNA binding domain like that of LexA or Gal4 ***(1–3)*** have been used extensively to detect protein/protein interactions. Even though the technique has also identified partners of transcription factors (for example ***[3–5]***), it has particular limitations in this application. This is because many transcription factors undergo conformational changes upon DNA binding and frequently engage in interactions with other factors through surfaces on their DNA-binding domains. When target factors are fused to the heterologous Gal4 or LexA DNA binding domains and exposed in their non-DNA-bound form, relevant interaction epitopes may be shielded or inappropriately folded. By contrast, in a situation where the transcription factor is bound to its authentic binding site on the DNA via its own DNA binding domain, its conformation should closely resemble that found in a normal promoter context and thus detect functionally relevant interactions with other factors. This concept has been instrumental in identifying coactivators of transcription factors that are transcriptionally inert in yeast ***(6–8)***. The majority of transcription factors, however, have transactivation potential in yeast. Therefore the author has designed a generally applicable yeast one hybrid screen that uses a transcriptionally disabled target protein to take advantage of the normal conformation of a DNA bound transcription factor (*see* **Fig. 1**). The utility of the approach is exemplified by the detection of cofactors for the Ets-1 protein ***(9,10)*** which changes conformation upon DNA binding ***(11,12)*** and is transcriptionally active in yeast ***(10)***. Tran-

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

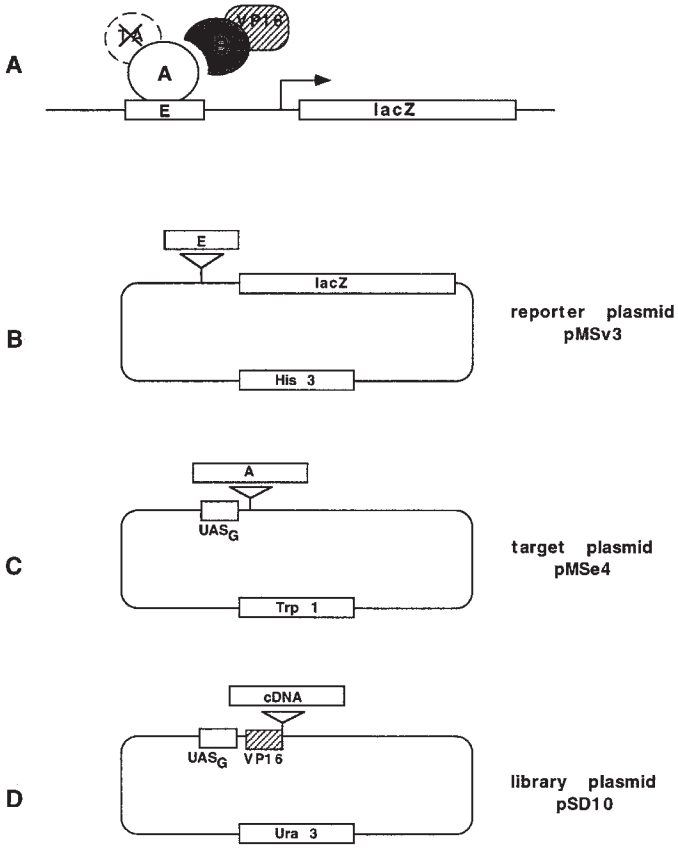


Fig. 1. **(A)** Principle of one hybrid screen with a transcriptionally disabled target transcription factor (A) interacting with VP16 tagged protein (B) encoded by a cDNA library. **(B)** Reporter plasmid pMSv3 with cloning site for binding site or enhancer element insertion (E). **(C)** Yeast expression plasmid pMSe4 for galactose inducible expression of the target transcription factor (A). **(D)** Example of a library plasmid (pSD.10) for galactose inducible VP16 tagged expression of a cDNA library.

scriptional inactivation can be achieved by deletion or mutation of the transactivation domain(s). To minimize potentially toxic or growth inhibitory effects, the target transcription factor is expressed from a galactose inducible promoter. As reporter, the ß-galactosidase gene on a high copy plasmid is used, which results in a higher sensitivity than integrated or auxotroph reporters. It can accommodate either multimerized binding sites for the target factor or, where available, authentic promoter regions. The reporter strength and background levels can be controlled by the choice of the copy number of the DNA

element. The yeast strain, expression, and reporter plasmids used are compatible with a plasmid library having either a URA3 or LEU2 auxotroph marker, which includes many commercially available yeast plasmid libraries.

In **Subheading 3.1.**, the author describes methods of how to generate a suitable reporter strain, which involves the construction of a reporter plasmid (*see* **Subheading 3.1.1.**) and a target expression plasmid (*see* **Subheading 3.1.2.**), as well as the transformation of these plasmids into yeast and the characterization of the background activity and sensitivity of the generated reporter strain (*see* **Subheading 3.1.3.**). After the transformation of this strain, with a suitable library and identification of positive clones (*see* **Subheading 3.2.**), false positives have to be eliminated (*see* **Subheading 3.3.**). The isolation and characterization of the cDNAs from the remaining positives is explained in **Subheading 3.4.** and finally **Subheading 3.5.** discusses the analysis of the interaction properties of the isolated clones with a quantitative colorimetric assay.

2. Materials

2.1. Strains and Growth Media

1. *Saccharomyces cerevisiae* strain W303-1A: *Mata, ho, his3-11,15; trp1-1; ade2-1; leu2-3,112; ura3; can1–100.*
2. *Escherichia coli* strain Xl-1 blue: *recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac[F' proAB lac*I^q*ZDM15 Tn10 (tetR)].*
3. YPD medium: (*see* **Note 1**): 1% (w/v) Bacto-yeast extract, 2% (w/v) bacto-peptone, 2% (w/v) dextrose in distilled water. Autoclave.
4. SD-His, -Trp medium: 0.67% yeast nitrogen base w/o amino acids (*see* **Note 2**), 2% dextrose, and an amino acid mix containing the constituents in **Table 1** but lacking histidine, and tryptophane (*see* **Note 3**). Autoclave.
5. SD-His, -Trp, -Ura medium: 0.67% yeast nitrogen base w/o amino acids (*see* **Note 2**), 2% dextrose, and an amino acid mix containing the constituents in **Table 1** but lacking histidine, tryptophane, and uracil (*see* **Notes 3** and **4**). Autoclave.
6. Sgal-His, -Trp, -Ura medium: 0.67% yeast nitrogen base w/o amino acids (*see* **Note 2**), 2% galactose and an amino acid mix containing the constituents in table 1 but lacking histidine, tryptophane and uracil (*see* **Notes 3** and **4**). Autoclave.
7. YPD plates: YPD medium containing 2% bacto-agar (w/v). Autoclave.
8. SD -His, -Trp -Ura 20cm × 20cm plates: SD -His, -Trp, -Ura (*see* **Note 4**) medium containing 2% (w/v) bacto-agar. Autoclave. Pour 250 mL per plate.
9. Sgal-His, -Trp -Ura + x-gal 20 cm x 20 cm plates: Sgal-His, -Trp -Ura (*see* **Note 4**) medium with 2.5%(w/v) bacto-agar (*see* **Note 5**). Autoclave and add under sterile conditions: 10% $NaPO_4$ buffer pH 7 (*see* **Note 6**) and 0.02% x-gal stock solution. Pour 250 mL per plate.
10. SD-His, -Trp, -Ura plates: SD-His, -Trp medium containing 2% (w/v) bacto-agar.

Table 1
Concentrations of Additives for Synthetic Complete Media (SC)

Constituent	Final Concentration (mg/L)
Adenin sulfate	20
Uracil	20
L-Tryptophan	20
L-Histidine-HCl	20
L-Arginine	20
L-Methionine	30
L-Tyrosine	30
L-Leucine	30
L-Isoleucine	30
L-Lysine-HCl	30
L-Phenylalanine	50
L-Glutamic acid	100
L-Aspartic acid	100
L-Valine	150
L-Threonine	200
L-Serine	375

11. SD-His, -Trp, -Ura + x-gal plates: SD-His, -Trp, -Ura (*see* **Note 4**) medium with 2.5% (w/v) bacto-agar (*see* **Note 5**). Autoclave and add under sterile conditions: 10% $NaPO_4$ buffer pH 7 (*see* **Note 6**) and 0.02% x-gal stock solution.
12. Sgal-His, -Trp, -Ura + x-gal plates: Sgal-His, -Trp, -Ura (*see* **Note 4**) medium with 2.5% (w/v) bacto-agar (*see* **Note 5**). Autoclave and add under sterile conditions: 10% $NaPO_4$ buffer pH 7 (*see* **Note 6**) and 0.02% x-gal stock solution.
13. SD-His, -Trp plates: SD-His, -Trp medium with 2% (w/v) bacto-agar.
14. SD-His, -Ura plates: SD-His, -Ura (*see* **Note 4**) medium with 2% (w/v) bacto-agar.
15. Sgal-His, -Trp, +x-gal plates: Sgal-His, -Trp medium with 2.5% (w/v) bacto-agar (*see* **Note 5**). Autoclave and add under sterile conditions: 10% $NaPO_4$ buffer pH 7 (*see* **Note 6**) and 0.02% x-gal stock solution.
16. Sgal-His, -Ura + x-gal plates: Sgal-His, -Ura (*see* **Note 4**) medium with 2.5% (*see* **Note 5**) (w/v) bacto-agar. Autoclave and add under sterile conditions: 10% $NaPO_4$ buffer pH 7 (*see* **Note 6**) and 0.02% x-gal stock solution.
17. LB broth: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, and 1.0% (w/v) NaCl. Autoclave.
18. LBamp plates: LB broth containing 1.5% (w/v) bacto-agar. Autoclave, cool to 55°C, and then add ampicillin to give a final concentration of 100 mg/mL.

2.2. Stock Solutions

1. 0.5 *M* EDTA. Autoclave.
2. 1 *M* Tris-Cl, pH 8.0. Autoclave.

3. TE: 10 m*M* Tris-HCl, pH 8.0; 1 m*M* EDTA. Autoclave.
4. 5 *M* NaCl. Autoclave.
5. 5 m*M* ATP. Prepare in sterile distilled H_2O. Store aliquots at –20°C.
6. 6X loading buffer: 0.25% bromphenolblue, 0.25% xylene xyanol, 30% glycerol. Store at 4°C.
7. 20XTBE: 1.8 *M* Tris-borate, 40 m*M* EDTA.
8. Polyacrylamide stock solution: 30% acrylamide, 08% *bis*-acrylamide in ddH_2O. Degas. Store at 4°C.
9. 10% ammonium persulfate in ddH_2O. Store in aliquots at –20°C.
10. 10 mg/mL ethidium bromide in H_2O. Store in the dark.
11. NaAc: 3 *M* NaAcetate pH 5.2. Autoclave.
12. Glycogen carrier: 20 mg/mL in sterile distilled H_2O. Store aliquots at –20°C.
13. 10% SDS: 10% sodium dodecyl sulfate in sterile distilled H_2O. Adjust pH to 7.2.
14. 10 *M* NAOH.
15. 20X SSC: 3 *M* NaCl, 3 *M* Na-citrate pH 7.0. Autoclave.
16. dNTPs (Roche Diagnostics, Indianpolis, IN): store aliquots of 1 m*M* solution in H_2O at –20°C.
17. TEN: 100 m*M* NaCl, 10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA.
18. Hybridization buffer: 6X SSC, 10 m*M* $NaPO_4$ pH 6.8, 0.5% SDS, 100 µg/mL sonicated salmon sperm DNA, 5X Denhardt's solution.
19. 50X Denhardt's solution: 1% Ficoll (type 400, Amersham Pharmacia Biotech, Uppsala, Sweden), 1% polyvinylpyrrolidone, 1% bovine serum albumin (fraction V, Sigma, St. Louis, MO). Filter and store in aliquots at –20°C.
20. Phenol: saturate with 10 m*M* Tris-Cl, pH 8.0.
21. Phenol/chloroform/isoamylalcohol (25:24:1): Saturate with 10 m*M* Tris-Cl, pH 8.0.
22. TE/LiAc: 0.1 *M* Li-acetate in TE.
23. Carrier DNA: 10 mg/mL salmon sperm DNA (Sigma) in sterile distilled H_2O. Dissolve for several hours at room temperature, sonicate, boil for 10 min, and store in aliquots at –20°C (*see* **Note 19**).
24. TE/LiAc/PEG: 40% Polyethylenglycol (MW4000) and 0.1 *M* LiAc in TE.
25. x-gal stock solution: 20 mg/mL in DMSO.
26. 1 *M* $NaPO_4$, pH 7.0: Mix 577 mL 1 *M* Na_2HPO_4, and 423 mL 1 *M* NaH_2PO_4. Adjust pH to 7.0. Autoclave.
27. 1 *M* $NaPO_4$, pH 6.8: Mix 463 mL 1 *M* Na_2HPO_4, and 537 mL 1 *M* NaH_2PO_4. Adjust pH to 6.8. Autoclave.
28. TENS: 100 m*M* NaCl, 10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA, 0.1% sodium dodecyl sulfate (SDS).
29. ß-gal lysis buffer: 0.1 *M* Tris-Cl (pH 7.5), 0.05% (v/v) Triton X-100
30. Z-buffer: 100 m*M* $NaPO_4$ (33 mL of 1 *M* NaH_2PO_4 and 66 mL of Na_2HPO_4 per liter), 10 m*M* KCl, 1 m*M* $MgSO_4$, adjust pH to 7.0. Autoclave. Add 2.7 µL β-mercaptoethanol/1 mL of buffer just prior to use.
31. ONPG stock solution: ortho-nitro-phenyl-galactoside (Sigma); 4 mg/mL in sterile, distilled H_2O. Store aliquots at –20°C.
32. 1 *M* Na_2CO_3: Prepare in sterile, destilled H_2O.

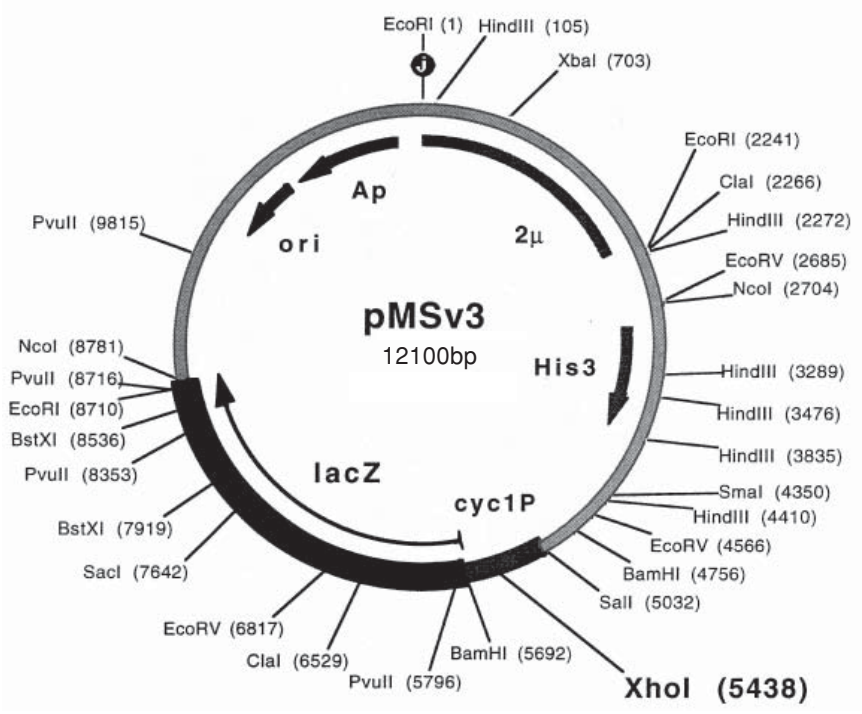


Fig. 2. pMSv3 reporter plasmid for insertion of a binding site or enhancer element into the XhoI site of the cyc1 promoter driving the lacZ gene.

2.3. Reagents/Special Materials

1. Yeast reporter plasmid pMSv3 (*see* **Note 7** and **Fig. 2**).
2. Sequencing primer for pMSv3: 5'-ATGTGTCAGCACTAAAGTTG-3' (non-coding strand reading back into XhoI site).
3. Yeast galactose inducible expression plasmid pMSe4 (*see* **Note 8** and **Fig. 3**).
4. Sequencing primer for pMSe4: 5'-GACACGCAAACACAAATAC-3'.
5. Yeast expression plasmid library in pSD.10 ***(6,10)*** or other compatible library (*see* **Note 9**).
6. Double-stranded sequencing kit (Amersham Pharmacia Biotech).
7. T4 DNA ligase and 10X ligase reaction buffer (Promega, Madison, WI): 2U/μL. Store at –20°C.
8. T4 polynucleotide kinase and 10X PNK reaction buffer (NEB, Beverly, MA): 10 U/μL Store at –20°C.
9. XhoI restriction endonuclease and 10X reaction buffer (NEB): 10U/μL. Store at –20°C.
10. SalI restriction endonuclease and 10X reaction buffer (NEB): 10U/μL. Store at –20°C.
11. Alkaline phosphatase and 10X reaction buffer (Roche Diagnostics): 1 U/μL. Store at 4°C.

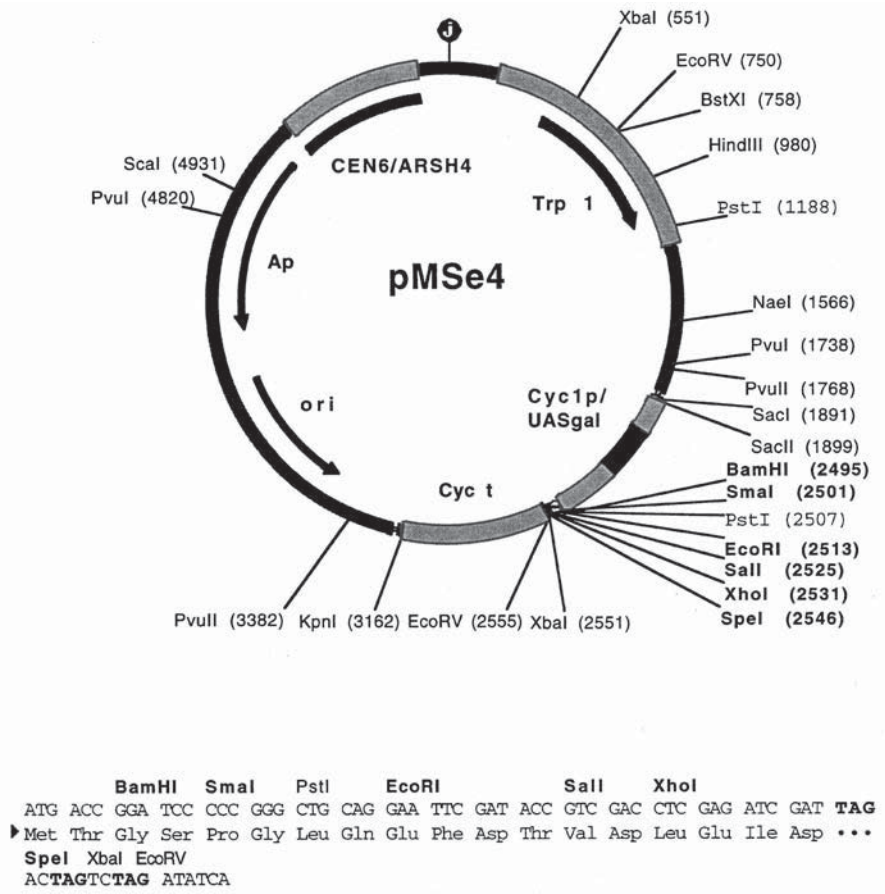


Fig. 3. pMSe4 galactose inducible expression plasmid for expression of target transcription factor. The reading frame of the poly linker is indicated.

12. Klenow fragment of *E.coli* DNA polymerase I (Roche Diagnostics): 1 U/µL. Store at –20°C.
13. α-[^{32}P]-dCTP (Dupont, Wilmington, DE): 3000 Ci/mmol; 10 mCi/mL.
14. Sephadex-G50 (Amersham Pharmacia Biotech): store in TEN buffer at 4°C.
15. X-ray film X-OMAT AR (Kodak, New Haven, CT).
16. DNA plasmid prep kit (Qiagen, Valencia, CA).
17. Gel extraction kit (Qiagen).
18. Molecular biology grade agarose (Sigma, St. Louis, MO).
19. Nitrocellose filter membranes, 20 cm × 20 cm and 10 cm diameter (Schleich & Schueler).
20. Sterile glass beads (~0.45 mm diameter), acid washed (Sigma).
21. TEMED: N,N,N,N-Tetramethylethylendiamine.

22. DMSO (dimethyl sulfoxide).
23. Chloroform.
24. Isoamylalcohol.
25. Ethanol: AR-grade ethanol (>98%).
26. 70% Ethanol: AR-grade ethanol diluted to 70% (v/v) with ddH_2O.

2.4. Equipment

1. 30°C incubator.
2. 37°C incubator.
3. 30°C shaker.
4. 37°C shaker.
5. Spectrophotometer equipped with 420 nm and 600 nm filters.
6. Low-speed centrifuge.
7. Microcentrifuge.
8. Hybridization oven.
9. Vacuum oven.
10. Waterbath with adjustable temperature.
11. Speedvac.

3. Methods

3.1. Generation of Reporter Strain

3.1.1. Construction of Reporter Plasmid

3.1.1.1. Cloning Protocol

1. Design oligonucleotides for the + and – strand of the consensus binding site or a promoter region containing a binding site for your transcription factor. Include recognition sequences for XhoI and SalI at the 5' and 3' end, respectively (*see* **Notes 10** and **11**). Design the equivalent oligonucleotides with mutations in the recognition sequence that abolish DNA binding of the transcription factor. Follow the protocol in parallel for wild type and mutated oligonucleotides.
2. Adjust oligonucleotide concentration to 5 mg/mL. Pipet together 100 µL + strand oligo, 100 µL-strand oligo, 2.5 µL 5 *M* NaCl, 47.5 µL TE. Place in metal heating block at 80°C for 5 min and let cool slowly to room temperature.
3. Pipet together 10 µL double-stranded oligonucleotide (20 µg), 2 µL ddH_2O, 2 µL 10X linker kinase buffer, 4 µL 5 m*M* ATP, 2 µL T4 polynucleotidekinase. Incubate for 1 h at 37°C (*see* **Note 12**).
4. Add 44 µL ddH_2O, 8 µL 10X ligase buffer (including ATP), 2 µL XhoI, 2 µL SalI, and 4 µL of T4 DNA ligase. Incubate for 2 h at 37°C (*see* **Note 10**).
5. Add 16 µL 6X loading buffer and split the sample into four wells on a nondenaturing 15% TBE polyacrylamide gel (12% if the oligonucleotide is longer than 30 nucleotides) ***(13)*** and run at 1–8 V/cm.
6. Stain the gel on the glass plate in a TBE/ethidium bromide bath with the fluid just barely covering the gel and *gently* agitate for 30 min. Lift gel out of the bath with

the glass plate, cover with plastic film, carefully invert onto UV light box, and cut out bands corresponding to 2, 3, and 5 copies of the oligonucleotide.

7. Place gel slice in an Eppendorf tube. Add 1 mL TE and incubate overnight at 37°C or for 3 h at 55°C. Remove TE and split into two tubes. Add 0.5 µL glycogen, 50 µL NaAc, and 1 mL ethanol, precipitate for 30 min at –20°C, spin for 15 min at 4°C, wash with 70% ethanol, dry pellet in a Speedvac, resuspend in 10 µL TE, and pool identical samples.
8. Cut 10 µg of pMSv3 with XhoI in a volume of 50 µL.
9. Add 1 µL (1U) of alkaline phosphatase and 10 µL of the 10X reaction buffer. Fill up to 100 µL and incubate for 30 min at 37°C.
10. Add 10 µL of 0.5 *M* EDTA. Heat at 68°C for 15 min to inactivate alkaline phosphatase.
11. Extract twice with an equal volume of phenol/chloroform/isoamylalcohol (25:24:1) and once with chloroform using 2 min spins in a microcentrifuge to separate phases.
12. Precipitate the DNA with 10 µL of NaAc and 250 µL ethanol at –20°C for 15 min. Recover the DNA by centrifugation for 15 min at 4°C. Wash with 70% ethanol and dry under vacuum.
13. Resuspend vector DNA in 25 µL of H_20. Set up ligations with 1 µL of vector DNA and, an estimated 10-fold molar excess of multimerized oligonucleotide (try a series of 0.5, 1, and 2 µL of recovered multimerized oligonucleotides), 2 µL 10X ligation buffer and 1 µL T4 DNA ligase in a total volume of 20 µL. Incubate at 16°C for 4 h.
14. Transform 100 µL of competent Xl-1blue *E. coli* with 2 µL of ligation reaction. Plate 100 µL aliquots on LBamp plates. Incubate overnight at 37°C.

3.1.1.2. Identification of Positive Clones (*see* **Note 13**)

1. Place nitrocellulose filter onto bacterial plate and mark orientation.
2. Lift off and place for 5 min onto Whatman 3MM paper soaked with 5% SDS (colony side up).
3. Place for 5 min onto 3MM paper soaked with 0.5 *M* NaOH/1.5 *M* NaCl (colony side up). Dry briefly on dry 3MM paper.
4. Let float for 5 min on the surface of a 0.5 *M* Tris pH 8.0/1.5 *M* NaCl solution. Submerge in the solution for another 5 min.
5. Dry upper and lower sides of the filter by placing it between two filter papers and pressing them together with a large pipet.
6. Wash in 6X SSC, dry on filter paper and leave in vacuum oven at 80°C for 1 h.
7. Calculate the melting temperature of the oligo. If it is shorter than 18 bp with the simplified formula $Tm = (A/T \times 2) + (G/C \times 4)$ or if it is longer with the formula $Tm = 81.5–16.6(\log[Na^+]) + 0.41(\%G/C) – 600/N$, where N is the number of nucleotides. The Na^+ concentration in 6X SSC is 1 *M*. The second term in the equation thus becomes zero.
8. Label oligonucleotide by Klenow fill in: Pipet together 100 ng of double-stranded oligonucleotide, 1 µL each of 1 m*M* solutions of dATP, dGTP, and dTTP, 50 µCi

α-[^{32}P]-dCTP and 1 µL Klenow fragment in a total volume of 20 µL. Incubate for 15 min at room temperature. Stop reaction by heating for 5 min at 70°C. Remove unincorporated dNTPs with a Sephadex G-50 spin column ***(13)***.

9. Wet filters in 6X SSC and prehybridize for 2 h at 10°C below the T_m of your probe in hybridization buffer, add probe at (2–5) × 10^6 cpm/mL, and hybridize 3–4 h at the same temperature.
10. Wash in 6X SSC, 2 × 5 min at room temperature and 2 × 5 min at the hybridization temperature.
11. Expose the filters to X-ray film for from 6 h to overnight and align the films with the plates to pick positive colonies. Grow bacterial cultures for DNA plasmid preps from the positive colonies and sequence the insert using the sequencing primer for pMSv3.

3.1.2. Construction of Target Expression Plasmid

1. Select restriction sites useful for cloning into the poly cloning site of pMSe4 (*see* **Notes 8** and **14,** and **Fig. 3)** that cut out a fragment from the target transcription factor cDNA, which lacks the transactivation domain(s) (*see* **Note 15**). In parallel identify useful restriction sites releasing the full-length cDNA of the transcription factor (*see* **Note 16**). If no sites are available, use a PCR strategy to generate the appropriate fragments. Cut out the fragments with the selected restriction enzymes and gel purify the fragment on a TBE/1% agarose gel.
2. Cut 2 µg of pMSe4 with the appropriate restriction enzymes in a total volume of 20 µL (*see* **Note 17**) and gel purify the vector fragment.
3. Set up a ligation reaction with 200 ng of pMSv4 and an two- to three-fold molar excess of the insert with 1 µL of T4 ligase in a total volume of 20 µL. Incubate for 4 h a 16°C.
4. Transform 100 µL of competent Xl-1blue *E. coli* with 2 µL of ligation reaction. Plate 100 µL aliquot on LBamp plates. Incubate overnight at 37°C.
5. Miniprep 12 colonies ***(14)*** and identify positive clones by restriction analysis. Verify the correct reading frame by sequencing the junction with the Cyc1 start codon using the pMSe4 sequencing primer.

3.1.3. Generation and Testing of Reporter Strain (see ***Note 18***)

1. Grow an overnight culture of W303-1A in 50 mL YPD at 30°C to a density of 0.8 OD^{600}
2. Centrifuge for 2 min at 3000*g* and resuspend the yeast cell pellet in 1 mL of ddH_2O. Transfer to an Eppendorf tube and centrifuge again for 2 min at 3000*g*.
3. Resuspend the pellet at a density of 50 OD^{600}/mL in TE/LiAc and shake at 30°C for 1 h.
4. Mix 2 µg of reporter plasmid, 2 µg of target expression plasmid, 10 µL of a 10 mg/mL solution of sheared salmon sperm carrier DNA (*see* **Note 19**) with 100 µL of yeast suspension and shake at 30°C for 30 min. Samples of the target expression plasmid containing either the full length, the mutated, or no transcription factor sequences should each be set up with combinations of reporter plasmids contain-

ing various copies of the DNA-binding site as well as a negative control plasmid containing no or a mutated binding site.
5. Add 700 µL of TE/LiAc/PEG. Mix well by pipeting up and down and shake for 30 min at 30°C.
6. Add 80 µL DMSO and heat shock for 10 min at 42°C. Pellet yeast cells by centrifugation at 3000*g* rpm in an Eppendorf centrifuge, wash with 1 mL of sterile H_2O, and pellet again. Discard 900 µL of supernatant, resuspend yeast cells in remaining volume, and plate out on SD-Trp-His plates. Incubate at 30°C for 2 d and purify several colonies of each condition by streaking out on a new SD-Trp-His plate.
7. Streak out several individual colonies from each condition on a Sgal-Trp-His + x-gal plate and incubate at 30°C (*see* **Note 20**). Check for color development after 2, 3, and 4 d. A reporter plasmid should be chosen that gives either no or negligible color development with the mutated transcription factor, but is active with the full-length factor (*see* **Note 21**).

3.2. Identification of Positive Clones from a cDNA Expression Library

3.2.1. Transformation of Plasmid Library (*see* **Note 9**)

1. Determine the transformation efficiency that can be achieved with the particular library, carrier DNA, and reporter strain in a small scale pilot transformation by following the protocol in **Subheading 3.1.3.** and transforming a serial dilution of the library. Note that the reporter strain has to be grown in selective SD-Trp-His medium.
2. For the large-scale transformation grow a 4 L culture of the reporter strain in SD-Trp-His to an OD^{600} of 0.8. Pellet yeast at 3000*g* for 5 min. Resuspend the yeast pellet in 4 L of YPD and grow to an OD^{600} of 1.0.
3. Centrifuge at 5000 rpm and resuspend the pellet in 800 mL sterile H_2O. Transfer to sixteen 50-mL polypropylene tubes and centrifuge again at 3000*g*. Resuspend each pellet in 5 mL TE/LiAc (corresponds to 50 OD^{600}/mL) and shake 1 h on a rotary shaker at 30°C.
4. Add 2–6 µg of library plasmid (*see* **Note 22**) and 500 µL of carrier DNA per tube and shake for an additional 30 min. Then add 35 mL TE/LiAc/PEG, mix well by pipeting and shake for another 30 min.
5. Add 4 mL DMSO and heat shock for 10 min at 42°C. Pellet yeast by centrifugation at 3000*g* for 5 min, wash with 50 mL of sterile H_2O, and pellet again. Resuspend pellets in sterile H_2O and pool to give a total volume of 80 mL.
6. Carefully layer nitrocellulose filter membranes onto 20 × 20 cm SD-Trp-His-Ura plates avoiding air bubbles. Plate out 4 mL of yeast solution per plate onto the filter membrane to yield a total of 20 plates (*see* **Note 23**). A large glass rod spreader should be used to achieve even distribution which will facilitate later identification of single colonies.
7. Plate out 100 µL of undiluted, 10^{-1}, and 10^{-2} dilutions of the yeast suspension on small 10-cm plates to monitor the transformation efficiency.

3.2.2. Isolation of Positive Colonies

1. Incubate at 30°C for 36 h until the first small colonies appear (*see* **Note 24**).
2. Carefully transfer the filter membrane with the small colonies to a Sgal-Trp-His-Ura + x-gal plate. Take care that neither of the two plates has any excess fluid to avoid smudging of the colonies and avoid air bubbles during the transfer (*see* **Note 25**).
3. Incubate at 30°C and pick blue colonies after 2 and 3 d of galactose induction. Purify the positive clones to single colonies on fresh SD-Trp-His-Ura plates.

3.3. Elimination of False Positive Clones

3.3.1. Testing for Galactose Dependence

1. Streak out two individual colonies each from the positive clones on SD-Trp-His-Ura + x-gal and Sgal-Trp-His-Ura + x-gal plates (*see* **Note 20**). This simple test will verify that the positive clones depend on the galactose induction of the target factor and the interacting library proteins and thus exclude reporter recombinations and false pickings.

3.3.2. Curing of Target and Library Plasmids

1. Grow a small colony of galactose dependent positive clones in 20 mL of YPD for 48 h (*see* **Note 26**) and plate 100 µL of a 1:500 dilution on a SD-Trp-His plate and on a SD-Ura-His plate. Incubate at 30°C for 2 d.
2. Replica plate colonies from each selection condition onto SD-Trp-His-Ura plates (*see* **Note 27**). Incubate both sets (original and replica plate) for an additional 2 d at 30°C.
3. Identify the colonies that grew on the SD-Trp-His plate, or the SD- Ura-His plate but not on the SD-Trp-His-Ura plates. These colonies should have lost the library plasmid (conferring Ura auxotrophy), or the target factor plasmid (conferring Trp auxotrophy) respectively. Restreak these colonies on SD-Trp-His vs SD-Trp-His-Ura and SD-Ura-His vs SD-Trp-His-Ura, respectively, to verify the phenotype.

3.3.3. Testing of Target and Library Plasmid Dependence

1. Restreak clones that have lost the target plasmid on Sgal-Ura-His + x-gal plates (*see* **Note 20**). Incubate for 48 h at 30°C. As a negative control, streak out a strain that has been transformed with the reporter plasmid and the empty library plasmid. For comparison of the color intensity also streak out on the same plate a colony that has *not* lost the target plasmid. Reporter activation between different clones may vary and the direct comparison between the conditions with or without the target plasmid is essential. This control should exclude the genes that can activate the reporter independently of the target protein (*see* **Note 28**).
2. Restreak clones that have lost the library plasmid on Sgal-Trp-His + x-gal plates and incubate for 48 h at 30°C. As a negative control include the original reporter strain and again also streak out a colony that has *not* lost the library plasmid for comparison of color intensity (*see* **Note 29**).

3. Identify the clones that require both the target plasmid and the library plasmid for reporter activity.

3.4. Isolation and Characterization of Positive Library Plasmids

3.4.1. Isolation of Library Plasmids (*see* **Note 30**)

1. To isolate the library plasmid coding for the interacting protein set up a 20-mL SD-Ura culture from a Trp plasmid-cured colony of the target- and library-dependent positive clones.
2. Plate out 100 µL of a 1:500 dilution on SD-Ura plates and incubate for 48 h at 30°C. Replica plate onto SD-Ura-His plates (*see* **Notes 26** and **31**) and incubate both plates for an additional 48 h. Identify the colonies that grow on SD-Ura but not on SD-Ura-His and thus have lost the His reporter plasmid.
3. Pellet 1 mL of an overnight culture from these colonies in a microcentrifuge tube and resuspend the pellet in 100 µL of TENS.
4. Add sterile glass beads until just below the level of the liquid and vortex vigorously for 1 min. Add 200 µL phenol and vortex for another minute. Separate the phases by centrifugation in a microcentrifuge and extract the aqueous layer one more time with phenol and once with phenol/chloroform/isoamylalcohol (25:24:1) (*see* **Note 32**).
5. Add 0.1 vol of 3 *M* NaAc, 2.5 vol of ethanol, and precipitate for 30 min at –20°C. Centrifuge for 15 min at 4°C, wash the pellet with 70% ethanol, dry in a Speedvac, and resuspend in 20 µL of TE.
6. Transform Xl-1 blue *E. coli* with 5 µL of this solution by chemical transformation or electroporation.
7. Miniprep several colonies from each transformation ***(14)*** and analyze by restriction digest. EcoRI should release the insert from the library plasmid pSD.10 and easily distinguish it from contaminations of the pMSe4-based target expression plasmid or the pMSv3-based reporter plasmid. Prepare a sequencing quality maxiprep from the library plasmids.

3.4.2. Characterization of Library Plasmids

1. Usually several versions of the same cDNA are isolated in screen. To avoid the sequencing of identical cDNAs, the isolated cDNAs should be first placed into homology groups. A rough estimate of the relatedness of the identified cDNAs can be obtained by digestion of the inserts with a frequently cutting restriction enzyme such as HaeIII.
2. The most reliable result is obtained by cross hybridization. For this purpose prepare several Southern blots (equaling the number of plasmids isolated) from gels containing EcoRI digests of all the isolated library plasmids ***(13)***. Then hybridize each blot with a different probe derived from the inserts of the library plasmids. Label the probes by the random priming method ***(15)*** and wash the blots under high stringency conditions using Church's buffer ***(16)***. The result of this should demonstrate whether a particular insert only hybridizes to itself or, also, to other isolated plasmids and thus place the different plasmids into homology groups.

3. Start by sequencing the largest insert of each homology group using a sequencing primers in the library plasmid (*see* **Note 33**). The inserts can also easily be recloned into simpler vectors such as the Bluescript™ series (Stratagene, La Jolla, CA) for more efficient sequencing.

3.5. Control Transformations with Positive Library Plasmids

3.5.1. Control Transformations

To finally verify that the isolated library plasmids indeed code for proteins interacting with the target protein, the original host strain should be retransformed with the reporter, target and isolated library plasmids. In addition, the reporter activation can be quantified using a colorimetric assay measuring ß-galactosidase activity.

1. Following the transformation protocol described in **Subheading 3.1.3.** transform the original host W303-1A with the following combinations for each of the isolated library plasmids (*see* **Note 34**):
   a) the isolated library plasmid, the pMSe4-based target vector, the pMSv3-based reporter plasmid;
   b) the ***empty*** library plasmid, the pMSe4-based target vector, and the pMSv3-based reporter plasmid;
   c) the isolated library plasmid, the ***empty*** pMSe4 vector, and the pMSv3-based reporter plasmid;
   d) the isolated library plasmid, the pMSe4-based target vector, and either the ***empty*** pMSv3 reporter plasmid or preferably a pMSv3-based reporter plasmid with a ***mutated*** binding site for the target factor.
2. Plate out on SD-Trp-His-Ura plates and incubate for 48 h at 30°C.
3. Purify three colonies from each condition by restreaking on new SD-Trp-His-Ura plates and incubating for an additional 48 h at 30°C.
4. Patch three individual colonies from each condition onto a Sgal-Trp-His-Ura + x-gal plate (*see* **Note 20**) and incubate 2–3 d at 30°C until a blue color develops. Alternatively, quantify reporter activity by the colorimetric assay **Subheading 3.5.2.**

3.5.2. Quantitative Colorimetric Assay for Reporter Activity (*see* **Note 35**)

1. Inoculate 3 mL of Sgal-Trp-His-Ura medium with a large individual colony (or several colonies purified from the same transformant; the OD^{600} should not be less than 0.25) from the SD- Trp-His-Ura plates (*see* **Note 36**) and incubate for 24 h at 30°C with vigorous shaking in small flask or 50 mL polypropylene tube.
2. Pellet 1.5 mL of the culture (should be between 0.4 and 1.2 OD^{600}) in an Eppendorf tube and wash once with sterile water. Resuspend the final pellet in 50–100 µL of ß-gal lysis buffer and freeze on powdered dry ice. Either store at –20°C or thaw on ice for further processing.
3. Prepare reaction tubes containing 500 µL of Z-buffer (with freshly added β-Mercaptoethanol) and 100 µL of ONPG (4 mg/mL) for each sample plus three blanks.

4. Add 10 µL (*see* **Note 37**) of the briefly vortexed cell suspension in precisely timed intervals. Mix the reaction tube and incubate at room temperature until yellow color develops (should be below 0.8 OD^{420} to be linear). Stop the reaction with 250 µL 1 *M* $NaCO_3$ with exactly the same time interval as before so that the incubation times are identical for all samples.
5. Spin tubes for 2 min in an Eppendorf centrifuge and measure the OD^{420} of the supernatant. Use a sample that did not contain any yeast lysate as a blank.
6. For normalization dilute 10 µL of cell suspension into 1 mL of H_2O and measure the OD^{600}. To conserve linearity use more extract if measurements are significantly below 0.1 or less extract if they are above 0.25. Vortex samples just prior to measurement, since the cells settle down quickly.
7. Units can be calculated according to the formula $U = 1000X\ OD^{420}/(CVt)$, where *C* is the concentration of extract measured as OD^{600}/vol, *V* is the volume of extract used in the assay, and *t* is the reaction time.

4. Notes

1. Because the yeast strain used has an *ade2* mutation, it may use up the adenine present in the YPD medium quite rapidly and accumulate a red colored metabolic intermediate that can be toxic to the cells at high concentrations. It is, therefore, advisable to add extra adenine (20 mg/L) if cells turn pink or brown in the YPD medium.
2. This synthetic medium is supplied complete with all necessary trace elements, vitamins and salts, for example, by Difco Laboratories (Detroit, MI).
3. The amino acid mixes lacking the appropriate amino acids can be made as 10X stock solutions, autoclaved for 15 min, and stored at 4°C. Complete premixed "drop out" media lacking specific amino acids are also commercially available from several suppliers, for example BIO101 (San Diego, CA).
4. Dependent on the library used: Leucine instead of uracil has to be lacking from selective plates and media if a plasmid library with a Leu2 auxotroph marker is used.
5. $NaPO_4$ buffer softens the agar; therefore the agar concentration has to be increased by 25%.
6. Yeast growth media have an acidic pH. For optimal ß-galactosidase activity and color development a neutral pH is required.
7. pMSv3 was derived from pLG670Z ***(17)***.
8. pMSe4 is similar to and derived from pSD.04 ***(6)***, which is based on the pRS yeast vector series ***(18)***.
9. Many available libraries are compatible with the system. VP16 tagged yeast expression libraries in pSD.10 containing an Ura3 auxotroph marker have been generated from random primed HeLa cell cDNA ***(6)*** and from polydT primed QT6 cell cDNA ***(10)***. The protocol has been written for this vector type. Several libraries have also been generated in the pACT vector ***(7,19)***, which expresses fusions with the Gal4 transactivation domain, uses a constitutive ADH promoter and has a Leu2 auxotroph marker. Because of the constitutive expression, growth inhibitory cofactors are unlikely to be identified with these reagents. Most

importantly, if these libraries are used, all media and plates mentioned in the protocol have to lack leucine instead of uracil. The commercially available Matchmaker™ libraries (Clontech, Palo Alto, CA) are based on very similar vectors and are thus also compatible with the system, but again require media lacking leucine instead of uracil. These libraries have been generated from many different cDNA sources. If none of the available libraries meet the particular needs, a new library has to be generated. In this case there are several options. If pSD.10 is chosen as a library vector the use of a RNAseH– reverse transcriptase such as Superscript™ (Life Technologies, Rockville, MD) for the generation of the cDNA is recommend. The ligation of nonpalindromic BstXI linkers (InVitrogen, San Diego, CA) to the cDNA for the cloning into pSD.10 has the advantage to minimize self-ligation of vector or cDNA insert without requiring dephosphorylation. It is crucial to optimize the ligation conditions with a BstXI test fragment before cloning the cDNA into the vector. It is also possible to use the pGAD vector series sold by Clontech (Palo Alto, CA) or the pAD-Gal4 phagemid vector from Stratagene. The latter has the advantage that the cDNA library can be generated in the vector HybriZAP II™, which can then be converted into a plasmid library by in vivo mass excision. Custom library services are also available for this vector.

10. XhoI and SalI have different recognition sequences but compatible overhangs. Thus, the two enzymes recognize 5' head-to-head and 3' tail-to-tail ligations of the oligonucleotides, respectively, but not the composite XhoI/SalI head-to-tail ligations. If these two enzymes are included in the ligation reaction, a directional multimerization of the oligonucleotides can be achieved.
11. Instead of multimerized consensus-binding sites, it is also possible to clone an authentic promoter segment into the reporter plasmid.
12. Oligonucleotides can also be ordered phosphorylated, in which case the phosphorylation step in the protocol can be omitted.
13. The author is giving a protocol for colony hybridization as a fast and generally applicable method to identify plasmids, which have oligonucleotide inserts. If the inserted sequences contain useful restriction sites, miniprep procedures and restriction analysis can be used as well.
14. Care has to be taken that the correct reading frame is maintained. The pMSe4 expression vector utilizes the *S. cerevisiae* Cyc1 ATG (*see* **Fig. 3**).
15. If the transcription factor under investigation has no or only weak transactivation activity, the full-length molecule may be used as well. Point mutations or small internal deletions known to abolish transactivation activity can also be used instead of full deletions of the transactivation domains.
16. An expression construct with the full-length factor is useful to identify the best reporter plasmid.
17. If possible, use two unique sites. If only one enzyme is used to open up the vector DNA, it has to be dephosphorylated as described **Subheading 3.1.1.1.**
18. The yeast transformation protocol in this section is a modification of the one published by Gietz and Schiestl ***(20)***.

19. The carrier DNA is an important factor for transformation efficiency. 1–3 × 10^4 transformants/µg of plasmid are routinely achieved. If the results are unsatisfactory, different sources or preparations of DNA that have been sonicated for different lengths of time (average length should be ~1 kb) should be tested for the best transformation efficiency. This is especially important for the library transformation.
20. The color development can be best observed when the yeast is streaked out on a nitrocellulose filter placed on top of the indicator plates.
21. The quantitative colorimetric assay with ONPG for ß-galactosidase activity given in **Subheading 3.5.2.** can also be used to characterize the backgound activity of the reporter system, but a reasonably low background in the plate assay with x-gal is essential, because this detection system is used later in the library screen.
22. The correct amount of library DNA has to be calculated from the transformation efficiency determined in the pilot experiment. An efficiency between 1–3 × 10^4 colonies/µg is usually achieved with the protocol and a total of 10^6 colonies is desired resulting in a total of 33–100 µg of library to be transformed.
23. Fifty thousand colonies can be separated as individual colonies on a 20 × 20cm plate. A total of 10^6 colonies can thus be screened with 20 plates.
24. Preincubating the transformants on glucose plates makes it possible to detect interaction partners that are growth inhibitory or mildly toxic to the yeast cells. Under the glucose conditions, the library proteins are not expressed, the transformants can grow up to a small colony, which can be identified as positive by reporter activation even if the cells stop growing after the library induction. The transformants should not be preincubated longer than 36 h, because the x-gal substrate will not diffuse well into larger colonies. If the studied target transcription factor has no negative effect on the yeast growth and libraries with constitutive expression of the cDNA are used (*see* **Note 9**), the preincubation on glucose plates may be skipped and the transformants can be plated directly on galactose plates with nitrocellulose filters.
25. The best result is achieved by allowing the filter to touch one end of the new plate and slowly lowering it until it completely covers the whole plate. It may be advisable to practice this with an empty filter before risking the precious library transformants.
26. The yeast strain does not require the auxotrophy marker genes for growth in full medium so that the plasmids carrying them will eventually be lost at a certain percentage.
27. Besides the classical method of replica plating with a piece of velvet cloth stretched over a pestle, colonies can also be transferred by placing a nitrocellulose filter on top of the colonies and then placing the filter onto a new plate with the colony side up.
28. This class of clones will typically include nonspecific and specific DNA-binding proteins that can recognize the promoter of the reporter plasmid directly. Clones that are only partially dependent on the target protein should not be discarded,

because transcription factor cofactors may have a certain affinity for DNA, but still bind specifically to the target factor. It is, therefore, most important to screen for a difference of reporter activity in the presence or absence of the target protein.

29. This class of false positives will typically include reporter recombinations that are galactose dependent or mutations of the target protein. They are usually much rarer than the first class of false positives.
30. This protocol is an adaptation of the one published by Hoffman and Winston ***(21)***.
31. The replica plating can also be skipped to prepare plasmids directly from the SD-Ura cultures. In this case, however, a majority of the isolated plasmids will be the reporter plasmid and significantly more DNA minipreps have to be prepared from plasmid transformed bacteria.
32. The purity of plasmid DNA can be increased by treating the aqueous supernatant with commercially available DNA binding glassmilk (for example, Geneclean™ [BIO101 Inc.]), but usually a sufficient number of *E. coli* transformants can be obtained without this step.
33. If pSD.10 is used as a library plasmid, for initial sequencing the following primers can be used; at the 5' end a primer in VP16: 5'-GCGCTCTGGATATGGCCG-3, and at the 3' end a primer in the cyc1 terminator: 5'-GTCACATAACTA-ATTACATG-3'.
34. Only the condition (a) should result in reporter activity. All the other conditions should be negative or at least significantly lower than (a), demonstrating that the reporter activity requires an intact binding site for the target factor, the target factor itself, and the isolated new protein.
35. This protocol is a modified version of the one published by Harshman et al. ***(22)***.
36. Use fresh plates. Colonies from plates stored at 4°C will grow and induce poorly in galactose.
37. The activity of samples can vary significantly. Up to 100 µL of extract can be used in the assay if the activity should be weak.

References

1. Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340,** 245–246.
2. Fields, S. and Sternglanz, R. (1994) The two-hybrid system: an assay for protein-protein interactions. *Trends Genet.* **10,** 286–292.
3. Zervos, A. S., Gyuris, J., and Brent, R. (1993) Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* **72,** 223–232.
4. Gallant, P., Shiio, Y., Cheng, P. F., Parkhurst, S. M., and Eisenma, R. N. (1996) Myc and Max homologs in Drosophila. *Science* **274,** 1523–1527.
5. Hurlin, P. J., Queva, C., and Eisenman, R. N. (1997) Mnt, a novel Max-interacting protein is coexpressed with Myc in proliferating cells and mediates repression at Myc binding sites. *Genes Dev.* **11,** 44–58.
6. Dalton, S. and Treisman, R. (1992) Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell* **68,** 597–612.

7. Gstaiger, M., Knoepfel, L., Georgiev, O., Schaffner, W., and Hovens, C. M. (1995) A B-cell coactivator of octamer-binding transcription factors. *Nature* **373,** 360–362 issn: 0028-0836.
8. Strubin, M., Newell, J. W., and Matthias, P. (1995) OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* **80,** 497–506.
9. Sieweke, M. H., Tekotte, H., Frampton, J., and Graf, T. (1996) MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell* **85,** 49–60.
10. Sieweke, M. H., Tekotte, H., Jarosch, U., and Graf, T. (1997) Cooperative interaction of Ets-1 with USF-1 required for HIV-1 enhancer activity in T cells. Embo J. **17,** 1728–1739.
11. Petersen, J. M., Skalicky, J. J., Donaldson, L. W., McIntosh, L. P., Alber, T., and Graves, B. J. (1995) Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science* **269,** 1866–1869.
12. Jonsen, M. D., Petersen, J. M., Xu, Q. P., and Graves, B. J. (1996) Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell Biol.* **16,** 2065–2073.
13. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* , 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
14. Zhou, C., Yang, Y., and Jong, A. Y. (1990) Mini-Prep in ten minutes. *Biotechniques* **8,** 172–173.
15. Feinberg, A. P. and Vogelstein, R. (1984) Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137,** 266–267.
16. Church, G. M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
17. Guarente, L. and Ptashne, M. (1981) Fusion of Escherichia coli lacZ to the cytochrome c gene of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* **78,** 2199–2203.
18. Sikorski, R. S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122,** 19–27.
19. Durfee, T., Becherer, K., Chen, P. L., Yeh, S. H., Yang, Y., Kilburn, A. E., Lee, W. H., and Elledge, S. J. (1993) The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* **7,** 555–569.
20. Gietz, R. D. and Schiestl, R. H. (1991) Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. *Yeast* **7,** 253–263.
21. Hoffman, C. S. and Winston, F. (1987) A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* **57,** 267–272.
22. Harshman, K. D., Moye-Rowley, W. S., and Parker, C. S. (1988) Transcriptional activation by the SV40 AP-1 recognition element in yeast is mediated by a factor similar to AP-1 that is distinct from GCN4. *Cell* **53,** 321–330.

6

Inverse PCR (IPCR) for Obtaining Promoter Sequence

Tony Triglia

1. Introduction

Obtaining the promoter sequence for a gene when only the cDNA sequence is available can be an arduous task, especially because genomic DNA libraries usually in λ phage vectors need to be screened. An easier method to obtain promoter sequence without the need for libraries is to use the technique of inverse polymerase chain reaction (IPCR) ***(1,2)***. IPCR leads to the amplification of previously unknown sequences because the primers that initially face away from each other on the linear template can be made to face each other as in normal PCR following circularization of the template (**Fig. 1**). Further amplification with nested primers ensures the integrity of the final product, which can be sequenced directly.

Subsequent to IPCR, a number of similar techniques, which can be used for walking in uncloned genomic DNA have been published. These include SSP-PCR ***(3)***, targeted gene walking PCR ***(4)***, RAGE ***(5)***, and vectorette PCR ***(6)***. Two of these methods specifically mention their use in obtaining promoter sequence. The mouse growth hormone releasing hormone gene promoter was obtained by using the RAGE technique ***(5)***, whereas the tissue type plasminogen activator gene promoter was obtained by a modification of the vectorette PCR technique ***(7)***. Without doubt many other promoter regions have also been obtained by one of the above methods, but the details most likely remain buried in the Materials and Methods section of the papers. This chapter describes a modification of IPCR that specifically leads to promoter sequence.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

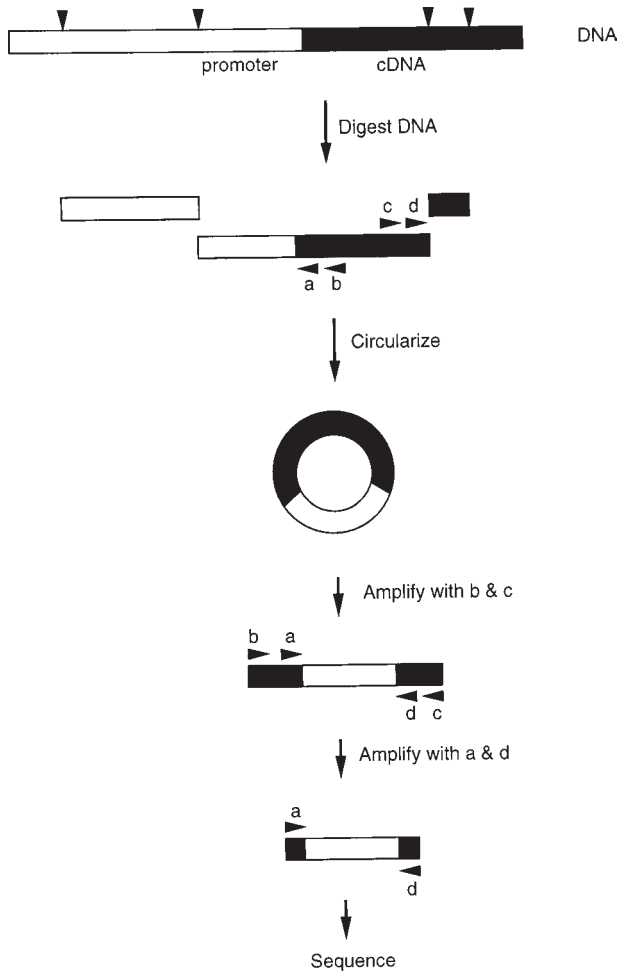


Fig. 1. Promoter sequence by Inverse PCR.

2. Materials

1. Genomic DNA in TE.
2. DNA buffer: 6 *M* guanidine, 0.1 *M* sodium acetate, pH 5.2.
3. TE: 10 m*M* Tris, pH 8.0, 1 m*M* EDTA.
4. Restriction enzymes that cut the genomic DNA relatively frequently.
5. 20-mer oligonucleotides (unpurified).
6. T4 ligase and 10X ligation buffer (Promega, Madison, WI).
7. Taq polymerase and 10X PCR buffer (Perkin-Elmer).
8. TaqStart™ Antibody (Clontech, Palo Alto, CA).
9. PCR cycling instrument.

10. 1.2% agarose/EtBr gels.
11. Geneclean™ (or similar) DNA purification kit (BIO101, Vista, CA).
12. Phenol-chloroform (1:1), TE-saturated.
13. 3 *M* sodium acetate, pH 5.2.
14. 70% ethanol.
15. 100% ethanol.

3. Method

3.1. Preparation of Genomic DNA (*see* Note 1)

This method for mammalian cells or tissues is a quick procedure suitable for multiple samples *(8)*.

1. Lyse cell suspensions in 7.5 vol of DNA buffer for 1 h.
2. Layer the lysed cells under 18 mL of ethanol.
3. Slowly stir the interface between the cell lysate and the ethanol with a pipet, collecting the adhering DNA to the pipet.
4. Wash the attached DNA with ethanol, dry slightly and allowed to rehydrate in TE. The dissolved DNA is ready for PCR.

3.2. Inverse PCR

1. Digest 1 µg DNA with a number of restriction enzymes in separate tubes in a total volume of 10 µL (*see* **Note 2**). Digest to completion.
2. Heat-inactivate the enzyme at 68°C for 10 min if it is heat-labile or alternatively extract with phenol-chloroform and ethanol precipitate with 0.1 vol 3 *M* sodium acetate pH 5.2 and 2.5 vol ethanol if heat-stable. After precipitation, wash the DNA pellet with 70% ethanol and resuspend to 10 µL TE.
3. Take 2 µL (i.e., 0.2 µg) of heat-inactivated or ethanol-precipitated digested DNA and setup a self-ligation reaction by adding 10 µL 10X ligation buffer, 10 µL 10 m*M* ATP, 4 µL T4 DNA ligase (3U/µL), and ddH_20 to 100 µL. Ligate 16 h at 15°C (*see* **Note 3**).
4. Following overnight ligation, remove 10 µL from each tube and add directly to a 100 µL PCR reaction containing the internal pair of the nested primers (i.e., b and c in **Fig. 1**) (*see* **Note 4**). Typical PCR reaction conditions are used such as 94°C/1 min, 50°C/1 min, 70°C/3 min for 35 cycles.
5. Ethanol precipitate the PCR reactions with 0.1 vol 3 *M* sodium acetate, pH 5.2 and 2.5 vol ethanol. Run each precipitated DNA in a single well of a 1.2% agarose/EtBr gel. Visualize the bands over a long-wavelength UV-transilluminator and use a toothpick to remove a small amount of amplified product into 5 µL of water *(9)* (*see* **Note 5**).
6. Use the total amount of first-round PCR product in a second round PCR reaction with oligonucleotides a and d. After the temperature cycling is complete, ethanol precipitate and run each product on an agarose gel as before in **step 6**. Purify the amplified DNA products using a Geneclean kit or similar (*see* **Note 6**).
7. Sequence the PCR product from both directions using primers a and d (*see* **Note 7**).

4. Notes

1. Genomic DNA needs to be clean enough to be readily digested by restriction enzymes and not to be inhibitory to ligation of the DNA.
2. In order to obtain sequence only from the upstream region, the enzymes chosen must cleave within the known sequence. The length of the fragments expected from each digest can be determined by Southern blots. However, in practice, it is easier to cleave with five different enzymes in order to ensure that at least one or two of the digests yield fragments that are neither too large to be amplified (2–3 kb with Taq polymerase) nor too small to be worth the effort in sequencing the fragment.
3. Diluting the digested DNA into a 100-µL self-ligation reaction ensures that the DNA concentration does not exceed 3 µg/mL. Experimentally, it has been found that at DNA concentrations below 3 µg/mL circles rather than linear concatamers tend to form ***(10)***. Also ligating in a 100 µL reaction ensures that any carried-over restriction enzyme buffer does not significantly alter the composition of the ligation reaction components.
4. The 5' ends of the primers b and c must be separated by at least 100 bp in order to ensure that a reasonable proportion of molecules are nicked in that region during the PCR temperature cycling steps. Only molecules nicked between the b and c primers can serve as templates for PCR with those primers.
5. Not all of the first round PCR reactions will work because either the product was too large to be amplified under the conditions used for PCR or the restriction enzyme digest and/or the self-ligation steps were not efficient. A recent paper published electronically ***(11)*** indicates that the technique of long and accurate PCR ***(12)*** can be combined with IPCR to generate products up to 4.5 kb. Running the PCR reactions on an agarose gel allows you not only to determine how many of the reactions have worked, but also removes the first round primers from the PCR product, which can be inhibitory or lead to spurious products if carried over to the second-round PCR reaction.
6. A much greater amount of the second round PCR products should be visible following EtBr staining compared to the first round products, and their size should have dropped by the difference between the a and b and the c and d primers. If both of these criteria are met, it is almost certain that the product derives from a specific amplification into the previously unknown promoter region.

   Note that it is necessary in most cases to nest only a single primer so that only primer a or d (not both) are needed in the second-round PCR.
7. Should it be necessary to clone the PCR product in order to obtain good sequence, either direct cloning of the product into a commercial T-vector or by digestion of the PCR product prior to cloning into an appropriate vector is suitable. However, the latter case requires the addition of restriction sites to the 5' ends of the a and d primers.

References

1. Triglia, T., Peterson, M. G., and Kemp, D. J. (1988) A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucl. Acids Res*. **16**, 8186.
2. Ochman, H., Gerber, A. S., and Hartl, D. L (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**, 621–625.
3. Shyamala, V. and Ames, G. F. (1989) Genome walking by single-specific primer polymerase chain reaction: SSP-PCR. *Gene* **84**, 1–8.
4. Parker, J. D., Rabinovitch, P. S., and Burmer, G. C. (1991) Targeted gene walking polymerase chain reaction. *Nucl. Acids Res*. **19**, 3055–3060.
5. Mizobuchi, M. and Frohman, L. A. (1993) Rapid amplification of genomic DNA ends. *Biotechniques* **15**, 215–216.
6. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., and Markham, A. F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome(YAC) clones. *Nucl. Acids Res*. **18**, 2887–2890.
7. Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A., and Lukyanov, S. A. (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucl. Acids Res*. **23**, 1087–1088.
8. Bowtell, D. D. (1987) Rapid isolation of eukaryotic DNA. *Anal. Biochem* **162**, 463–465.
9. Kadokami, Y. and Lewis, R. V. (1994) Repeated PCR of a gel band can be used to obtain a single PCR band. *Biotechniques* **17**, 438.
10. Collins, F. S. and Weissmann, S. M. (1984) Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method. *Proc. Natl. Acad. Sci. USA* **81,** 6812–6816.
11. Leister, D., Thompson, R., and Gebhardt, C. (1997) Production of long inverse PCR products facilitates the isolation of YAC insert termini. *Tech. Tips Online* (http://www.elsevier.com/locate/tto) T40066.
12. Barnes, W. M. (1994) PCR amplification of up to 35kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA* **91,** 2216–2220.

7

PCR-Directed Linker Scanning Mutagenesis

Kurt Gustin and Robert D. Burk

1. Introduction

Linker scanning mutagenesis is a powerful method with which to assay the contribution of individual DNA sequence elements within a transcriptional control region ***(1)***. By replacing discrete segments of DNA with heterologous segments of the same length, the topological and spatial organization of the DNA helix is maintained. This allows the contribution of individual DNA-binding motifs to be determined in the context of the native DNA helix configuration. Although most commonly employed in the analysis of promoter and enhancer regions, linker scanning mutagenesis has also been adapted for the analysis of protein coding regions ***(2)***.

The original method for generating linker scanning mutants was time consuming and labor intensive ***(1)***. This method relied on partial exonucleolytic digestion and synthetic linker ligation to generate two pools of mutants, deleted from opposite ends of the region of interest. Clones from these pools were then sequenced and pairs were selected and combined to replace the desired sequence. The advent of the polymerase chain reaction (PCR) has led to the development of a number of different strategies for generating linker scanning mutants that are faster and easier to perform than the original method ***(3–8)***.

In this chapter we describe a method that was used to generate a panel of linker scanning mutants within the hepatitis B virus enhancer I region (*see* **Fig. 1** and **ref. *9***). The reader is encouraged to examine alternative protocols, several of which have appeared in previous volumes of *Methods in Molecular Biology* ***(10,11)***, to determine which method is best suited for each experimental objective.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

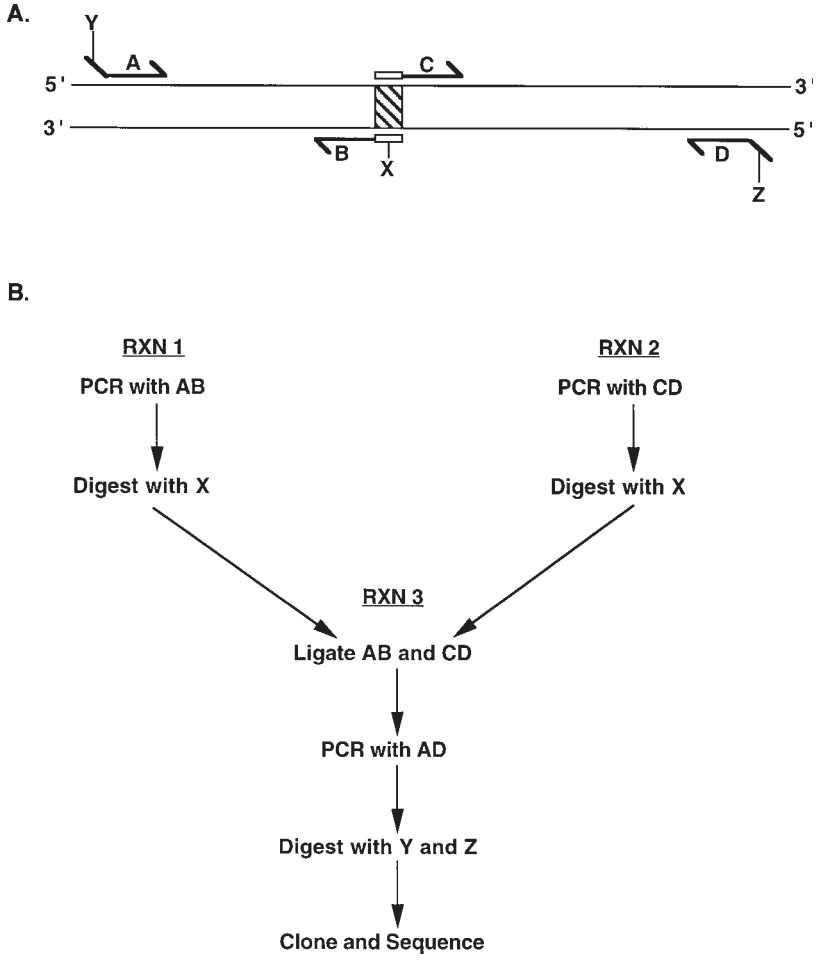


Fig. 1. **(A)** Schematic representation of primers used to generate linker scanning mutants. The targeted region of DNA is represented by thin black lines with the 5' and 3' ends indicated. The region to be replaced by heterologous sequences is shown as a hatched box. Primers are designated A, B, C, and D, and their approximate binding sites are shown. Open boxes on primers B and C represent the linker sequences that will be inserted. X, Y, and Z designate restriction sites used in constructing and characterizing the mutants. **(B)** Flow chart for PCR-directed linker scanning mutagenesis. The three PCR reactions are designated RXN 1, 2, and 3. (See text for details.)

2. Materials

1. Template DNA: Containing the region to be mutagenized (*see* **Note 1**).
2. Oligonucleotide primers: diluted to 10 pmol/µL (*see* **Note 2**).

3. 10X dNTPs: Prepare a solution that is 2 m*M* each dNTP.
4. Thermostable DNA polymerase and 10X buffer (*see* **Note 3**).
5. Thermocycler and appropriately sized, thin-walled PCR tubes.
6. Restriction endonucleases for digesting the PCR products and vector DNA.
7. T4 DNA ligase and 5X buffer.
8. Reagents for the purification of PCR products (*see* **Note 4**).
9. Vector DNA: Appropriately digested and prepared for ligation with the mutagenized fragment.
10. Competent cells and selective medium for the isolation of recombinant clones.
11. Reagents for the isolation and analysis of recombinant plasmids.
12. Gel electrophoresis equipment: gel boxes, power supplies, agarose or acrylamide, molecular weight markers, running buffer, and stain to analyze PCR products and recombinant clones.
13. Reagents and equipment to sequence recombinant DNA clones.

3. Methods

3.1. PCR Reactions 1 and 2

Prepare two reactions: one with primers AB and one with primers CD.

1. To a PCR tube on ice add 5 µL of 10X PCR buffer, 5 µL of 10X dNTPs, 5 µL of primer A (or C), 5 µL of primer B (or D), 1 ng of template DNA, 1.25 U of thermostable polymerase, and water to 50 µL (*see* **Note 5**).
2. Program the thermalcycler. Generally, we perform an initial denaturation at 94°C for 5' followed by 25–30 cycles with the following parameters: denaturing at 94°C for 30 s, 30 s at the annealing temperature (*see* **Note 6**), and extending at 72°C for 30 s. Following the last extension, the reactions are incubated at 72°C for 10 min and stored at 4°C until needed.
3. Start the cycler. Once the block has reached 94°C, place the reaction tubes in the machine.
4. Following completion of the PCR, remove 5 µL of the reaction to a new tube, add 2.5 µL of the appropriate 10X reaction buffer, 16.5 µL of water, and 1 µL (10 U) of restriction endonuclease cleaving at the engineered restriction site "X" (**Fig. 1**). Incubate at the recommended temperature for 1 h (*see* **Note 7**).
5. Inactivate the restriction endonuclease by heating at 65°C for 20 min (*see* **Note 8**).

3.2. PCR Reaction 3

1. Remove 5 µL of the digested PCR fragment from each tube and combine.
2. Heat at 65°C for 1 min and place on ice.
3. Add in order 9 µL of water, 5 µL of 5X ligase buffer, and 1 µL (1 U) of T4 DNA ligase. Mix and incubate at 25°C for 1 h.
4. To a PCR tube on ice, add 5 µL of 10X PCR buffer, 5 µL of 10X dNTPs, 5 µL of primer A, 5 µL of primer D, 1 µL of the ligation mix, 1.25 U of thermostable polymerase and water to 50 µL (*see* **Note 9**).

5. Program the thermalcycler and carry out the PCR as described above (**Subheadings 3.1.2.** and **3.1.3.**).
6. Analyze the reaction products on an appropriate gel (*see* **Note** 7). If a single product is observed, proceed to **step 7**. If multiple bands are detected, gel purify the appropriately sized fragment and proceed to the next step.
7. Adjust the salt concentration and pH of the remaining PCR reaction, add 10 U of each enzyme (X and Y, *see* **Fig. 1**) and incubate at the recommended temperature for 2 h.
8. Purify the PCR product and ligate with the prepared vector DNA (*see* **Note 4**).
9. Transform, screen, and identify recombinants ***(12)***.
10. Sequence to confirm that the expected mutations have been introduced and that no secondary mutations have occurred.

4. Notes

1. Generally, plasmid DNA prepared by any of the standard Methods ***(12)***, or using commercially available purification kits, will work well.
2. The design of primers B and C determines the size and sequence of the "linker" insertion. These primers should have a 3' end complementary to the target region, and 5' ends that are complementary to each other and contain the heterologous linker sequence. The 5' ends should also include a restriction endonuclease cleavage site (X in **Fig. 1A**) that can be used to facilitate screening of recombinants. For cloning purposes, it is convenient to incorporate restriction sites into the 5' ends of primers A and D (i.e., sites Y and D, respectively) that allow directional cloning of the mutagenized fragment. Alternatively, sites that are internal to primers A and D can be used, or the fragment can be cloned directly into a PCR product cloning vector.
3. Thermostable DNA polymerase is available from a variety of sources and the manufacturer's recommended reaction conditions will work well for most primer/template combinations. If problems are encountered with the amplification, conditions may have to be optimized by adjusting the concentration of $MgCl_2$. Use of a thermostable polymerase that possesses proofreading capability (*pfu*, *pwo*, *vent*, etc.) will reduce the likelihood of nonspecific mutations being introduced and facilitate blunt-end cloning of PCR products.
4. PCR products can be purified either by column chromatography or from gel slices following electrophoresis. Provided that the PCR reactions yield a single band, column chromatography is the fastest method for removing primers, salt, and proteins. If multiple bands are observed, it is usually worthwhile to gel purify the correct fragment before proceeding.
5. If nonspecific priming appears to be a problem, adding the polymerase after the initial denaturation step can alleviate this problem.
6. The annealing temperature should be 4°C below the predicted *Tm* (i.e., melting temperature) of the primers being used. The Tm can be calculated by using the equation: $Tm = 4(G + C) + 2(A + T)$, where G, C, A, and T represent the total

number of occurrences of the respective nucleotide in the primer. Note that during the initial cycles, the primers will have relatively short regions of homology confined to their 3' ends, thus requiring lower annealing temperatures. After a few cycles have been completed, template with homology that extends the length of the primer will be present and the *Tm* can be increased.

7. Run 5–10 µL of the original PCR reactions on an agarose or acrylamide gel to confirm that only a single fragment was produced. If more than one product is visible, it may be possible to gel purify the correct product. Alternatively, the reaction conditions can be optimized by adjusting the temperature, salt concentration, or DNA template concentration.
8. If the restriction endonuclease cannot be heat inactivated, the reaction must be extracted once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and once with an equal volume of chloroform:isoamyl alcohol (24:1) before proceeding to the ligation. Alternatively, three freeze-thaw cycles are often effective.
9. If a single product is not detected, it may help to vary the amount of the ligation used in the PCR reaction.

References

1. McKnight, S. L. and Kingsbury, R. (1982) Transcriptional control signals of a eukaryotic protein-coding gene. *Science* **217,** 316–324.
2. Cunningham, B. C. and Wells, J. A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244,** 1081–1085.
3. Nelson, R. M. and Long, G. L. (1989) A general method of site-specific mutagenesis using a modification of the *thermus aquaticus* polymerase chain reaction. *Anal. Biochem.* **180,** 147–151.
4. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77,** 51–59.
5. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucl. Acids Res.* **16,** 7351–7367.
6. Perrin, S. and Gilliland, G. (1990) Site-specific mutagenesis using asymmetric polymerase chain reaction and a single mutant primer. *Nucleic Acids Research* **18,** 7433–7438.
7. Li, X-M. and Shapiro, L. J. (1993) Three-step PCR mutagenesis for 'linker scanning'. *Nucl. Acids Res.* **21,** 3745–3748.
8. Gustin, K. E. and Burk, R. D. (1993) A rapid method for generating linker scanning mutants utilizing PCR. *BioTechniques* **14,** 22–23.
9. Gustin, K. E., Shapiro, M., Lee, W., and Burk, R. D. (1993) Characterization of the role of individual protein binding motifs within the hepatitis B virus enhancer I on promoter activity using linker scanning mutagenesis. *Virology* **193,** 653–660.
10. Harlow, P. P., Hobson, G. M., and Benfield, P. A. (1996) Construction of linker-scanning mutations using PCR. *Methods in Molecular Biology* **57,** 287–295.

11. Schanke, J. T. (1997) Linker scanning mutagenesis by three-step PCR. *Methods in Molecular Biology* **67,** 197–202.
12. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. eds. (1989) *Current Protocols in Molecular Biology*, New York: Wiley.

8

Transfection Technologies

Elaine T. Schenborn

1. Introduction: Historical Background and Current Applications

Experimental transfer of discrete nucleic acids to eukaryotic cells has allowed us to take giant strides in our basic understanding of eukaryotic gene expression and regulation. Viral particles transfer nucleic acid naturally to cells by the process of infection; the process of nucleic acid transfer to cells by artificial, nonviral methods is referred to as "transfection." Transfection to mammalian cells by chemical, nonviral methods was introduced more than 30 years ago ***(1)***. Plasmid DNA is currently the type of nucleic acid most frequently transfected, but oligonucleotides, yeast artificial chromosomes, and RNA can each be transfected to mammalian cells using appropriate methods. Nucleic acid delivery or gene transfer has been expanded to the whole organism level and is central for our ability to produce transgenic animal models and to engage in clinical studies for the purpose of correcting metabolic disorders and treating cancer ***(2)***.

Historically, the first type of nucleic acids transferred to mammalian cells were portions of RNA and DNA viruses ***(1,3)***. Vaheri and Pagano in 1965 pioneered the use of diethylaminoethyldextran (DEAE-dextran) for transfection of poliovirus RNA ***(1)***. The calcium phosphate precipitation method, still the most widely used transfection technique for cells in vitro, was published by Graham and Van der Eb in 1973 as a means to transfer adenviral DNA ***(3)***. Progress in transfection technology was relatively slow until extra tools for molecular biology became available. Plasmid vectors, together with cloning and amplification methods, provided researchers with "molecular tweezers, scissors, and glue" to design and manipulate DNA sequences that could be purified in milligram amounts.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

Transfection methods are now routinely used for studies of cellular metabolism, transcriptional control points, and protein function. The understanding of eukaryotic gene regulation benefited greatly by the codevelopment of improved transfection reagents together with genetic reporter and selection systems. Transcriptional activity, in a defined cellular context, can be studied by transfection of reporter genes linked operationally to promoters and transcriptional elements of interest. The ideal reporter gene codes for a protein not found within the cell, and is associated with an assay that is sensitive, quantitative, rapid, and reproducible (for a general review, *see* **ref. 4**). In addition to transcriptional regulation analyses, reporter gene systems can be used for other applications, such as protein:protein interaction studies. An example of this type of application is the widespread use of the yeast or mammalian two-hybrid systems to identify and characterize interacting protein domains ***(5,6)***. Cellular protein trafficking, another type of application marrying transfection and reporter technologies, can be studied by fusing the protein of interest with a fluorescing reporter gene protein such as the green fluorescent protein, and following the protein movements microscopically ***(7)***.

Selection of cells that maintain the transfected DNA for multiple generations, either by chromosomal integration or episomal maintenance, is facilitated by vectors expressing a gene that allows positive or negative selection. For example, the marker gene for neomycin phosphotransferase used together with the drug G418 ***(8)*** provides positive selection, and the thymidine kinase gene marker in combination with nucleotide analogs ***(9)*** provides a negative selection strategy. Multiple types of applications are possible with isolated cell clones that are stably transfected. For example, screening assays for drugs that modulate specific promoter activity are made possible by engineered cell lines containing an integrated reporter gene linked to the regulatory region under investigation. Also, cell lines can be engineered to express high levels of a pharmaceutical value recombinant protein.

Transfection of RNA synthesized in vitro allows the examination of elements affecting, for example, RNA stability, translation efficiency, splicing, and processing, as well as the effects of ribozyme activity upon cellular processes ***(10,11)***.

1.1. Classification of Transfection Technologies

Many techniques have been developed for the purpose of transferring nucleic acids to cells. The following characteristics describe desirable attributes for a successful transfection technology: highly efficient delivery of nucleic acid to the appropriate cellular locus (nucleus for plasmid DNA and cytoplasm for RNA); low cellular toxicity; minimal interference with normal cellular physiology; ease of use; and reproducibility. For gene therapy applications, there are

additional considerations such as biological half-life and cell-type-specific delivery.

Transfection technologies can be classified into two categories: chemical reagents and physical methods. Chemical reagents include cationic polymers and chemicals such as DEAE-dextran, calcium phosphate, and cationic liposomes. Physical methods for introducing nucleic acids into cells include electroporation, and biolistic particle delivery. This is not an exhaustive listing of transfection technologies, but represents methods most commonly used and available today for transfections. Each will be briefly described in the following sections. A summary of applications for these transfection technologies is given in **Table 1**. For each cell type, it is critical to select the best method and reagents for gene transfer and to optimize conditions for that particular method. Use of plasmid DNA vectors with reporter genes facilitates the transfection optimization procedures.

2. Chemical Reagents

One common feature of chemical reagents used for transfection is their cationic nature. The positively-charged molecules associate with the negatively charged phosphates of the nucleic acid backbone. For in vitro transfections, an overall net positive charge of the nucleic acid/reagent mixture is typically effective. This presumably allows closer apposition of these nucleic acid complexes with the outer cell membrane having negative surface charges. Calcium phosphate coprecipitation and DEAE-dextran methods are the oldest chemical methods for transfection, and are still in popular use. Other types of chemical transfection reagents used today are polycationic molecules such as polyamidoamine dendrimers ***(12,13)***; polypeptides such as polylysine, polyornithine, and conjugates ***(14–16)***; and polyethylenimine ***(17)***.

2.1. Calcium Phosphate Coprecipitation

Theory: Calcium associates with the negative charges of the nucleic acid. Addition of a buffered phosphate solution results in precipitates of DNA with calcium and phosphate. These particles may enter cells by phagocytosis ***(18)***.

Protocol and Critical Parameters: 0.25 *M* $CaCl_2$ and DNA (20–80 µg/mL) are mixed together in a sterile tube. An equal volume is added of HEPES buffered saline (HBS; 50 m*M* HEPES, 280 m*M* NaCl, 1.5 m*M* sodium phosphate, pH 7.1). The precipitate forms at room temperature for 30 min and is added drop-wise to cultured cells. The DNA precipitate volume equals approx 10% of the final volume of the cell culture medium. The culture medium is replaced the next morning, or earlier if toxicity is evident. A brief incubation with glycerol or dimethylsulfoxide prior to the media change increases transfection efficacy in some cell types ***(19)***.

Table 1
Transfection Technology Applications

	Calcium Phosphate	DEAE-dextran	Cationic Liposome	Electroporation	Biolistic Particle Delivery
Transient expression	+	+	+	+	+
Stable expression	+	Not efficient	+	+	+
Oligonucleotide transfer	Not efficient	Not efficient	+	+	+
RNA transfer	Not efficient	+	+	+	+
In vivo applications	Not efficient	Not efficient	+	Not efficient	+

One modification of this protocol replaces HBS with BES Buffered Saline at pH 6.95 ***(20,21)***. The precipitate forms slowly after addition of the DNA to the cell culture medium, during exposure to 2–4% CO_2 concentration in the incubator. This method is suitable for stable transfections using supercoiled DNA, for fibroblasts, and for epithelial cells.

A second modification of the basic procedure uses a higher calcium concentration. In the presence of increased calcium, the optimal precipitation interval is decreased ***(22)***.

Controlling the precipitate size is critical for successful, reproducible transfections using any of the calcium phosphate coprecipitation methods. Smaller precipitates have been correlated with higher transfection efficacy ***(22)***. Precipitate size is sensitive to the effects of calcium concentration, DNA concentration, pH, temperature of the reagents during the precipitation step, and how the buffered saline and DNA solutions are mixed. In order to maintain a high DNA concentration, a functionally inert carrier plasmid DNA, or high-molecular-weight genomic DNA, can be included ***(23)***.

Advantages:

- The reagents are readily available, inexpensive, and easy to prepare. Tested premixed reagents are available from biotechnology suppliers such as Promega (Madison, WI) (cat. no. E1200), and Sigma (St. Louis, MO) (cat. no. CAPHOS).
- Calcium phosphate coprecipitation functions with a wide variety of standard cultured cell lines, and is suitable for generating stable transfectants.
- Calcium phosphate coprecipitation methods do not require specialized equipment.

Limitations:

- Calcium phosphate coprecipitation is not appropriate for in vivo applications.
- Some cell types that are not transfected efficiently by this technique.

2.2. DEAE-Dextran

Theory: Diethylaminoethyldextran, or DEAE-dextran, is a polymeric, polycationic reagent that binds to and transfects negatively charged nucleic acids ***(1,24)***. The DEAE-dextran/DNA complex may be taken-up by the cell via endocytosis or phagocytosis ***(25,26)***. Chloroquine treatment increases expression of transfected DNA in some cell types, presumably by neutralizing the pH of lysosomes in which DNA becomes internalized ***(26)***.

Protocol and Critical Parameters: A standard procedure involves the mixing from 0.1 to 0.5 mg/mL DEAE-dextran (MW of approx 500,000 Daltons) with 5–30 µg/mL DNA in a phosphate buffered saline (PBS) solution. The solution is added to washed cells and incubated for 30 min at 37°C. After this time, the cells are overlaid with complete growth medium (with or without 80 µ*M* chloroquine) for up to 2.5 h, at which time the media is replaced with fresh, complete growth medium.

A modifed protocol by Al-Moslih *(27)* recommends exposing cells first to a solution of 1 mg/mL DEAE-dextran in PBS for 9 min at room temperature. The DEAE-Dextran solution is removed, DNA in PBS is added and incubated with the cells at 37°C for 30 min. Complete serum-containing medium is added, with or without chloroquine.

It is critical that the cellular morphology be closely monitored during the transfection period because DEAE-dextran can be very toxic to cells.

Advantages:

- The reagents used in the protocol are inexpensive and easy to prepare. Tested and pre-mixed reagents are available from biotechnology suppliers such as Promega (cat. no. E1210) and Sigma (cat. no. DEDEX).
- Specialized equipment is not required.

Limitations:

- This reagent is not efficient for generating stable transfectants.
- This method is not suitable for in vivo gene transfer.

2.3. Cationic Liposomes

Theory: The first synthetic cationic lipid published as a nucleic acid delivery reagent was 1,2-dioleoyloxypropyl-3-trimethyl ammonium bromide (DOTMA) developed by Felgner and colleagues *(28)*. Since then, several different types of synthetic cationic lipids for transfection have been designed along similar themes to DOTMA, having the general structure of a mono- or polycationic head group, chemically linked to a lipophilic moiety. The cationic region of the molecule associates with the negative charges of nucleic acid. An overall positive or neutral charge of the lipid:DNA complexes is generally correlated with higher transfection efficacy for cultured cells in vitro. It is hypothesized that the net neutral or positive charge effectively reduces electrostatic repulsion between the nucleic acid and the negatively charged cellular membrane. The lipid:DNA complexes appear to be endocytosed into cells. Nucleic acid entrapment in endosomal vesicles can be relieved by the presence of other lipids, such as dioleoylphosphatidylethanolamine (DOPE) in the liposome formulation *(29)*.

Protocol and Critical Parameters: Cationic liposomes and nucleic acids are mixed, generally in serum-free media, and allowed to associate for approximately 15 min. The lipid:DNA complexes are added to cells and incubated for 1–24 h, and complete serum-containing medium is added with or without removal of the transfection mixture. Specific recommendations for transfection conditions differ based upon the specific type of lipid used and the supplier's instructions.

For any liposome to be used for transfection, there are three essential and critical parameters requiring optimization. Every different cell type tested requires that an optimal amount of DNA be determined in combination with the concentration of cationic lipid. An optimal window of DNA and lipid

concentration exists for a particular type of cell, with excessive toxicity at higher concentrations and reduced efficacy at lower concentrations. The recommended amounts of lipid: DNA ratios vary for the different commercially available lipids. Another parameter that requires optimization is the length of the transfection interval. Depending upon the specific reagents and conditions tested, 1–24 h are suggested time intervals. Longer incubations are generally more toxic.

Advantages:

- The protocol for using cationic liposomes is relatively easy; often the reagents are mixed in a single tube prior to transfection.
- Many different cationic lipids are commercially available for gene transfer applications.
- Cationic liposome-mediated transfection can be used to generate transient and stable transfectants.
- The method is versatile for the delivery of nucleic acids. RNA is effectively transferred to cells ***(30)***, and DNA in sizes ranging from oligonucleotides to yeast artificial chromosomes can be transfected successfully ***(31,32)***.
- Transfection with cationic liposomes is effective with some types of cells that are difficult to transfect with other methods ***(33)***.
- The reagents can be used for in vivo gene transfer. Clinical trials have employed liposomes for gene delivery ***(2)***.

Limitations:

- The number and range of parameters for optimization can be daunting.
- The reagents are more costly than calcium phosphate or DEAE-dextran chemicals.

3. Physical Methods

Physical methods of nucleic acid delivery to cells rely upon penetrating the cell membrane for entry of the macromolecules. Compromising the integrity of the membrane by electrical shock or electroporation is the most commonly used physical method for gene transfer to cell populations. Biolistic particle delivery is a promising method for delivery of nucleic acid for gene delivery to animals, especially for vaccination purposes. Microinjection is used to introduce macromolecules to a small number of cells. This method is tedious, but, with skill, can result in the transfer of DNA into virtually all the injected cells ***(34)***. Other physical methods are also being developed, such as shock wave permeabilization ***(35)***. The more commonly used physical methods of electroporation and biolistic particle delivery will be further discussed.

3.1. Electroporation

Theory: Electroporation uses electric field pulses to create holes and pores in the cellular membrane, thus making the cell permeable to macromolecules. The electrical pulses decay exponentially, or are controlled to produce square

wave patterns, depending upon the type of capacitors and control units used for electroporation. The cell permeabilization is reversible, and the pores reclose during recovery after the pulses (for general reviews, *see* **refs. *36,37***).

Protocol and Critical Parameters: Cells are suspended at a concentration of 10^6–10^7 cells/mL in electroporation buffer, such as calcium and magnesium-free phosphate buffered saline, HEPES buffered saline, or serum-free culture medium. (Adherent cells first need to be trypsinized.) Cells are transferred to a sterile cuvet and DNA is added. For transient expression experiments, supercoiled DNA is added at 10–40 µg/mL; for stable transfection experiments, linear DNA is added at 1–20 µg/mL. The cells and DNA are electroshocked one or more times, and cells are plated after a short recovery period.

The most critical parameters to optimize for any cell type are the magnitude of the voltage and duration of the current pulse. Combinations of relatively high voltages with a short time constant, or lower voltages with longer time constants, are effective. The optimal conditions depend upon the inherent characterisitics of the cell, together with conductivity of the electroporation buffer, and the temperature. Electroporation of cells can be carried-out at room temperature or on ice, but the optimal conditions for a particular cell need to be empirically determined ***(38)***. Excessive voltage and/or duration of pulses result in cell death. The balance between cell death and efficacy of gene transfer should be monitored during the establishment of optimal conditions for electroporation of any given cell type. Optimal transfection may occur under conditions that cause approximately 50% cell death.

Advantages:

- Electroporation is effective for either transient or stable transfections. Transient applications can utilize supercoiled plasmid DNA. Stable transfections are more effective when linear DNA is electroporated ***(39)***. The conditions for electroporation can be adjusted to result in many or single copies of the recombinant gene per cell.
- Many different types of cells can be transfected by this method, including those that are difficult to transfect with chemical reagents ***(39–41)***. This is often the method of choice for transfection of plant protoplasts ***(42)***.
- Commercial apparatuses for electroporation are available from several biotechnology supply companies including Life Technologies (Gaithersburg, MD), Bio-Rad (Hercules, CA), and BTX (San Diego, CA). Disposable cuvettes are also available.

Limitations:

- Empirical optimization to determine the voltage, duration, and number of pulses is required for each cell type.
- Electroporation is not useful for in vivo applications.
- Higher numbers of cells and more DNA are required for electroporation compared to chemical methods for transfection.

3.2. Biolistic Particle Delivery

Theory: DNA is precipitated onto microscopic gold particles. The gold/DNA particles are projected into the target cell using a specialized device sometimes referred to as a "gene gun." Within the cell, the nucleic acid dissociates from the gold particles.

Protocol and Critical Parameters: Several different types of instruments have been designed for this application. Nucleic acids are first precipitated onto gold microcarrier beads and deposited onto a Mylar sheet. The gold/DNA or gold/RNA particles are placed in the device, accelerated by a high voltage discharge, and projected onto the cells or tissue site of choice. Commercially available devices use a low pressure helium pulse to accelerate the gold microcarrier beads. With these devices, samples can be prepared using tubing that is precoated with gold particles.

Advantages:

- Many different types of cultured cell lines and primary cells have been transfected by the biolistic particle delivery approach ***(43–45)***.
- This method has been used for delivery of DNA in vivo. Delivery of DNA has been successful to organs such as liver and muscle. The epidermis is a targeted area for the injection of DNA for vaccine development ***(46)***.
- Plant cells that are rather difficult to transfect by other methods have been transfected by biolistic particle delivery ***(47)***.
- The Helios Gene Gun System is available commercially from BioRad (Hercules, CA).

Limitations:

- The equipment for biolistic particle delivery is expensive.

References

1. Vaheri, A. and Pagano, J. S. (1965) Infectious poliovirus RNA: a sensitive method of assay. *Virology* **27,** 434–436.
2. Felgner, P. L., Tsai, Y. J., Sukhu, L., Wheeler, C. J., Manthorpe, M., Marshall, J., and Cheng, S. H. (1995) Improved cationic lipid formulations for in vivo gene therapy. *Ann. New York Acad. Sci.* **772,** 126–139.
3. Graham, F. L. and van der Eb, A. J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52,** 456–467.
4. Groskreutz, D. and Schenborn, E. (1997) Reporter systems, in *Recombinant Protein Protocols: Detection and Isolation* (Tuan, R., ed.), Humana, Totowa, NJ, pp.11–28.
5. Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340,** 245–246.
6. Fearon, E. R., Finkel, T., Gillison, M. L., Kennedy, S. P., Casella, J. F., Tomaselli, G. F., Morrow, J. S., and Dang, C. V. (1992) Karyoplasmic interaction selection strategy: a general strategy to detect protein-protein interactions in mammalian cells. *Proc. Natl. Acad. Sci. USA* **89,** 7958–7962.

7. Cheng, L., Fu, J., Tsukamoto, A., and Hawley, R. G. (1996) Use of green fluorescent protein variants to monitor gene transfer and expression in mammalian cells. *Nature Biotechnol.* **14,** 606–609.
8. Southern , P. J. and Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Gen.* **1,** 327–341.
9. Cheng, Y. C., Huang, E. S., Line, J. C., Mar, E. C., Pagano, J. S., Dutschman, G. E., and Grill, S. P. (1983) Unique spectrum of activity of 9-[(1,3-di-hydroxy-2-propoxy)methyl]guanine against herpes viruses in vitro and its mode of action against herpes simplex virus type 1. *Proc. Natl. Acad. Sci. USA* **80,** 2767–2770.
10. Wilson, T., Paphadjopoulos, D., and Taber, R. (1979) The introduction of poliovirus RNA into cells via lipid vesicles (liposomes). *Cell* **17,** 77–84.
11. Leopold, L. H., Stone, S. K., Newkirk, T. A., Reddy, R. M., and Reddy, E. P. (1995) Multi-unit ribozyme-mediated cleavage of bcr-abl mRNA in myeloid leukemias. *Blood* **85,** 2162–2170.
12. Haensler, J. and Szoka, F. C., Jr. (1993) Polyamidoamine cascade polymers mediate efficient transfection of cells in culture. *Bioconj. Chem*. **4,** 372–379.
13. Kukowska-Latallo, J. F., Bielinska, A. U., Johnson, J., Spindler, R., Tomalia, D. A., and Baker, J. R. (1996) Efficient transfer of genetic material into mammalian cells using Starburst polyamidoamine dendrimers. *Proc. Natl. Acad. Sci. USA* **93,** 4897–4902.
14. Farber, F. E., Melnick, J. L., and Butel, B. S. (1975) Optimal conditions for uptake of exogenous DNA by Chinese hamster lung cells deficient in hypoxanthine-guanine phosphoribosyltransferase. *Biochim. Biophys. Acta* **390**, 298–311.
15. Bond, V. C. and Wold, B. (1987) Poly-L-ornithine-mediated transformation of mammalian cells. *Mol. Cell Biol.* **7,** 2286–2293.
16. Wu, G. Y. and Wu, C. H. (1987) Receptor-mediated in vitro gene transformation by a soluble DNA carrier system. *J. Biol. Chem.* **262,** 4429–4432.
17. Boussif, O., Lezoualch, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., and Behr, J.-P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and *in vivo*: polyethylenimine. *Proc. Natl. Acad. Sci. USA* **92,** 7297–7301.
18. Loyter, A., Scangos, G. A., and Ruddle, F. H. (1982) Mechanisms of DNA uptake by mammalian cells: fate of exogenously added DNA monitored by the use of fluorescent dyes. *Proc. Natl. Acad. Sci. USA* **79**, 422-426.
19. Wilson, S. P. and Smith, L. A. (1997) Addition of glycerol during DNA exposure enhances calcium phosphate transfection. *Anal. Biochem.* **246,** 148–150.
20. Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7,** 2745–2752.
21. Chen, C. A. and Okayama, H. (1988) Calcium phosphate-mediated gene transfer: A highly efficient transfection system for stably transforming cells with plasmid DNA. *Biotechniques* **6,** 632–638.
22. Jordan, M., Schallhorn, A., and Wurm, F. M. (1996) Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucl. Acids Res.* **24,** 596–601.

23. Strain, A. J., Wallace, W. A. H., and Wyllie, A. H. (1985) Enhancement of DNA-mediated gene transfer by high-Mr carrier DNA in synchronized CV-1 cells. *Biochem. J.* **225,** 529–533.
24. McCutchan, J. H. and Pagano, J. S. (1968) Enhancement of the infectivity of Simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* **41,** 351–357.
25. Yang, Y. W. and Yang, J. C. (1997) Studies of DEAE-dextran-mediated gene transfer. *Biotechnol. Appli. Biochem.* **25,** 47–51.
26. Luthman, H. and Magnusson, G. (1983) High efficiency polyoma DNA transfection of chloroquine treated cells. *Nucl. Acids Res.* **11,** 1295–1308.
27. Al-Moslih, M. I. and Dubes, G. R. (1973) The kinetics of DEAE-dextran induced cell sensitization to transfection. *J. Gen. Virol.* **18,** 189–193.
28. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M., and Danielsen, M. (1987) Lipofection: a highly efficient, lipid mediated DNA transfection procedure. *Proc. Natl. Acad. Sci. USA* **84,** 7413–7417.
29. Farhood, H., Sevina, N. S., and Huang, L. (1995) The role of dioleoylphosphatidylethanolamine in cationic liposome mediated gene transfer. *Biochim. Biophys. Acta* **1235,** 289–295.
30. Malone, R. W., Felgner, P. L., and Verma, I. M. (1989) Cationic liposome-mediated RNA transfection. *Proc. Natl. Acad. Sci. USA* **86,** 6077–6081.
31. Felgner, J., Bennett, R., and Felgner, P. L. (1993) Cationic lipid mediated delivery of polynucleotides. *Methods* **5,** 67–75.
32. Lee, J. T. and Jaenisch, R. (1996) A method for high efficiency YAC lipofection into murine embryonic stem cells. *Nucl. Acids. Res.* **24,** 5154–5155.
33. Muller, S. R., Sullivan, P. D., Clegg, D. O., and Feinstein, S. C. (1990) Efficient transfection and expression of heterologous genes in PC12 cells. *DNA Cell Biol.* **9,** 221–229.
34. Capecchi, M. R. (1980) High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* **22,** 479–488.
35. Lauer, U., Burgelt, E., Squire, Z., Messmer, K., Hofschneider, P. H., Gregor, M., and Delius, M. (1997) Shock wave permeabilization as a new gene transfer method. *Gene Therapy* **4,** 710–715.
36. Shigekawa, K. and Dower, W. J. (1988) Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *BioTechniques* **6,** 742–751.
37. Weaver, J. C. and Chizmadzhev, Y. A. (1996) Theory of electroporation: a review. *Bioelectrochem. Bioenerget.* **41,** 135–160.
38. Chu, G., Hayakawa, H., and Berg, P. (1987) Electroporation for the efficient transfection of mammalian cells with DNA. *Nucl. Acids Res.* **15,** 1311–1326.
39. Potter, H., Weir, L., and Leder, P. (1984) Enhancer-dependent expression of human k-immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc. Natl. Acad. Sci. USA* **81,** 7161–7165.
40. Wong, T. K. and Neumann, E. (1982) Electric field mediated gene transfer. *Biochem. Biophys. Res. Comm.* **107,** 584–587.

41. Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P. H. (1982) Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* **7,** 841–845.
42. Fromm, M., Taylor, L. P., and Walbot, V. (1985) Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc. Natl. Acad. Sci USA* **82,** 5824–5828.
43. Yang, N. S., Burkholder, R., Roberts, B., Martinell, B., and McCabe, D. (1990) In vivo and in vitro gene transfer to mammalian somatic cells by particle bombardment. *Proc. Natl. Acad. Sci. USA* **87,** 9568–9572.
44. Cheng, L., Ziegelhoffer, P. R., and Yang, N. S. (1993) *In vivo* promoter activity and transgene expression in mammalian somatic tissues evaluated by using particle bombardment. *Proc. Natl. Acad. Sci. USA* **90,** 4455–4459.
45. Qiu, P., Ziehelhoffer, P., Sun, J., and Yang, N. S. (1996) Gene gun delivery of mRNA in situ results in efficient transgene expression and genetic immunization. *Gene Therapy* **3,** 262–268.
46. Haynes, J. R., McCabe, D. E., Swain, W. F., Widera, G., and Fuller, J. T. (1996) Particle-mediated nucleic acid immunization. *J. Biotechnol.* **44,** 37–42.
47. Klein, T. M., Wolf, E. D., Wu, R., and Sanford, J. C. (1987) High velocity microprojectiles for delivering nucleic acids into living cells. *Nature* **327,** 70–73.

9

The Use of Particle-Mediated Gene Transfer for the Study of Promoter Activity in Somatic Tissues

Charles M. Nicolet and Ning-Sun Yang

1. Introduction

Research in molecular genetic studies and in transgenic biotechnology has resulted in excellent advancements in exploiting gene transfer techniques. Basic research in transcriptional and/or translational control, protein engineering, cellular immunological mechanisms, and developmental genetic processes using transient or stable transgenic approaches have generated the main body of information in our current understanding of various molecular biology and cell biology disciplines. Parallel with this development, the gene transfer and transgene expression techniques have also been extremely useful in their application to plant and animal (including mammalian) biotechnologies. Specifically, analyses of promoter activity and usage have played important roles in generation of genetically engineered plants, transgenic animals, and recombinant or transgenic protein products in various microbial fermenters or plants or in mammalian cell bioreactors. More recently, various gene transfer and promoter usage technologies have also been evaluated extensively for application to gene therapy research.

Gene therapy has potential value in the treatment of a variety of diseases. Gene therapy strategies have been proposed in both treatment and vaccination for infectious diseases, for cancer, and for correction of inborn metabolic disorders. Effective therapeutic strategies are expected to require extensive knowledge of the gene expression patterns directed by various promoters in different tissues, and other factors that influence the expression levels of introduced transgenes. It has been a technical challenge to develop systems that address these issues in a more biologically and physiologically relevant fashion. For example, many current gene delivery strategies rely on viral systems that

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

require recipient cells to be actively dividing, a condition difficult to achieve in most differentiated tissues. Several nonviral physical–chemical techniques such as electroporation, liposome-mediated transfection, or calcium–phosphate transfection, are effective in many in vitro systems, but are often untenable or too complex for in vivo applications. Ideally, one would like to employ a technique that could be broadly applied to in vitro, ex vivo (i.e., clinical or primary cell samples), and in vivo applications, and that could allow the rapid assessment of numerous expression parameters in a biologically relevant fashion (e.g., in intact organs, fresh tissue explants, or related cell explant systems). This would more closely mimic a clinical, in vivo gene therapy approach.

One gene transfer approach that has shown great potential in this regard is the particle-mediated gene delivery technology. This method utilizes physical means to intracellularly introduce naked DNA into cells, and so is virtually independent of cell membrane structures, replication capacity of the cell, presence of particular membrane receptors, or other factors not under control of the investigator or clinician. Particle mediated gene transfer (PMGT), also known as biolistic, ballistic, gene gun, or microprojectile gene transfer, employs microscopic gold particles (1–3 μm in diameter) as delivery vehicles to introduce DNA into targeted cells ***(1,2)***. DNA is precipitated onto the gold beads by well-defined standard chemical means, then these beads are accelerated into target cells using a motive force generated by a shockwave. The delivery system discussed in this chapter utilizes a pressurized pulse of helium to accelerate the DNA-coated particles. Originally developed for transfection of plant cells, PMGT was later found to have widespread application to animal systems, as discussed in the present chapter.

Numerous experimental models have shown the broad based utility of PMGT (*see* ***3, 4, 5***). Because of the ability of PMGT to deliver genes into a wide range of biologically relevant systems, notably organs and other hard-to-transfect cell types, many of the models investigated are directly relevant to a gene therapy approach. This is particularly true in the context of promoter analysis, because it has been possible to conveniently and effectively analyze the activity of a number of different promoters in various target organs, thus providing a rational preclinical data set for use in designing related gene therapy protocols. **Fig. 1** shows an example of such an analysis in a rat model system ***(6)***. In vivo promoter activity was investigated by assaying the transient activity levels of a luciferase reporter gene fused downsteam of several different viral or mammalian cellular promoters. Five different rat tissues were transfected in vivo and *in situ* with these different promoter constructs. The data show that all organ types gave readily detectable levels of expression, although clear differences in the transcriptional capacity of various tissues was noted. Also, interesting and potentially important tissue-specific differences in

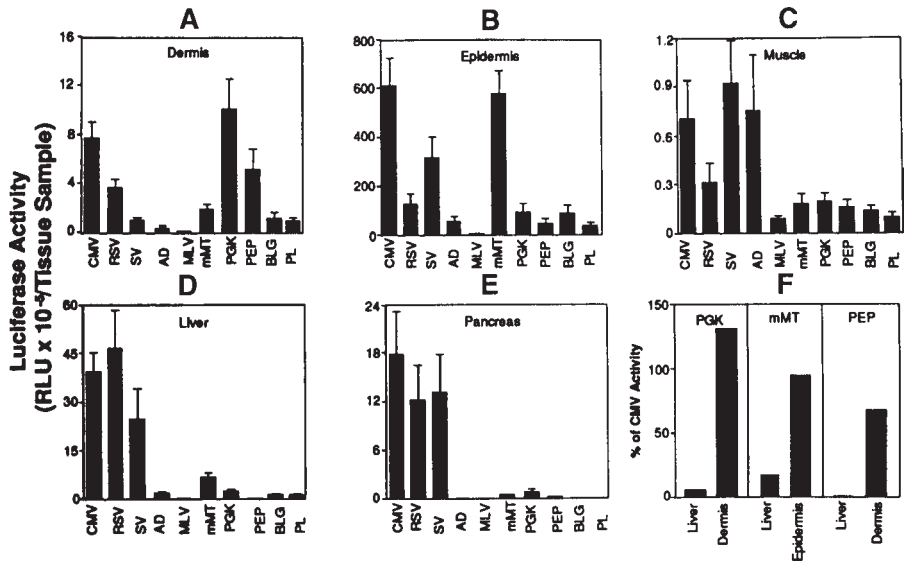


Fig. 1. Comparison of in vivo transgenic luciferase activities driven by various viral and cellular promoters in different rat tissues. Luciferase activity was obtained from tissue extracts of abdominal dermis (A), abdominal epidermis (B), abdominal muscle (C), liver (D), and pancreas (E) after gene delivery with pCMV-luciferase (cytomegalovirus immediate/early promoter, CMV), pRSV-luciferase (Rous sarcoma virus LTR, RSV), pAD-luciferase (adenovirus 2 major late promoter, AD), pMLV-luciferase (murine leukemia virus LTR, MLV), pmMT-luciferase (murine metallothionein promoter, mMT), pPGK-luciferase (murine phosphoglycerate kinase promoter, PGK), pPEP-luciferase (murine phosphoenolpyruvate carboxykinase promoter, PEP), pBLG-luciferase (bovine beta-lactoglobulin promoter, BLG), or pPL-luciferase (bovine prolactin promoter, PL). (F) Relative activity of different promoters in various tissues compared to pCMV-luciferase, employed as an internal standard (expression defined as 100%) for each test tissue. Graphed values represent an average of the luciferase activity ± SEM per target tissue. Six to 12 separate tissue samples were collected from three to six experimental rats, with no more than two samples from each animal. Control tissue samples that were recipients for naked gold beads (no coated DNA) showed no detectable luciferase activity.

promoter activity were observed ***(6)***. To our knowledge, related data have not yet been obtainable by nongene gun delivery methods.

The PMGT procedure also provides a convenient assay system to investigate other basic mechanisms of gene expression in cell biological and physiological systems. An example of this was a study investigating the functional regions of the Harvey-ras promoter in mammary epithelial tissue explants ***(7)***. Different deleted derivatives of the Ha-ras promoter region fused upstream to a

Table 1
Experimental Systems Utilizing the Gene Gun

Species	Tissue Type	Reference
Bombyx mori	Silk gland	***(8)***
Xenopus	Embryos	**(9)**
Newt	Regenerating limb	***(10)***
Rat	Mammary epithelial cells	***(7,11)***
Rat	Fetal and adult brain	***(12)***
Rat	Oligodendrocytes	***(13)***
Rat	Hepatocytes	***(6)***
Mouse	Macrophages, splenocytes	***(14)***
Mouse	Epidermis	***(15)***
Canine	Oral buccal mucosa, epidermis	***(16)***
Porcine	Epidermis	***(17)***
Human	T cells, peripheral blood mononuclear cells	***(14)***
Human	Prostate tumor xenograft in nude mouse	***(18)***
Human	Freshly excised tumor tissue	***(19)***

luciferase reporter gene were introduced into primary mammary epithelial tissue by PMGT, then the expression quantitated. This allowed a dissection of the important regulatory regions of an oncogene in primary cells that are a target for transformation by that same oncogene. Short of creating transgenic animals using the same deleted promoter derivatives, this information might not have been generated otherwise. **Table 1** summarizes some of the applications of PMGT in a variety of experimental systems.

Currently, the only commercially available gene gun for in vivo gene transfer is distributed by Bio-Rad (Hercules, CA) and is known as the Helios device ***(20)***. Bio-Rad also distributes most of the reagents and equipment used in the preparation of the gold beads and tubes. The basic concept, methods, and procedures for particle delivery utilized by the Bio-Rad system closely follow, but are not necessarily identical to, the protocols for the Accel® device (designed by Dr. Dennis McCabe) that we and many other academic investigators have used in the experiments discussed in this chapter. For either the Helios or Accel devices, gene delivery parameters that yield satisfactory results must be determined initially and fine-tuned in each laboratory, for each model system under study. This chapter should serve as a reference guide, in combination with technical literature from Bio-Rad, and of course the user's own observations.

Many of the procedures and protocols discussed in the sections were closely adapted from the Accel® standard operation procedures manual. This manual was developed by a research team at the Auragen/Geneva Co., Middleton, WI,

where one of the authors (N.-S. Yang) previously worked as a key investigator in the cancer gene therapy group.

2. Materials

2.1. Special Equipment

1. Gene gun and supplies: The only commercially available devices for particle-mediated gene transfer are distributed by Bio-Rad. The devices are the Helios gene gun and Biolistic PDS-1000/He chamber. Both utilize the principle of a motive force propelling DNA-coated gold beads into tissue, but differ in certain details. As the Helios device is directly analogous to the Accel® gene gun, the hardware with which we have the most experience, we have limited the subsequent discussion to this type of device. All specialized materials required for using the Helios gene gun can be purchased as a complete system, and the various consumable components can be purchased individually as well. The complete system includes:
   a. the gene "gun," or particle acceleration hardware.
   b. the tube turner, used for preparation of cartridges.
   c. supplies for attachment to the helium tank (generating the motive force for particle delivery) and the nitrogen tank (used in tube preparation).
   d. Tefzel tubing, gold particles of different sizes (called microcarriers in the catalog).
   e. various accessories, including a tube cutter, extra O-rings, dessicant pellets, and gene transfer optimization kits.
2. Ear protection, for operator of gene gun.
3. Ultrasonic cleaner: for dispersing particle clumps.
4. Helium and nitrogen tanks.
5. Luminometer: To be used if luciferase is used as reporter system.

2.2. Solutions/Reagents

1. 100% EtOH: stored at –20°C.
2. 50 m*M* Spermidine: stored at –20°C.
3. 1 *M* $CaCl_2$: stored at room temperature.
4. PVP: 1 mg/mL in 100% EtOH, stored at –20°C (*see* **Note 1**).
5. DNA preparation reagents (*see* **Note 2**).
6. Reporter plasmids: ß-galactosidase and luciferase expression constructs are available from numerous suppliers, including Promega (Madison, WI) and Clontech (Palo Alto, CA) (*see* **Note 3**).
7. Luciferase assay reagents: These are commercially available from Promega. Activity can also be assayed using noncommercial reagents ***(12,13)***.
8. 1.5 % glutaraldehyde.
9. X-gal staining solution: 400 µg X-gal, 1 m*M* $MgCl_2$, 5 m*M* potassium ferricyanide in phosphate buffered saline (PBS).

10. Tissue extraction reagent: PBS with 0.5% Triton X-100 and protease inhibitors (e.g., 1 mM PefaBloc [Boehringer-Mannheim, Indianapolis, IN] or 10 µg/mL aprotinin).
11. Tissue dissociation solution: 200 U/mL collagenase I, 200 U/mL hyaluronidase, 100 U/mL DNAse I in Dulbecco's modified eagle medium (DMEM).
12. Animal husbandry (*see* **Note 4**).

3. Methods

3.1. Cartridge Preparation

3.1.1. Bead Preparation

1. Weigh out gold particles in a 1.5 mL microfuge tube. The amount depends on bead-loading rate (*see* **Note 5**).
2. Add 100 µL 50 m*M* spermidine, resuspend gold thoroughly by vortexing.
3. Sonicate gold–spermidine in an ultrasonic cleaner for 3–5 s.
4. Add the DNA in volume less than 100 µL. Amount depends on DNA loading rate (*see* **Note 5**).
5. Mix DNA, spermidine, and gold thoroughly.
6. Add 100 µL 1 *M* $CaCl_2$ dropwise while vortexing the tube.
7. Allow the mixture to precipitate at room temperature for 10 min.
8. Briefly spin to pellet gold. Pour off the supernatant.
9. Resuspend the pellet in remaining supernatant, wash pellet three times with 1 mL 100% EtOH using a short spin between each wash. Between washes, pour off supernatant, resuspend the pellet in remaining supernatant before adding larger volume of EtOH.
10. Resuspend the pellet finally in 1 mL 100% EtOH.
11. Pipet the 1 mL EtOH bead suspension into a volume of EtOH to get the desired bead-loading rate (*see* **Note 5**).
12. This suspension is ready for tube preparation, or it can be sealed and stored at –20°C.

3.1.2. Tube Preparation

1. Purge tubing with nitrogen at 0.5 L/min (LPM) in the tube turner apparatus for at least 15 min to remove all water .
2. Cut an approx 28-in. length of tubing with bevelled ends, and attach one end to a syringe fitted with an adapter piece that can slide over the tubing.
3. Bead suspension should be at the appropriate volume (e.g., 3.5–4 mL for one 28-in. tube) and at room temperature. PVP can be added at this point, to the final desired concentration (*see* **Note 1**). Sonicate suspension with ultrasonic cleaner then vortex prior to next step.
4. Draw up bead suspension into 28-in. length of tubing using the syringe fitted to the tubing.
5. Slide tubing containing the suspension into the tube turner. MAKE SURE NITROGEN IS OFF.
6. Allow beads to settle for 3 min. Draw off EtOH with syringe slowly and evenly.

7. Begin rotation of the tube turner, proceed for 30 s with nitrogen still off.
8. Turn nitrogen on to 0.3–0.4 LPM while turners still rotating.
9. Allow the tube to dry completely, 3–4 min, with turner rotating.
10. Remove from turner, and cut off ends of tubes that do not show even distribution of beads. Cut remaining tubing into 0.5 in. tubes with a razor blade, scalpel, or the Bio-Rad cutter.
11. Carry out quality control on tubes (*see* **Notes 5** and **6**).
12. Store tubes at 4°C with a dessicant pellet (available from Bio-Rad) in a sealed vial. These cartridges are good for several months.

3.2. Particle Delivery

In the following section we discuss several PMGT protocols that we have used to analyze different somatic tissues in a variety of situations. These protocols have been chosen because they are considered to be applicable to numerous other experimental models assaying gene expression in different mammalian somatic tissues. The apparatus should be set up prior to animal or final tissue preparation. The tubes are loaded into the cartridge, the cartridge loaded into the gun, the gun plugged into the electrical outlet and connected to the helium, and the regulator adjusted to the appropriate pressure (*see* **Note 7**).

3.2.1. Murine Epidermis

1. Anesthetize the animal. For mouse skin or externally exposed tissue this is usually not required.
2. Clip fur over delivery area using Oster clippers with #40 blade.
3. A commercial depilatory such as Nair used over the clipped area can increase expression. If Nair is used, the animal will have to be anesthetized (to prevent them from licking or smearing the cream). Wipe the depilatory cream thoroughly, then further clean the area with warm water and blot dry with gauze.
4. Place barrel of loaded gun (i.e., loaded with the cartridges containing the DNA coated gold beads) against skin and discharge (*see* **Note 8**).

3.2.2. Canine Oral Mucosa.

1. Anesthetize the animal.
2. Wipe dry the oral buccal mucosa with gauze.
3. Place the barrel of loaded gun onto the target tissue and discharge.

3.2.3. Human Tumor Cells (*see* **Note 9**)

1. Mince solid tumor tissues in small fragments in approx 2 mL of tissue dissociation enzyme solution.
2. Transfer mixture to a 50-mL flask containing 4-mm glass beads and 20 mL of enzyme solution. Incubate for 20 min at 37°C and 5% CO_2.
3. Separate undigested tissue and clumps from cell suspension with a 30-mesh stainless steel screen, and place clumps in fresh enzyme solution for additional digestion.

4. Repeat up to three times as necessary to dissociate sufficient numbers of cells from the tissue clumps. Wash and resuspend in culture medium, and measure cell number and viability.
5. Place cells at concentration of 5×10^7 cells/mL.
6. Prewet 35 mm cell culture dish with medium, then remove medium.
7. Place 20 µL of cell suspension (1×10^6 cells total) onto the wetted surface of the dish; this volume often spreads out to an area almost exactly the size of the gene gun delivery site.
8. Place barrel up to the cell suspension on dish; discharge gun.
9. Immediately add 2 mL culture medium, incubate for desired time to assay expression (*see* **Subheading 3.3.** for discussion of time course).

3.2.4. In vivo Delivery to an Organ

1. Perform surgery on animal to expose organ.
2. Wipe barrel of gene gun with EtOH to sterilize.
3. Place barrel of gun against organ, discharge.
4. Complete surgery.

3.3. Analysis of Expression

Choose appropriate time course to analyze expression following delivery (*see* **Note 10**).

3.3.1. Tissue Lysis

1. Harvest the tissue sample that served as a target for gene delivery. Suspension cells can be spun down, adherent cells can be trypsinized and pelleted. Following delivery to skin or other solid tissues, a region encompassing the delivery site is surgically removed following euthanasia of test animals. Alternatively, skin punch biopsies can be taken following gene delivery into larger animals or if time course studies are to be carried out.
2. Snap freeze tissues sample immediately following harvest.
3. Add tissue extraction buffer to frozen tissue, start scissor-mincing right away (*see* **Note 11**). Typically, we add 0.5 mL of extraction buffer to a 1-cm-diameter skin tissue patch in a 1.5-mL microfuge tube.
4. Homogenize tissue as thoroughly as possible with sharp surgical scissors, pellet tissue, re-mince the pellet.
5. Remove debris by centrifugation in a microfuge, assay supernatant for transgenic product. For example, this crude tissue extract can be used directly in luciferase assays or in an ELISA.

3.3.2. Histochemistry and Immunohistochemistry (*see* **Note 12**)

1. For visualization of transgenic β-galactosidase activity, prepare 10 µm tissue sections and fix in 1.5% glutaraldehyde.
2. Incubate fixed slides in X-gal staining solution for 1–8 h at 37°C. The length of time will depend on the amount of activity present.

Table 2
Troubleshooting Particle Mediated Gene Transfer

Problem	Possible Cause	Solution
Poor or No Expression	Nonoptimal ballistic parameters	Confirm appropriate pressures and particle densities are being used
	Incorrect or degraded vector	Confirm expression vector is correct and intact
	Inappropriate DNA solution	Make sure DNA is resuspended in TE or distilled water, not PBS
	Gold was not delivered	Make sure cartridges are loaded in chamber and properly positioned for firing
Gold does not penetrate skin/organ	Inadequate discharge pressure	Check helium supply, increase discharge pressure
	Insufficient particle momentum	Use larger gold particles
Problems with hardware gene gun device, usage, the DNA/gold particles/cartridge assembly		Consult the Bio-Rad user's manual or technical services representative

3.3.2. Bioassay of Immunological and/or Cell Biological Activities

1. Collect culture supernatants from in vitro or ex vivo transfected cells (e.g., culture supernatants of dissociated tumor cells or normal tissues).
2. Dilute culture supernatants and apply to appropriate responder cells (*see* **Note 13**).
3. Cell lysates prepared using tissue extraction buffer are NOT suitable for most cell biological bioassays. (The Triton used in the extraction buffer is inhibitory to most cell functions even when diluted to very low levels.) For secreted products such as cytokines, culture supernatants are collected then used with no further manipulations.

3.4. Troubleshooting

3.4.1. See Table 2

4. Notes

1. Polyvinylpyrrolidone (PVP), avg. MW 360,000, is included with the Bio-Rad set up, and is used to help attach the beads to the inside wall of the Tefzel tubing. The gold particles will stay on without the PVP, but its use often helps ensure that the initial small pulse of helium gas going through the unit, which is less than the discharge pressure required to get tissue penetration, does not remove many of the gold particles in an unproductive fashion. That is, only the full discharge pressure will lift the particles from the tube and accelerate them into the recipient targeted tissue or cells. The concentration of PVP used in the final bead–ethanol slurry, prior to drawing up into the tube, may vary depending on the ultimate application of those tubes, but we generally begin with a concentration of 0.01 mg/mL, diluted from a concentrated stock of 1 mg/mL made in 100% EtOH. It is important that no water get into this solution, so it is advisable to generate new stock solutions every couple of months.
2. As with most gene transfection protocols, the purity and quality of the DNA preparation is an important parameter in the efficiency of gene expression. Plasmid DNA can be prepared by a number of standard techniques, including ion-exchange chromatography (as in the Qiagen system and numerous other related products), equilibrium density CsCl-EtBr gradients, or scale-up of the alkaline lysis protocol with additional organic solvent extractions. DNA has been used from all such purification procedures with satisfactory results. We primarily use the Qiagen Plasmid Maxi or Mega Preparation kits (Qiagen Inc., Chatsworth, CA).
3. The particular choice of reporter gene construct will depend on the system under study. For example, in the Promega system the option exists to clone any whole promoter region upstream of the luciferase gene or to clone just enhancer regions upstream of a minimal promoter that is already fused to luciferase. Luciferase is a very convenient transgenic product to assay for expression following PMGT delivery for a number of reasons: the luciferase assay is extremely sensitive; the assay is linear over a several log range of enzyme concentration; the luciferase mRNA and enzyme have short half-lives, so the activity measured at a given point is directly and proportionally indicative of the amount of transgene expression occurring at that point (i.e., there is not an ongoing accumulation of transcription or translation products); background activity in all tissues is virtually negligible; the assay is relatively straightforward if access to a luminometer is available; and enzymatic activity can be measured in a wide range of tissue lysis solutions. We have typically carried out experiments using luciferase fused to the CMV promoter, particularly when evaluating different gene delivery conditions.

   The use of ß-galactosidase as a reporter gene is particularly useful when histology is to be carried out in the targeted tissue, as the activity of this enzyme can be colorimetrically detected with the enzymatic substrate X-gal. It is a simple manipulation to stain tissue, organs, or cells with X-gal that have served as recipients for gene delivery using PMGT, in order to measure the pattern and fate of transgene expression.

4. A description of the requirements for animal work is beyond the scope of this chapter, but clearly depends on the investigational system in use and institutional requirements. Gene gun experiments have been carried out in a wide range of species including mouse, rat, dog, monkey, and pig.
5. An important consideration in PMGT is the dosage of DNA delivered into the cells. This is dependent on two parameters: bead-loading rate (BLR), the amount of gold beads present in each cartridge, and DNA-loading rate (DLR), the amount of DNA precipitated onto each bead. Recommended BLRs range between 0.25 and 1.0 mg particles/cartridge. The value chosen is determined partially by design and partially by experience, based on the target cell population and the results of trial experiments with reporter expression vectors. For example, higher particle densities are often used to transfect suspension cells due to the high cell densities in the aqueous smear, whereas adherent cells and intact tissues may require reduced particle densities to balance between effective transfection and minimal cell/tissue damage. For an 0.5 mg/tube (or target) BLR, typically used by our laboratory in a number of applications, the BLR is set at 7 mg gold/mL of EtOH. The 28- in. length of tubing used requires approx 4 mL of bead/EtOH slurry, corresponding to 28 mg of gold particles weighed out into the 1.5 mL microfuge tube.

   DLR can be varied over a fairly wide range of concentrations. At DLRs of from 0.25 to 2.5 µg DNA/milligram gold, the expression in target cells, both tissue culture lines and several in vivo systems, may show only small differences if the expression plateaus at lower values. We typically use 1–2 µg DNA/mg gold. At DNRs approaching 5 µg/mg gold, the precipitation of the DNA can result in bead clumping, which gives poor tube preparations. The similarities of gene expression over a range of DLR conditions underlie a major advantage of the particle delivery system: different plasmid DNAs can be effectively coprecipitated onto the same bead, and as a result a cell acquiring one plasmid is virtually certain to acquire the other. Thus, various cotransfection experiments can be carried out with the expectation that the expression level of one gene is constant even as the amount of DNA encoding that gene is changed to allow the deposition of a different plasmid onto the same bead.
6. Tube preparations can be examined by simple procedures that ensure the integrity of the DNA, and by carrying out test discharges onto surfaces that can be microscopically examined. To examine the DNA 10–15 µL of water or TE is delivered with a Pipetman (or related device) into the tube. By pumping the liquid up and down, and moving the liquid from one side of the tube to the other, most of the gold beads will be released from the side of the tubing (though at PVP concentrations higher than 0.01 mg/mL, release into the liquid is not efficient). The small aliquot of liquid is spun in a microfuge, then the supernatant, containing all the DNA resolubilized from the beads, is removed and the DNA in it visualized by agarose gel electrophoresis. This can reveal the integrity of the precipitated DNA, and also gives an indication of the efficiency of the precipitation reaction. Also, when testing the strength of different promoters following PMGT, it is important to demonstrate that the same amounts of DNA are being

delivered from each tube. This is readily accomplished by isolating the DNA from the different tube preparations and comparing their fluorescent intensities on gels.

Gold bead discharge patterns can be examined in two and three dimensions. The gun can be loaded and discharged onto a plastic surface such as a tissue culture dish, then this dish can be examined microscopically to ensure that an even spread of particles was effected. In addition, discharge can be carried out over a 1% agar in water gel, then a gel slice taken through the target area. The slice is mounted on a slide, and again examined microscopically to determine areas of high-, medium-, and low-bead density across the section.

7. Helium discharge pressure is dependent on a number of variables in the experimental system under study. Skin delivery requires high pressures to effect bead penetration into epidermal cell layers that can confer good expression. In contrast, penetration into soft organs (e.g., pancreas), dissected organoids (e.g., mammary gland organoids), and cell suspensions often require lower pressures. The exact pressure utilized also depends on the particular hardware configurations in use. The user manual for the Bio-Rad device should serve as a very good guide for initial pressures to attempt, with the optimal pressures determined empirically by the user.
8. Discomfort following skin treatment is apparently minimal. In general, macroscopic disruption of the skin, external bleeding, or hematoma are not observed. Some animals may show transient erythema or inflammation at the treatment site. After two to three days, some sloughing of the transfected skin layer may be evident.
9. The protocol outlined for disaggregated human tumor cells is also used for a variety of suspension cell and adherent (detached, counted, and resuspended accordingly) tissue culture cell lines that have served as the recipients for PMGT.
10. The time course of expression following PMGT is dependent on the reporter gene used, the sensitivity of the detection assay, and the nature of the recipient cells or tissue. Measurable reporter gene activity can often be noted at 6–8 h following DNA delivery. Expression in murine epidermis can be detected up to 3–4 days following PMGT, then it falls off rapidly. Different animal species will likely have distinguishable time course patterns that will have to be determined by the investigator. Expression has been noted in rat dermis to last for up to one year, but this would have to be considered a fairly exceptional case.
11. As with most of the procedures discussed, the volume of extraction buffer used will depend on the system under study. Regarding the delivery to murine epidermis, a tissue patch roughly 1 cm in diameter is resuspended in 0.5 mL of the tissue extraction buffer and processed as described.
12. A detailed discussion of immunohistochemistry methods is beyond the scope of this review, but we would note that this technique has been used succesfully to assay the transgenic expression of various cytokines in vivo in epidermal tissues that have served as the recipient for gene gun transfection.

13. The choice of responder cells used in a bioassay of specific proteins produced by the transfected cells (including ex vivo tumor cell samples) depends on the transgenic protein being expressed. We have used growth-factor-dependent cell lines to assay the presence of biologically active GM-CSF and IL-2 in target cells that have been transfected with expression constructs encoding these genes. The procedure works best when culture supernatants from transfected cells can be used. We have found it impossible to remove the Triton X-100 present in the tissue extraction buffer by either dialysis or ammonium sulfate precipitation, and also have been unable to dilute the Triton to the point where the responder cells remain viable but bioactivity is still detectable.

References

1. Klein, T., Wolf, E., Wu, R., and Sanford, J. (1987) High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature* **327,** 70–73.
2. McCabe, D.E., Swain, W.F., Martinell, B., and Christou, P. (1988) Stable transformation of soybean (*Glycine max*) by particle acceleration. *Bio/Technol.* **6,** 923–926.
3. Yang, N.-S., McCabe, D. E., and Swain, W. F. (1995) Methods for Particle-Mediated Gene Transfer into Skin, in *Methods in Molecular Medicine, Gene Therapy Protocols,* (Robbins, P., ed.) Humana, Totowa, NJ.
4. Sanford, J. C., Smith, F. D., and Russell, J. A. (1993) Optimizing the biolistic process for different biological applications. *Meth. in Enzymol.* **217,** 483–509.
5. Yang, N.-S. and Sun, W. H. (1995) Gene gun and other non-viral approaches for cancer gene therapy. *Nature Med.* **1,** 481–483.
6. Cheng, L., Ziegelhoffer, P. R., and Yang, N.-S. (1993) In vivo promoter activity and transgene expression in mammalian somatic tissues evaluated by using particle bombardment. *Proc. Natl. Acad. Sci. USA* **90,** 4455–4459.
7. Ariazi, E. A., Thompson, T. A., Burkholder, J. K., Yang, N. S., and Gould, M. (1995) Transcriptional regulatory and response mapping of the rat Ha-ras upstream sequence using primary mammary epithelial cells. *Carcinogenesis* **16,** 965–968.
8. Nony, P., Prudhomme, J. C., and Couble, P. (1995) Regulation of the P25 gene transcription in the silk gland of Bombyx. *Biol. Cell.* **84,** 43–52.
9. Koster, J. G., Eizema, K., Peterson-Maduro, L. J., Stegeman, B. I., and Destree, O. H. (1996) Analysis of Wnt/engrailed signaling in Xenopus embryos using biolistics. *Dev. Biol.* **173,** 348–352.
10. Pecorino, L. T., Lo, D. C., and Brockes, J. P. (1994) Isoform-specific induction of a retinoid-responsive antigen after biolistic transfection of chimaeric retinoic acid/thyroid hormone receptors. *Development* **120,** 325–333.
11. Thompson, T. A., Gould, M. N., Burkholder, J. K., and Yang, N.-S. (1993) Transient promoter activity in primary rat mammary epithelial cells evaluated using particle bombardment gene transfer. *In vitro Cell. Dev. Biol.* **29A,** 165–170.
12. Jiao, S., Cheng, L., Wolff, J. A., and Yang, N.-S. (1993) Particle bombardment-mediated gene transfer and expression in rat brain tissues. *Bio/Technol.* **11,** 497–502.

13. Guo, Z., Yang, N.-S., Jiao, S., Sun, J., Cheng, L., Wolff, J. A., and Duncan, I. D. (1996) Efficient and sustained transgene expression in mature rat oligodendrocytes in primary culture. *J. Neuroscience Res.* **43,** 32–41.
14. Burkholder, J. K., Decker, J., and Yang, N.-S. (1993) Rapid transgene expression in lymphocyte and macrophage primary cultures after particle bombardment mediated gene transfer. *J. Immunol. Methods* **165,** 149–156.
15. Nicolet, C. M., Burkholder, J. K., Gan, J., Culp, J., Kashmiri, S. V. S., Schlom, J., Yang, N.-S., and Sondel, P. M. (1995) Expression of a tumor reactive antibody-interleukin 2 fusion protein after in vivo particle mediated gene delivery. *Cancer Gene Therapy* **2,** 161–170.
16. Keller, E. T., Burkholder, J. K., Shi, F., Pugh, T. D., McCabe, D., Malter, J. S., MacEwen, E. G., Yang, N.-S., and Ershler, W. B. (1996) In vivo particle mediated cytokine gene transfer into canine oral mucosa and epidermis. *Cancer Gene Therapy* **3,** 186–191.
17. Fuller, J. T., Fuller, D. H., McCabe, D., Haynes, J. R., and Widera, G. (1996) Gene gun mediated DNA immunization with HBV antigens: immune responses in mice, monkeys, and pigs. *Vaccine* **96,** 87–90.
18. Cheng, L., Sun, J., Pretlow, T. G., Culp, J., and Yang, N.-S. (1996) CWR22 xenograft as an exvivo human tumor model for prostate cancer gene therapy. *J. Natl. Cancer Instit.* **88,** 607–611.
19. Mahvi, D. M., Burkholder, J. K., Turner, J., Culp, J., Malter, J. S., Sondel, P. M., and Yang, N.-S. (1996) Particle-mediated gene transfer of granlulocyte-macrophage colony stimulating factor cDNA to tumor cells: implications for a clinically relevant tumor vaccine. *Human Gene Therapy* **7,** 1535–1543.
20. Bio-Rad Life Sciences 1997 Catalog.

10

Optimizing Electroporation Conditions for the Transformation of Mammalian Cells

William C. Heiser

1. Introduction

Electroporation is a process in which a controlled electrical pulse is applied to cells, inducing a transient destabilization of the cell membrane. During this time, the cells are highly permeable to exogenous substances in the surrounding media. DNA, proteins, and small molecules are all taken up by cells during electroporation; introduction of DNA into cells is the most common application. Gene transfer by electroporation offers many advantages for analysis of gene expression. The technique is simple, rapid, and reproducible. It is especially suited to suspension cultures and certain cell types that are poorly transfected by other means. Because all cells are transfected instantaneously, and essentially simultaneously, it is particularly suited to quantitation of gene transfer. Finally, single-copy, stable transfectants can often be isolated ***(1)***. Whereas the basic mechanisms of electroporation are still largely unknown, optimizing the conditions for electroporation of any particular cell type is primarily empirical.

For additional discussions on the theory and parameters to consider in optimizing electroporation see the articles by Potter ***(2,3)***, Shigekawa and Dower ***(4)***, Forster and Neumann ***(5)***, and Spencer ***(6)***. For examples of optimizing electroporation conditions in specific cell lines, see the articles by Chu, et al. ***(7)***, Knutson and Yee ***(8)***, Andreason and Evans ***(9)***, Anderson, et al. ***(10)***, and Baum, et al. ***(11)***.

1.1. Electroporation Theory and Electroporation Parameters

The process of electroporation occurs when cells are placed between two electrodes and a high voltage pulse is used to induce a temporary breakdown of

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

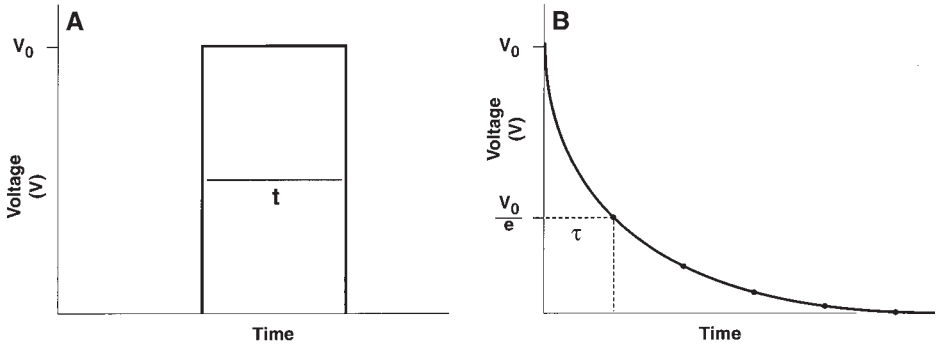


Fig. 1. **(A)** Square wave pulse. **(B)** Exponential decay pulse. In the case of square waves, the pulse length t describes the length of time the cell is subjected to the field strength of V_0. In the case of exponential decay waves, the time constant τ describes the length of time required to reduce the field strength to 1/*e* of the initial value.

the cell membrane. The most commonly used types of electrical pulse are the rectangular or square-wave pulse and the exponential decay or capacitance discharge pulse (**Fig. 1**). Square waves are generated by momentarily connecting a high voltage power supply with a switch. Exponential waveforms are produced when a charged capacitor is allowed to discharge into the sample containing the cells and DNA. In both of these pulses, the key instrument parameters that characterize the pulse are the field strength and the pulse length (see below). A third type of waveform using an electric field oscillating at a radio frequency has also been described ***(12,13)***. Because most commercial instruments utilize a capacitance discharge device, this discussion and the experimental protocols outlined below are limited to this type of instrument. With slight modifications, the procedures can also be used to optimize electroporation using a square-wave pulse generator.

Exponential decay circuits used for electroporation have the advantage of being simple systems with easily controlled parameters. The nature of the electric pulse is determined by the internal capacitors, resistors, and power supply, as well as by the circuitry of the instrument. The capacitor is a device used for storing electrical charge and, in part, determines the energy and the length of time that the electroporation pulse is delivered to the cells. The capacitance *C* (typically expressed in microfarads, μF) is a measure of the amount of electrical charge *Q* in a capacitor at a particular voltage *V*, according to the equation: $C = Q / V$. When a capacitor is discharged, the voltage delivered to the sample decays exponentially with time, t, as indicated by the following equation:

$$V = V_0 e^{-t / RC}$$

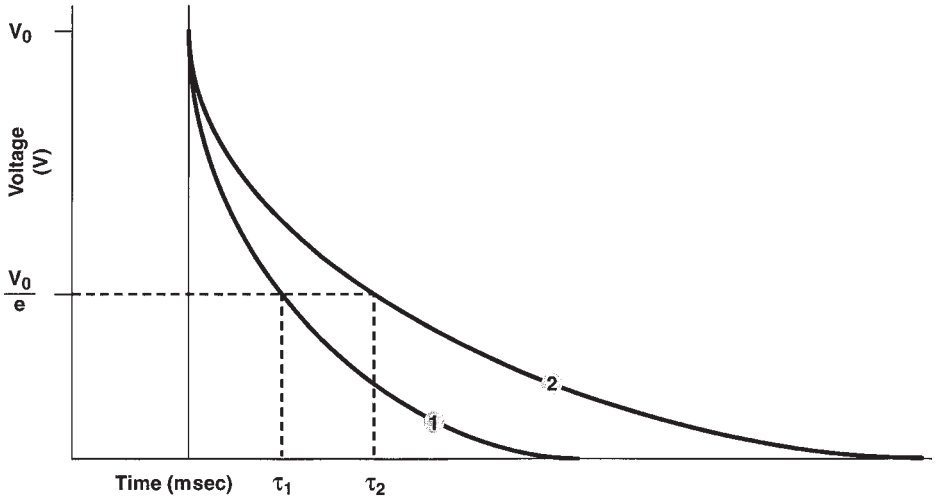


Fig. 2. Voltage decay from a capacitance discharge system according to the equation $V = V_0\, e^{-t/RC}$. In discharge 1 the voltage decreases from V_0 to V_0 / e at $t = \tau_1 = RC$. In discharge 2 the voltage decreases from V_0 to V_0 / e at $t = \tau_2 = RC$, such that $\tau_2 = 2\,\tau_1$. For discharge 2, the longer time constant can be produced by using a capacitor with twice the capacitance as in discharge 1, or by using electroporation media with twice the resistance as in discharge 1 (neglecting ionization effects in the media). Alternatively, by decreasing the sample volume of cuvet 1 by one-half, the resistance of the circuit will increase by approx twofold (see text for details).

V_0 is the voltage of the capacitor at time 0, R is the resistance of the circuit (expressed in ohms, Ω), and e is the base of the natural logarithm ($e = 2.718$). **Fig. 2** shows how the voltage decreases with time and how a change in capacitance or resistance can affect the voltage decay. The time required for the voltage to decay to V_0 / e (or 37% [$^1/_{2.718}$] of the initial value) occurs when $t = RC$ and is referred to as the "time constant," τ; it is usually expressed in milliseconds. By changing the capacitor setting on the instrument or the resistance of the circuit (e.g., by changing the ionic composition of the electroporation medium or by changing the volume of the buffer in the cuvet) it is possible to increase or decrease the time that the cells are exposed to the electric field (**Fig. 2**).

The parameter that describes the electrical environment in the electroporation chamber is the electric field strength E measured in V/cm. The electric field is created by the application of a potential difference (voltage) between two metal electrodes. For the electrode geometry of two parallel plates separated by a distance d (cm), E is given by the equation:

$$E = V / d$$

In practical terms, as the field strength increases, the electrical conductance between the cell membrane and the cell interior increases, causing the membrane to undergo a conformational change. When the potential across the biological membrane reaches 1 V, the membrane begins to breakdown ***(14,15)***. As the field strength increases further, the cell membrane may be irreversibly damaged by the electric field. Therefore, the optimum field strength must be sufficient to cause a local breakdown in the cell membrane without irreversibly damaging it. This is further complicated by the three-dimensional spherical shape of the cell, the fact that all of the cells in a population are not of uniform diameter, and the fact that a cell subjected to an electric field between two parallel electrodes will not experience a uniform voltage over the entire surface of the membrane.

While it is possible for an individual to build an electroporation instrument that will successfully electroporate cells, there are a number of significant concerns. First, and most important, the voltage and current applied during electroporation are more than sufficient to cause ventricular fibrillation or cardiac arrest in humans. Second, unless correct parts are used, it may be difficult to generate reproducible electric fields. Finally, without adequate knowledge of the resistance of the circuit, it may be difficult to determine the actual voltage applied to the system ***(5)***. Most commercial capacitance discharge electroporation instruments have built-in monitors that measure the time constant and the actual peak voltage that has been applied to the sample. However, even some of these instruments may not take into account the effect of internal resistors that have been installed in the instrument to protect the circuitry from arcing, which may occur at high field strengths (> 4000 V/cm). The Gene Pulser II Electroporator (Bio-Rad Laboratories, Hercules, CA) contains resistors within the circuitry to reduce the current flow during arcing and also to protect the instrument from damage should arcing occur. Additionally, an algorithm within the software of this instrument compensates for any voltage drop due to the circuitry (e.g., protection resistors and sample resistance) so that the applied voltage is equal to the set voltage.

1.2. Parameters Affecting Electroporation

Because many parameters should be considered when electroporating a cell line for the first time, published conditions that work for a similar cell type may be used as a starting point. **Table 1** lists electroporation conditions reported for several cell types. These variables need to be adjusted for each cell line and for the particular instrument to achieve optimal transfection efficiencies. These parameters are described briefly here.

Table 1
Conditions Used for Electroporating Plasmids into Several Cell Lines

Cell type	Species	Instrument	Capacitance (μF)	Voltage (V)	Cuvet (cm)	Volume (mL)	Buffer	Time const (ms)	Field strength (V/cm, calculated)	Temp (°C)	Ref
U937 (macrophage)	Human	B/R GP	960	200	0.4	0.4	RPMI 1640	28	500	RT	27
HuT 78 (T-cell)	Human	B/R GP	250	150	0.2	0.2	RPMI 1640	4	750	RT	27
GCT (macrophage)	Human	B/R GP	960	200	0.4	0.4	RPMI 1640	29	500	RT	27
293	Human	B/R GP	960	200	0.45	0.4	DMEM[a]	30	444	RT	28
HeLa	Human	B/R GP	960	200	0.45	0.4	DMEM[a]	30	444	RT	28
Foreskin fibroblasts	Human	B/R GP	960	230	0.4	0.8	He-BS[b]	13–15	575	RT	29
Lymphoblasts	Human	N R	960	270	0.4	0.6	PBS	N R	675	0	30
Lymphocytes	Human	B/R GP	960	250	0.4	0.25	RPMI 1640[c]	51	625	0	31
Jurkat	Human	B/R GP	960	250	0.45	0.4	Opti-MEM[d]	30	556	RT	28
Vero	Monkey	B/R GP	960	250	0.4	0.3	RPMI 1640	33–38	625	RT	32
CV-1	Monkey	Homemade	1080	300	0.38	1.0	HeBS[b]	N R	789	RT	7
COS-7	Monkey	B/R GP	960	300	0.4	0.3	RPMI 1640	33–38	750	RT	32
L929 (fibroblasts)	Mouse	B/R GP	500	350	0.4	0.3	RPMI 1640	33–38	875	RT	32
3T3 (fibroblasts)	Mouse	Homemade	1080	275	0.38	1.0	He-BS[b]	N R	724	RT	7
CHO	Hamster	B/R GP	960	300	0.4	0.8	PBS	13	750	RT	33
Rat-3 (fibroblasts)	Rat	B/R GP	250	330	0.4	0.4	Opti-MEM[e]	14–16	825	0	34
GH (pituitary)	Rat	N R	25	240–300	0.4	0.8	He-Suc[f]	N R	600–750	RT	35
Hepatocytes	Rat	N R	960	160	N R	0.8	PBS[g]	20–25		RT	36

[a]Dulbecco's MEM + 10% newborn calf serum.
[b]HEPES-buffered saline with 6 m*M* dextrose.
[c]RPMI 1640 + 20% fetal calf serum.
[d]Opti-MEM + 10% fetal bovine serum.
[e]Opti-MEM + 0.3 *M* sucrose + 10% calf serum.
[f]HEPES-buffered sucrose (8 m*M* HEPES, pH 7.4, 272 m*M* sucrose).
[g]PBS + 5% fetal calf serum.
N R = not reported.
B/R GP = Bio-Rad Gene Pulser.

1.2.1. Electroporation Buffer

The choice of buffer is critical for optimizing electroporation conditions because the electrical resistance of the medium is a major factor in determining the time constant (τ) of the pulse delivered. The resistance of high-ionic-strength (i.e., high-salt) buffer (e.g., phosphate-buffered saline or cell culture medium) is approximately 20 Ω, while the resistance of low-ionic-strength buffer (e.g., phosphate- or HEPES-buffered isotonic sucrose) is approximately 200 Ω. Therefore, if all other electroporation parameters are fixed, cells electroporated in a high-ionic-strength environment experience a shorter time constant than those electroporated in a low-ionic-strength environment. The most efficient transfection of mammalian cells by electroporation is generally achieved in high-ionic-strength buffer at high capacitance settings; however, the optimal medium for any specific cell line must be determined empirically. Cell culture medium without serum is recommended as the first choice for electroporation buffer because it contains the essential components for cell viability.

1.2.2. Temperature

Because electroporation involves the transient formation of pores in the cell membrane ***(15)***, and because application of an electric field causes cell heating, especially at high field strengths, the temperature at which the process occurs might be expected to have a role in the efficiency of transformation. In theory, keeping the cells at low temperature during electroporation should reduce cell heating, thereby increasing cell viability. Keeping the cells at low temperature following electroporation should also allow the pores to remain open longer, thereby allowing the DNA in the medium more time to enter the cell. Alternatively, raising the temperature after the pulse may help reseal the pores created during the pulse and increase cell survival. Additionally, changing the temperature of a solution changes the conductivity of the solution. With increasing temperature the conductivity of a solution increases, resulting in more current I passing through the solution at any given voltage. Because resistance R changes according to the equation $R = E / I$, as temperature (and hence, current) increases, the time constant τ decreases. Therefore, changing the temperature of the sample has multiple effects on cells. In practice, some cell types are more efficiently transformed at low temperature ***(16)*** while other types are more efficiently transformed at ambient temperature ***(7,9)***. This determination may be made simply by comparing gene expression following electroporation of cells either chilled on ice or maintained at room temperature ***(10)***.

1.2.3. Plasmid

There are numerous reports describing the effect of using linear or supercoiled plasmid for transformation. In transient expression assays, gene expression is generally higher in cells electroporated with supercoiled plasmid than in cells electroporated with linearized plasmid ***(2,17,18)***. However, for stable expression, linearizing the plasmid increases the number of permanently transfected cells, possibly because the linearized plasmid may be more easily integrated into the genomic DNA ***(7,16)***. To obtain the maximum transformation efficiency, plasmid should be of the highest purity, though the use of carrier DNA may reduce the need for ultrapure plasmid (*see* **Subheading 1.2.4.**). Cells electroporated with plasmid purified by CsCl ultracentrifugation generally express higher levels of the transfected gene than do cells electroporated with the same plasmid purified by other methods. However, such purification is often not feasible when assaying numerous plasmids. Generally, as long as the plasmids are prepared in the same manner, changes in expression levels are due to differences in transcription or translation of the gene of interest ***(17)***.

1.2.4. Additions to the Electroporation Buffer

Adding carrier DNA to the electroporation buffer has been shown to improve the reproducibility and sensitivity of transient assays ***(17)*** as well as to increase the efficiency of stable transformation ***(18)***. A large excess of carrier DNA in the electroporation buffer increases the transformation efficiency of low amounts of reporter plasmid, and also increases the range over which reporter gene activity is directly proportional to the concentration of reporter plasmid in the electroporation buffer. Different types of carriers are effective, including salmon sperm DNA ***(7)***, calf thymus DNA ***(18)***, and plasmid ***(17,18)***. Because both the concentration and purity of reporter plasmid will affect the level of transient expression, the addition of a large excess of DNA as carrier during electroporation may minimize this variability ***(19)***. By using carrier DNA, gene expression may be quantitated with miniprep DNA ***(17)***. Including a constant amount of a second reporter plasmid allows for correction for variations in electroporation efficiency in individual samples ***(17)***. Finally, electroporating cells in the presence of DEAE-dextran has been shown to result in an increase in transient expression up to 150-fold in several cell lines ***(20)***. This enhanced transformation efficiency is probably due to DEAE-dextran causing DNA to adhere to the cell surface, resulting in a higher DNA concentration in the local environment of the cells and thus allowing DNA to more readily enter cells upon electroporation.

1.2.5. Growth State

The highest gene expression following electroporation is obtained using cells which are actively growing and dividing rather than in stationary growth phase ***(10)***. For optimum cell recovery, the cell density in the cuvet should be in the range of 10^6–10^7 cells/mL; higher cell concentrations may result in undesirable cell fusion. Cells synchronized in the G2/M phase of the cell cycle have also been reported to have a higher transfection efficiency ***(21,22)***; however, this is not practical for routine electroporation.

1.2.6. Postelectroporation Treatment

Several reports have indicated that treatment of cells following electroporation with certain chemicals may increase expression several fold. Although the results of this effect have been limited to only a few cell lines, a stimulatory effect of sodium butyrate has been documented to result in a several fold increase in gene expression in a number of cases ***(21,23–25)***.

1.2.7. Manipulating the Time Constant and Field Strength

As discussed above, the time constant and field strength are the most important parameters that need to be optimized in order to obtain maximum gene transfer by electroporation. Although the time constant ($\tau = RC$) cannot be selected directly, it can be manipulated in several ways. First, the time constant can be varied most easily by changing the capacitance, which is set directly on a capacitance discharge electroporator. Second, by changing between a low-resistance buffer and a high-resistance buffer, the resistance of the electroporation solution, and hence, the time constant, may be altered. Finally, because the resistance of the circuit is also determined by the volume of liquid in the cuvet, changing the buffer volume also affects the time constant. Decreasing the volume of buffer in the cuvet increases the resistance, resulting in a nearly proportional increase in τ. Note that it is possible to decrease the volume in the electroporation cuvet yet maintain the same time constant by proportionally decreasing the capacitance (s*ee* **Table 2**).

As is the case with the time constant, the field strength ($E = V / d$) may also be manipulated in several ways. First, the voltage, which is directly proportional to the field strength, may be set directly on the instrument. This is the parameter over which the operator has most control. Hence, optimizing electroporation conditions is most easily accomplished by choosing a particular capacitance setting (for a specific time constant) and varying the voltage to identify the corresponding optimum field strength. This is the basis of the method described for optimizing electroporation in **Subheadings 3.2.1.** and **3.2.2.** Second, the electrode gap may be varied by using electroporation cuvets

Table 2
Compensating Capacitance and Voltage to Maintain Similar Values for the Time Constant and Field Strength as Cuvet Gap Width and Buffer Volume are Varied

	Case 1	Case 2	Case 3
Cuvet gap width (cm)	0.4	0.4	0.2
Buffer volume (mL)	0.8	0.4	0.2
Capacitance (μF)	1000	500	1000
Voltage (V)	300	300	150
Time constant, τ (msec)[a]	15	15	15
Field strength, E (V/cm)	750	750	750
Cell density (cells/mL)	0.5×10^7	1×10^7	2×10^7
Cells/cuvet[b]	0.4×10^7	0.4×10^7	0.4×10^7
Plasmid concentration (μg/mL)[b]	10	10	10
Plasmid/cuvet (μg)	8	4	2

[a]Time constants are for low resistance media at room temperature.
[b]In this example, the cell number and plasmid concentration are similar in all three cases.

with different gap widths. Electroporation cuvets are commercially available with gap widths of 0.1, 0.2, 0.4, and 0.9 cm. For electroporation of mammalian cells, the 0.2 and 0.4 cm gap width cuvets are most useful, although for electroporating large numbers of cells, 0.9 cm gap cuvets may be used. When the gap width is changed, it is possible to maintain the same field strength as long as the voltage is changed proportionally. **Table 2** illustrates how the capacitance and voltage may compensate to maintain the same values of τ and *E* when the electrode gap and the volume of media are altered.

For mammalian cells, there is generally a tradeoff between values of τ and *E* at which optimum gene transfer occurs. As the value of τ rises the applied voltage must be reduced to compensate. In general, as the field strength increases, viability decreases, but those cells that do survive are more likely to take up DNA.

1.3. Approaches to the Optimization of Electroporation

Electroporation conditions for a number of common cell lines and cell types are given in **Table 1**. These parameters are most applicable if the type of instrument being used is the same as the one for which the parameters are reported. If a different instrument is used, it is advisable to confirm the reported conditions in an initial experiment. Early reports using electroporation to transform mammalian cells employed high field strengths (2000–10,000 V/cm) and very short time constants (0.005–0.05 ms). However, later studies showed that

higher levels of expression could be obtained by using lower field strengths (500–2000 V/cm) and longer time constants (10–30 ms). The later conditions, in low-resistance media, are recommended starting points for optimizing electroporation.

If the electroporation conditions for the cell line of interest are not known, the following steps are suggested to arrive at the optimum electroporation conditions. A preliminary experiment should be performed to evaluate the capacitance and voltage settings. As mentioned above, these parameters, along with the choice of buffer, are the most crucial to determine the optimal values of τ and E when beginning work with a new cell line. The approximate settings for the capacitance and voltage may be determined by checking cell viability as a function of capacitance and voltage as described in Protocol 1.

Once the capacitance and voltage settings and the electroporation buffer (and, hence, the initial values for τ and E) that result in approximately 50% cell killing have been determined, these conditions are then used to assay gene expression as described in Protocol 2. Holding the capacitance (and hence, τ) constant, the voltage (and hence, E) is varied about 30% in both directions of the voltage setting that resulted in 50% cell killing. It is advisable to repeat this step a few times to verify reproducibility and to obtain the most reliable choice of voltage. Next, by holding E constant it is possible to further optimize the value of τ either by changing the capacitance setting (by about 30% in either direction) or by altering the resistance by changing the volume of electroporation buffer in the cuvet.

When the capacitance and voltage settings that result in maximum gene expression have been determined, other parameters may be assessed to further optimize expression. Those parameters that are most likely to have an effect are the temperature at which the cells are electroporated, the amount of carrier DNA in the electroporation buffer, the presence of DEAE-dextran in the electroporation buffer, and the postelectroporation conditions. These may be most easily assayed in a single experiment as outlined in Protocol 3. If an increase in expression is found in cells electroporated under two or more of the conditions tested in Protocol 3, a final experiment should be performed incorporating all of the procedures that resulted in increased gene expression.

2. Materials

1. Electroporation buffers:
   a. Phosphate-buffered saline (PBS): 137 m*M* NaCl, 2.7 m*M* KCl, 9.5 m*M* sodium phosphate pH 7.3; sterilize by autoclaving.
   b. HEPES-buffered saline: 10 m*M* HEPES, pH 7.3, 140 m*M* NaCl; sterilize by autoclaving.
   c. Opti-MEM I (Life Technologies, Gaithersburg, MD).

d. Cell culture medium (e.g., Life Technologies; Hyclone, Logan, UT): As required for each cell type.
e. Phosphate-buffered sucrose: 272 m*M* sucrose, 10 m*M* sodium phosphate, pH 7.3; filter sterilize.
f. Hepes-buffered sucrose: 272 m*M* sucrose, 10 m*M* HEPES, pH 7.3; filter sterilize.

2. Growth medium: Cell culture medium supplemented with serum and antibiotics, as required for each cell type.
3. Trypsin/EDTA (Life Technologies): 0.05% trypsin, 0.53 m*M* EDTA.
4. Trypan Blue (0.4%) in PBS.
5. TE: 10 m*M* Tris, pH 7.5, 1 m*M* EDTA, pH 8.0; sterilize by autoclaving.
6. DEAE-dextran (500,000 avg. mol wt, Pharmacia, Piscataway, NJ): Prepare a 50 mg/mL stock in water; sterilize by autoclaving.
7. Sodium butyrate (butyric acid, sodium salt, Sigma, St. Louis, MO): Prepare a 500 m*M* stock in water; filter sterilize.

3. Methods

3.1. Electroporation

The procedures in **Subheadings 3.1.1.** and **3.1.2.** describe the basic protocols for electroporation of attached and suspension cells, respectively. Protocols for optimizing electroporation are detailed in **Subheading 3.2.**

3.1.1. Attached Cells

1. One or two days prior to electroporation, subculture the cells so that they will be 50–75% confluent on the day of the experiment (*see* **Note 1**).
2. On the day of the experiment, aspirate the growth medium, then wash the cells twice with PBS. Add 0.5 ml of trypsin/EDTA per 75 cm^2 flask to detach the cells. Monitor the cells with an inverted microscope. When the cells have rounded up, tap the flask gently to loosen the cells from the plastic.
3. Add 5–10 mL of growth medium per 75 cm^2 flask. Resuspend the cells by pipeting, then transfer them to a 50-mL sterile, disposable polypropylene centrifuge tube.
4. Pellet the cells by centrifugation at 500*g* for 10 min. Aspirate the medium, then resuspend the cells in 25 mL of electroporation buffer.
5. Re-pellet the cells. Aspirate the buffer, then resuspend the cells in cold electroporation buffer at a concentration of (1–10) $\times$ 10^6 cells/mL (*see* **Note 2**). Keep the cells on ice until ready to electroporate.
6. Add the plasmid DNA(s) to the electroporation cuvets, then add the cells (*see* **Notes 3–5**).
7. Gently tap the sides of the cuvet several times to mix the cell/plasmid suspension. Place the cuvet in the electroporation shocking chamber and pulse the cells at the desired voltage and capacitance settings. Record the time constant for the pulse.
8. Remove the cuvet from the shocking chamber (*see* **Note 6**). Using a Pasteur pipet containing ~0.5 mL of growth medium, transfer the cells from the cuvet to a

35-mm tissue culture plate or a six-well tray. Add growth medium to bring the volume to 2 mL. Place the dishes into a humidified CO_2 incubator at 37°C.
9. Assay transient expression of the reporter gene 24–48 h following electroporation.

3.1.2. Suspension Cells

1. One or two days prior to electroporation, subculture the cells so that they will be in mid- to late-log phase, generally about $(5–15) \times 10^5$ cells/mL, on the day of the experiment. (About $(1–5) \times 10^6$ cells will be needed per electroporation.)
2. On the day of the experiment, harvest the cells by transferring the appropriate number of cells to a 50-mL sterile, disposable, polypropylene centrifuge tube. Pellet the cells by centrifugation at 500*g* for 10 min.
3. Aspirate the medium, then resuspend the cells in 25 mL of electroporation buffer.
4. Follow steps 5–9 in **Subheading 3.1.1.**

3.2. Optimizing Electroporation

Protocols 1, 2, and 3 offer suggestions on optimizing electroporation conditions for any cell line. If nothing is known about electroporating a cell line, use Protocol 1 to determine ranges of τ and E that may be used for electroporation. This procedure is a simple and rapid assay based on cell viability as an indicator of potential electroporation efficiencies. Use Protocol 2 to determine more precisely the settings of capacitance and voltage to use in assessing reporter gene expression. This is the most important series of experiments to perform to optimize electroporation. Finally, Protocol 3 offers suggestions on fine-tuning and further optimizing ancillary parameters to maximize gene expression.

3.2.1. Protocol 1: Preliminary Estimation of Electroporation Parameters.

Optimum electroporation conditions are generally associated with a 40–80% loss in viability ***(7,9)***. A first approximation of the parameters of capacitance and field strength required for electroporation may be assessed in a preliminary experiment by estimating cell viability one day following electroporation of cells under several conditions. Because most cells are optimally electroporated in low-resistance buffer (e.g., cell culture medium, PBS, or similar buffer), the initial electroporation experiment should be performed in one of these media unless it is known that the cells are easily killed using these conditions, in which case high-resistance buffer should be used. It is important to include a control sample of cells which were not subjected to an electric pulse.

1. Use Opti-MEM or cell culture medium (without serum), or PBS or HEPES-buffered saline without Mg^{2+} and Ca^{2+} (which in some cases may reduce transformation ***[26]***). However, if cells are sensitive to electroporation conditions required for low-resistance buffer, use HEPES- or phosphate-buffered sucrose.

2. For electroporation use 0.4-cm cuvets with 0.4 mL cells/cuvet at room temperature.
3. Electroporation settings for low-resistance media (with approximate values of τ at room temperature):
   Capacitance = 25 μF; voltage = 1200, 1300, 1400, 1500, 1600, 1700 V (τ ≈ 0.7 ms)
   Capacitance = 500 μF; voltage = 250, 300, 350, 400, 450, 500, 550 V (τ ≈ 15 ms)
   Capacitance = 1000 μF; voltage = 100, 150, 200, 250, 300, 350, 400 V (τ ≈ 30 ms)
   Electroporation settings for HEPES- or phosphate-buffered sucrose:
   Capacitance = 25 μF; voltage = 100, 150, 200, 250, 300, 350, 400 V (τ ≈ 9 ms)
4. The day following electroporation monitor cell viability. It is not important to be quantitative, but only to assess the voltage that results in approximately a 50% decrease in cell density relative to the control that was not electroporated. For attached cultures, examine the cells microscopically and estimate the cell density of the electroporated samples relative to the control. For suspension cells, it may be necessary to stain the cells with trypan blue to distinguish the viable from the nonviable cells.
5. Interpretation of results: As voltage increases, cell viability decreases. The purpose of this experiment is to determine the approximate voltage that results in 50% cell death *relative* to the control. The electroporation conditions that reduce viability by 50% are also the conditions at which gene expression is near maximal.
6. In the event that there is much more than 50% cell death even at the lowest voltage, repeat this experiment using lower voltage settings at 25 and 500 μF and, in addition, perform the electroporation in phosphate-buffered sucrose.

3.2.2. Protocol 2: Determination of Capacitance and Voltage Yielding Maximum Gene Expression

Having determined from Protocol 1 the capacitance and voltage settings that result in a 50% decrease in viability, or having obtained suggested field strength and time constants from **Table 1** or other literature references, it is now practical to verify the instrument conditions required for optimum gene expression by electroporation. Include as a control a sample of cells that are not electroporated.

1. Use cuvets with the same electrode gap, the same electroporation buffer, and the same volume of buffer used in Protocol 1, or indicated in **Table 1** or other reference.
2. Use 1–10 μg of reporter plasmid/cuvet in either TE or water at room temperature (*see* **Note 7**).
3. Electroporation settings: Use the capacitance and voltage settings identified in Protocol 1 that result in 50% cell killing, or the optimum settings given in **Table 1**, or other reference; vary the voltage approximately 30% higher and lower in 20–25 V increments. For example, if the voltage that results in 50% cell killing (from Protocol 1) or if the optimum voltage (from **Table 1**) is 250 V, electroporate seven samples, at 175, 200, 225, 250, 275, 300, and 325 V (*see* **Note 8**).
4. Monitor cell viability (as described in Protocol 1) and assay expression of the transfected gene the day following electroporation.

Table 3
Optimizing Ancillary Conditions in Electroporation[a].

Sample	Temperature[b] Pre	Post	Carrier[c]	DEAE-dextran[d]	Butyrate[e]
1	RT	RT	–	–	–
2	RT	0°C	–	–	–
3	0°C	RT	–	–	–
4	0°C	0°C	–	–	–
5	RT	RT	+	–	–
6	RT	RT	–	+	–
7	RT	RT	–	–	+

[a]See Procedure 3 for details.
[b]Temperature of cells prior to electroporation (pre) or following electroporation (post) (see text for details).
[c]Cells electroporated with 50 µg of pUC or pTZ carrier.
[d]Cells electroporated in buffer containing 5 or 10 µg/mL DEAE-dextran.
[e]Cells plated in media containing 5 or 10 m*M* sodium butyrate following electroporation.

5. Interpretation of results: At low voltages, gene expression increases with increasing voltage. As the voltage is further increased, cell killing continues to increase; as a result, gene expression plateaus, then drops. This experiment should determine the voltage required for maximum gene expression at several capacitance settings. It should also confirm the cell viability results of Protocol 1.

3.2.3. Protocol 3: Optimizing Ancillary Parameters to Improve Electroporation Results

Gene expression in electroporated cells may be further increased by changing the temperature at which cells are electroporated, by adding carrier DNA and DEAE-dextran, and by treating the cells after electroporation with sodium butyrate. In this case, the starting electroporation conditions should be those determined in Protocol 2. Include the modifications listed in **Table 3** and described below.

1. For room temperature electroporations, add cells and plasmid to an electroporation cuvet and allow the cells to equilibrate at room temperature for about 5 min before pulsing. For electroporation at 0°C, chill the electroporation cuvet on ice before adding plasmid and cells, then incubate an additional 5 min on ice before pulsing.
2. For postelectroporation treatment at room temperature, use growth medium at room temperature to transfer the cells to a six-well tray or 35-mm plate. For postelectroporation treatment at 0°C, immediately after pulsing the cells place the cuvet on ice for about 5 min before transferring the cells to a six-well tray or 35-mm plate.

3. Add 50 µg of carrier plasmid DNA (such as pUC or pTZ) and 5 or 10 µg/mL of DEAE-dextran (final concentration) to the cells prior to electroporation.
4. After electroporation, plate cells in medium containing 5 or 10 m*M* sodium butyrate for 16–24 h.
5. Interpretation of results: The effects of temperature, carrier DNA, DEAE-dextran, and sodium butyrate on gene expression in any cell type may only be determined empirically. These changes may result in either higher or lower gene expression compared to the control (**Table 3**, Sample 1). Should several of these conditions result in higher gene expression compared to the control, an additional experiment combining those conditions should be performed to further optimize the electroporation conditions.

4. Notes

1. Depending on the cell size and the final cell density, a 75 cm^2 flask contains about $(5–15) \times 10^6$ cells when the cells are about 50% confluent. About $(1–5) \times 10^6$ cells will be needed per electroporation cuvet.
2. The cell concentration is dictated by the number of cells per electroporation and the volume of buffer used in the electroporation cuvet. As a starting point, prepare cells at a concentration of 5 x 10^6 cells/mL and use either 0.4 ml of cells in a 0.4 cm cuvet or 0.2 mL of cells in a 0.2 cm cuvet.
3. Isolate plasmid using standard procedures ***(37,38)*** or a commercially available kit (e.g., Bio-Rad Laboratories, Hercules, CA, or Qiagen, Chatsworth, CA). If results of electroporation are being used to determine differences between plasmids (e.g., promoter strength), it is important to prepare the plasmids by the same method. The plasmid used for electroporation should be resuspended in sterile water or sterile TE at a concentration of at least 0.5 µg/µL. The amount of plasmid required per electroporation sample will depend on the type of cells being studied. Adding 1–10 µg of plasmid per sample is a good starting point; for any given electroporation conditions, gene expression is usually proportional to the concentration of plasmid in the sample.
4. If electroporation is performed at room temperature, add the cells to the cuvet and allow them to equilibrate for 5 min before pulsing. If the electroporation is performed at 0°C, place the cuvet on ice for a few minutes before adding the cells.
5. If the same plasmid is being used in all electroporations (such as when optimizing the procedure), mix the plasmid with the cells in a single tube; this reduces the number of pipeting steps and minimizes sample variation.
6. Chilling the cells to 0°C at this step may improve cell viability, resulting in higher transformation efficiency. In this case, incubate the cuvet containing the electroporated cells on ice for 5 min before transferring them to growth medium.
7. The optimum conditions for electroporation may be most easily determined using a plasmid containing a reporter gene that can be easily assayed. Reporter genes for which vectors are widely available include ß-galactosidase (ß-gal from *Escherichia coli*), luciferase (luc from the firefly *Photinus pyralis*), and chloram-

phenicol acetyltransferase (CAT from *E. coli*). The luc assay is by far the easiest and most rapid. It is extremely sensitive, and it can be quantitated over five orders of magnitude using a luminometer. Luc is the reporter gene of choice for analyzing the transcriptional activity of weak promoters. ß-gal can be assayed histologically using 5-chloro-4-bromo-3-indoyl-ß-D-galactoside (X-gal), biochemically using *o*-nitrophenyl-ß-D-galactoside (ONPG), or in a chemiluminescence assay. Staining cells with X-gal provides a simple, rapid, and qualitative method for assessing the voltage and capacitance in a preliminary experiment. Luc and ß-gal may be conveniently assayed in the same cell extract, permitting the use of one as a reporter gene and the other as an internal control. CAT is probably the most widely used reporter gene, however, it is more difficult to quantitate than either luc or ß-gal. The CAT assay is also more time consuming, and requires the use of radioactively labeled chloramphenicol, acetate, or acetyl-CoA. All of these reporter genes are highly expressed 24–48 h. following electroporation.

8. It is advisable to assess gene expression at all of the capacitance settings used in Protocol 1 that resulted in a 50% cell killing because cells electroporated at different capacitance settings experience different time constants, a major factor in optimizing electroporation conditions.

References

1. Boggs, S. S., Gregg, R. G., Borenstein, N., and Smithies, O. (1986) Efficient transformation and frequent single-site, single-copy insertion of DNA can be obtained in mouse erythroleukemia cells transformed by electroporation. *Exp. Hematol.* **14,** 988–994.
2. Potter, H. (1988) Electroporation in biology: methods, applications, and instrumentation. *Anal. Biochem.* **174,** 361–373.
3. Potter, H. (1993) Application of electroporation in recombinant DNA technology. *Methods Enzymol.* **217,** 461–478.
4. Shigekawa, K. and Dower, W. J. (1988) Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *BioTechniques* **6,** 742–751.
5. Forster, W. and Neumann, E. (1989) Gene transfer by electroporation, in *Electroporation and Electrofusion in Cell Biology* (Neumann, E., Sowers, A. E., and Jordan, C. A., eds.), Plenum, New York, pp. 299–315.
6. Spencer, S.C. (1991) Electroporation technique of DNA transfection, in *Methods in Molecular Biology*, Vol. 7 (Murray, E. J., ed.), Humana Press, Totowa, NJ, pp. 45–52
7. Chu, G., Hayakawa, H., and Berg, P. (1987) Electroporation for the efficient transfection of mammalian cells with DNA. *Nuc. Acids Res.* **15,** 1311–1326.
8. Knutson, J. C. and Yee, D. (1987) Electroporation: parameters affecting transfer of DNA into mammalian cells. *Anal. Biochem.* **164,** 44–52.
9. Andreason, G. L. and Evans, G. A. (1989) Optimization of electroporation for transfection of mammalian cell lines. *Anal. Biochem.* **180,** 269–275.

10. Anderson, M. L. M., Spandidos, D. A., and Coggins, J. R. (1991) Electroporation of lymphoid cells: factors affecting the efficiency of transfection. *J. Biochem. Biophys. Methods* **22,** 207–222.
11. Baum, C., Forster, P., Hegewisch-Becker, S., and Harbers, K. (1994) An optimized electroporation protocol applicable to a wide range of cell lines. *BioTechniques* **17,** 1058–1062.
12. Chang, D. C. (1989) Cell poration and fusion using an oscillating electric field. *Biophys. J.* **56,** 641–652.
13. Chang, D. C., Gao, P. Q., and Maxwell, B. L. (1991) High efficiency gene transfection by electroporation using a radio-frequency electric field. *Biochim. Biophys. Acta* **1992,** 153–160.
14. Baker, P. F. and Knight, D. E. (1983) High-voltage techniques for gaining access to the interior of cells: application to the study of exocytosis and membrane turnover. *Methods Enzymol.* **98,** 28–37.
15. Chang, D. C. (1992) Structure and dynamics of electric field-induced membrane pores as revealed by rapid-freezing electron microscopy,in *Guide to Electroporation and Electrofusion* (Chang, D. C., Chassy, B. M., Saunders, J. A., and Sowers, A. E., eds.), Academic, San Diego, CA, pp. 9–27.
16. Potter, H., Weir, L., and Leder, P. (1984) Enhancer-dependent expression of human κ immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc. Nat. Acad. Sci. USA* **81,** 7161–7165.
17. Showe, M. K., Williams, D. W., and Showe, L. C. (1992) Quantitation of transient gene expression after electroporation. *Nucl. Acids Res.* **20,** 3153–3157.
18. Nickoloff, J. A. and Reynolds, R. J. (1992) Electroporation-mediated gene transfer efficiency is reduced by linear plasmid carrier DNAs. *Anal. Biochem.* **205,** 237–243.
19. Showe, M. K. and Showe, L. C. (1995) Quantitation of transient gene expression, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J. A., ed.), Humana, Totowa, NJ, pp. 211–224.
20. Gauss, G. H. and Lieber, M. R. (1992) DEAE-dextran enhances electroporation of mammalian cells. *Nucl. Acids Res.* **20,** 6739–6740.
21. Goldstein, S., Fordis, C. M., and Howard, B. H. (1989) Enhanced transfection efficiency and improved cell survival after electroporation of G2/M synchronized cells and treatment with sodium butyrate. *Nucl. Acids Res.* **17,** 3959–3971.
22. Yorifuji, T., Tsuruta, S., and Mikawa, H (1989) The effect of cell synchronization on the efficiency of stable gene transfer by electroporation. *FEBS Lett.* **245,** 201–203.
23. Gorman, C. M., Howard, B. H., and Reeves, R. (1983) Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nucl. Acids Res.* **11,** 7631–7648.
24. Tatsuka, M., Orita, S., Yagi, T., and Kakunaga, T. (1988) An improved method of electroporation for introducing biologically active foreign genes into cultured mammalian cells. *Exp. Cell Res.* **178,** 154–162.
25. Satyabhama, S. and Epstein, A. L. (1988) Short-term efficient expression of transfected DNA in human hematopoietic cells by electroporation: definition of parameters and use of chemical stimulators. *DNA* **7,** 203–209.

26. Neumann, E., Schafer-Ridder, M., Wang, Y., and Hofschneider, P. H. (1982) Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* **1,** 841–845.
27. Barry, P., Pratt-Lowe, E., and Luciw, P. (1989) Electroporation of T-cell and macrophage cell lines. Bio-Rad Technical Bulletin 1349.
28. Maxwell, I. H. and Maxwell, F. (1988) Electroporation of mammalian cells with a firefly luciferase expression plasmid: kinetics of transient expression differ markedly among cell types. *DNA* **7,** 557–562.
29. Ray, F. A. (1995) Electroporation of plasmid DNA into normal human fibroblasts, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J. A., ed.), Humana, Totowa, NJ, pp. 133–140.
30. Xia, F. and Liber, H. L. (1995) Electroporation of human lymphoblastoid cells, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J.A., ed.), Humana Press, Totowa, NJ, pp. 151–160.
31. Cann, A. J., Koyanagi, Y., Koyonagi, Y., and Chen, I. S. Y. (1988) Efficient electroporation of primary human lymphocytes. Bio-Rad Technical Bulletin 1348.
32. Barry, P., personal communication.
33. Taghian, D. G. and Nickoloff, J. A. (1995) Electroporation of Chinese hamster ovary cells, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J. A., ed.), Humana, Totowa, NJ, pp. 115–121.
34. Dagher, S. F., Conrad, S. E., Werner, E. A., and Patterson, R. J. (1992) Phenotypic conversion of TK-deficient cells following electroporation of functional TK enzyme. *Exp. Cell Res.* **198,** 36–42.
35. Paulssen, R. H., Paulssen, E. J., and Gautvik, K. M. (1995) Electroporation of rat pituitary cells, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J. A., ed.), Humana, Totowa, NJ, pp. 123–131.
36. LeCam, A. (1995) Electroporation-mediated gene transfer into hepatocytes, in *Methods in Molecular Biology*, Vol. 48 (Nickoloff, J. A., ed.), Humana, Totowa, NJ, pp. 141–150.
37. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 1.21–1.52.
38. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1989) *Current Protocols in Molecular Biology*, Wiley, New York, pp. 1.6.1–1.7.15.

11

Calcium Phosphate Transfection of Mammalian Cultured Cells

Elaine T. Schenborn and Virginia Goiffon

1. Introduction

Calcium phosphate and DEAE-dextran reagents were incorporated into the first chemical methods that successfully transferred nucleic acid directly to cultured mammalian cells in a process referred to as transfection ***(1–4)***. Early transfection studies used viral RNA ***(1)*** and DNA ***(2,4)***, which, at that time, were relatively easy to propagate and purify, and allowed phenotypic discrimination of the transfected mammalian cells. Calcium phosphate coprecipitation and DEAE-dextran methods became widely used after cloning and manipulation of plasmid DNA became routine, and it was demonstrated that plasmid DNA was effectively transferred to cultured cells using these methods. Together with advancements in vector development came the introduction of additional transfection methods using chemical reagents such as liposomes ***(5,6)***, dendrimers ***(7,8)***, and cationic polymers ***(9,10)***, plus physical methods such as electroporation ***(11)*** and biolistic microparticle bombardment ***(12)***. However, the calcium phosphate coprecipitation technique still remains one of the most widely used in vitro transfection methods.

The calcium phosphate coprecipitation technique consists of mixing DNA with calcium chloride and a phosphate buffer, forming a fine precipitate, and distributing over cultured cells. In a classical publication of Graham and van der Eb ***(4)***, the effect of calcium concentration, DNA concentration, addition of carrier DNA, pH, and transfection interval were systematically studied using infectivity of transferred adenovirus DNA as the assay endpoint. Applications using the calcium phosphate coprecipitation technique were extended by Wigler et al. ***(13)*** using endonuclease-restricted herpes simplex virus (HSV) DNA to assay for long-term transformation by the HSV *tk* gene. The mecha-

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

nism for transfection of DNA with calcium phosphate may involve phagocytosis by the cells of the coprecipitates ***(14)***. The DNA also achieves some protection from extracellular and intracellular nucleases by its association with calcium phosphate ***(14)***.

Continued popularity of the calcium phosphate coprecipitation technique is based upon several features. The method works with a variety of cultured cell types and generally requires minimal optimization of the protocol. The procedure does not require specialized equipment. The standard components are calcium chloride and a HEPES-buffered saline. These reagents are simple to prepare from relatively inexpensive, readily available components, or can be purchased premade and qualified from biotechnology reagent companies. The method is versatile, being suitable for transient or stable transfection paradigms, and for adherent cells or cells growing in suspension culture ***(15)***. Disadvantages of the technique include 1) the precise conditions for precipitate formation are difficult to control and reproduce between experiments ***(16)***, even for experienced researchers; 2) the technique is limited to in vitro applications; and 3) the technique is not universally successful with all cell types.

2. Materials

All solutions should be prepared with deionized, nuclease-free water and tissue-culture-grade reagents. Solutions should be sterilized by filtration through a 0.2 µm filter.

1. Growth media: suitable for the cell type (*see* **Note 1**).
2. Transfection-quality DNA (*see* **Notes 2–4**): Containing the reporter, selectable marker, or other genetic element of interest.
3. 2X HEPES-Buffered Saline (2X HBS): 50 m*M* HEPES (e.g., Sigma [St. Louis, MO] cat. no. H9136), 280 m*M* NaCl, and 1.5 m*M* (anhydrous) Na_2HPO_4 . Adjust the final pH to 7.1 with 1 *N* NaOH. An exact pH is important for efficient transfection. Filter sterilize and store in 5 mL aliquots at –20°C. This solution may be thawed and refrozen several times. Commercial suppliers of pretested reagents are available (*see* **Note 5**).
4. 2 *M* $CaCl_2$: Filter sterilize and store in 5 mL aliquots at –20°C. This solution may be thawed and refrozen several times.
5. Transfection-quality water (deionized, nuclease-free, sterilized).
6. Phosphate-buffered saline (PBS; 1X): 137 m*M* NaCl, 2.7 m*M* KCl, 4.3 m*M* Na_2HPO_4, and 1.47 m*M* KH_2PO_4. Adjust the final pH to 7.1. Filter sterilize and store at room temperature. A 10X stock may be prepared for convenience and stored at room temperature.
7. BES-buffered saline (BBS): 50 m*M* BES (Calbiochem [La Jolla, CA] cat. no. 391334), 280 m*M* NaCl, and 1.5 m*M* Na_2HPO_4. Adjust to a final pH of 6.95 using 1 *N* NaOH. An exact pH is important for efficient transfection. Filter ster-

ilize and store in 5mL aliquots at –20°C. This solution can be thawed and refrozen several times.
8. 2.5 *M* $CaCl_2$(10X): Filter sterilize and store in aliquots at –20°C.
9. Staining solution: 2% methylene blue, 60% methanol.
10. Glycerol shock solution: 15% Glycerol in HBS.
11. Dimethyl sulfoxide (DMSO) shock solution: 1X PBS, 10% tissue culture grade DMSO (e.g., Sigma cat. no. D2650); the DMSO should be from a freshly opened ampule.

3. Methods

3.1. Plating Adherent Cells for Transfection (*see* Note 6)

The following protocols are designed for 60-mm tissue culture dishes. The number of cells, reagent volumes, and DNA amounts may be scaled up or down proportionately for different sized culture dishes. To simplify calculations, **Table 1** lists the growth areas for standard sizes of tissue culture dishes relative to the area of a 60-mm dish.

1. The day before the transfection experiment, trypsinize exponentially growing cells.
2. Neutralize the trypsin with serum-containing medium and pipet the cells gently to obtain a single-cell suspension.
3. Count the cells using a hemacytometer and dispense approx 3×10^5 cells per 60-mm culture dish.
4. Rock the plates gently from side to side to distribute the cells evenly over the growth surface of the dish. Avoid a swirling motion, which can concentrate the cells on the outer edge of the dish.
5. Incubate the cells overnight in growth medium in a humidified CO_2 incubator. The plating density may be adjusted depending on how rapidly the cells divide and how toxic the transfection treatment. In general, the cells should be 50–80% confluent the day of the transfection, and nearly confluent at the time of harvest or dilution into selective media.

3.2. Standard Calcium Phosphate Coprecipitation (*see* Notes 7–9)

An overview of workflow for the standard calcium phosphate transfection technique is shown in **Fig. 1**. In the following protocols, volumes and amounts are provided for 60-mm-size tissue culture dishes. The cell densities, reagent volumes, and DNA amounts may be scaled up or down proportionately for different sized dishes.

1. The day prior to transfection, plate the cells as described in **Subheading 3.1.**
2. One to 3 h prior to the transfection, remove the medium from the cells and replace it with 4.5 mL fresh growth medium. Bring all transfection reagents to room temperature.

Table 1
Relative Area for Standard Sized Tissue Culture Plates

Size of Plate	Growth Area (cm^2)[a]	Relative Area [b]
96 well	0.32	0.015
24 well	1.88	0.09
12 well	3.83	0.18
6 well	9.4	0.45
35 mm	8.0	0.38
60 mm	21	1
100 mm	55	2.6

[a] This information was calculated for Corning™ culture dishes.
[b] Relative area is expressed as a factor of the total growth area compared to a 60 mm plate. To determine the approximate cell number to plate, multiply 3×10^5 cells by this factor.

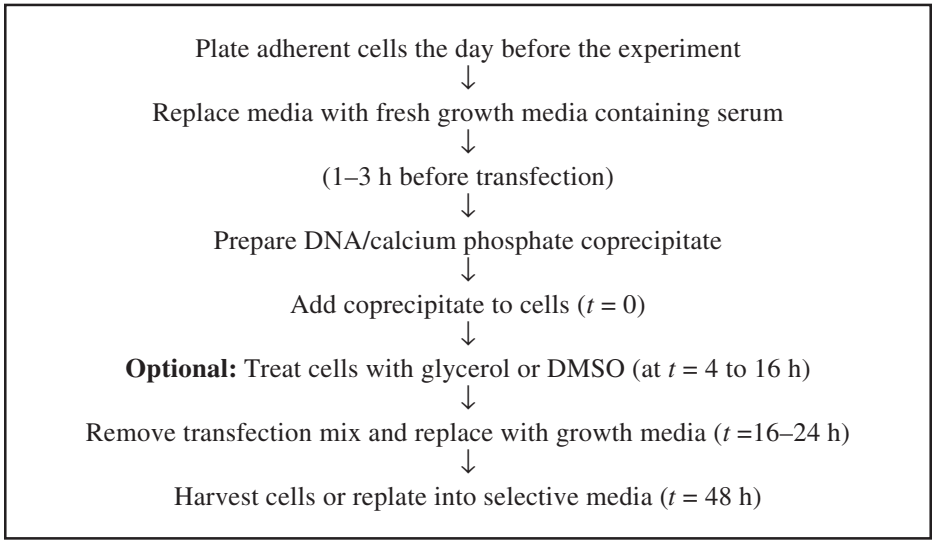


Fig. 1. General overview of the calcium phosphate protocol.

3. For each 60-mm dish to be transfected, aliquot nuclease-free water, 6–12 µg DNA, and 37 µL 2 *M* $CaCl_2$ in a final volume of 0.3 mL to one sterile tube. Mix well.
4. To a second sterile tube, aliquot 0.3 mL 2X HBS. It is helpful to use a clear tube for the 2X HBS in order to visualize the precipitate that forms after step 5.
5. In a tissue culture hood, gently vortex the tube containing the 2X HBS solution. Adjust the speed of the vortex such that the tube contents can be vortexed safely with the cap off.

6. While continuing to vortex, slowly add the prepared DNA solution dropwise to the 2X HBS (*see* **Note 10**). Recap the tube securely and place the tube where it will not be bumped or jostled.
7. Incubate for 30 min at room temperature. Note that many tissue culture hoods tend to be warmer than room temperature. After the incubation, a fine calcium phosphate-DNA coprecipitate should be visible.
8. Vortex the transfection solution and add it dropwise to the cells (0.6 mL per 60-mm dish). Gently rock the dish of cells back and forth to evenly distribute the transfection mix, and return to a 37°C CO_2 incubator. The calcium phosphate/DNA coprecipitate may be microscopically visible on the cells as black specks.
9. Replace tissue culture medium after 16–24 h. If the transfection appears to be particularly toxic to the cells, the transfection mixture may be removed after 4 h or more.
10. For transient applications, harvest cells approx 48 h after the start of transfection (*see* **Note 11**).

3.3. Calcium Phosphate Coprecipitation (BES Modification) (*see* Note 12)

BES-buffered saline replaces the HBS from the standard protocol. This method is recommended for stable transfections using supercoiled DNA, and is suitable particularly for fibroblasts and epithelial cells ***(17,18)***. It also works for transient expression, although only at the same relative efficiency as the standard protocol using HEPES-buffered saline. This method is not enhanced by further treatment with glycerol or DMSO ***(19)***.

1. The day prior to transfection, plate the cells as described in **Subheading 3.1.**
2. Cover cells with 4 mL medium per 60-mm plate. If drug sensitivity is being used as a selective marker for stable transfection, prepare an extra plate of nontransfected cells as a test for sensitivity to the drug of interest.
3. Dilute an aliquot of the 2.5 *M* $CaCl_2$ stock to a final concentration of 0.25 *M*. In a sterile tube, add 4–12 µg of plasmid DNA to 0.2 mL 0.25 *M* $CaCl_2$. Mix well.
4. In a tissue culture hood, add 0.2 mL of 2X BBS to the DNA solution. Mix well.
5. Incubate 10–20 min at room temperature.
6. Add the DNA/$CaCl_2$/BBS mixture dropwise to cells, and rock the plate gently to ensure even distribution.
7. Incubate for 15–24 h at 37°C in an incubator under 2–4% CO_2 (*see* **Note 13**).
8. Remove the medium, wash cells twice with growth medium or PBS, and add fresh growth medium. Incubate for an additional 24 h at 37°C under the CO_2 concentration usually used for the cells.
9. For transient applications, harvest cells approximately 48 h after the start of transfection (*see* **Note 11**).

3.4. Stable Transfection

For stable transfections, the cells are often transfected with a plasmid DNA containing a selectable marker, such as the neomycin phosphotransferase gene. Either the standard calcium phosphate coprecipitation protocol or the BES modification may be used. It is helpful to use a dish of nontransfected cells as a control for drug sensitivity.

3.4.1. Plating of Cells

1. Trypsinize the cells and replate at several dilutions (typically between 1:20 and 1:200) in media containing the selective agent (e.g., the antibiotic G-418 when selecting for cells expressing the neomycin phosphotransferase gene) 48 h after the start of the transfection. If clones will be stained to calculate the transfection efficiency, dedicate extra plates for this purpose.
2. Replace the selective medium every 2–4 d, for about 14 d, or until all control cells (e.g., nondrug-resistant, if applicable) have died and distinct clones, or "islands," of surviving cells are clearly visible under the microscope.
3. Individual colonies may be selected using standard techniques and materials, such as cloning rings. Guidelines are provided in ***(20)***.

3.4.2. Determining the Transformation Efficiency

For a permanent record of stable transformation efficiency, the percentage of transfectants may be calculated by staining and counting the colonies in the following manner.

1. After clones are visible under the microscope, remove the media from the cells.
2. Add sufficient methylene blue staining solution to cover the cells, and incubate for 5 min.
3. Remove the stain solution, which may be reused. Rinse the cells thoroughly by dipping the culture dish into a reservoir of fresh tap water. Discard the wash.
4. Repeat this rinsing procedure several times, until the rinse water is nearly clear. Shake off excess moisture and air-dry the plates on paper towels. After drying, the plates may be stored at room temperature.
5. Count the number of visibly stained colonies for each dilution (*see* **Note 14**).
6. Calculate the percentage of cells transfected based on the number of stained colonies, the known dilution of cells, and the original number of cells plated.

3.5. Auxiliary Procedures

The utility of performing a glycerol ***(21–23)*** or DMSO ***(24,25)*** shock step should be determined for each cell line because not all cells respond favorably to these toxic treatments (*see* **Note 15**). In our lab, we found a glycerol shock step to be advantageous for Chinese hamster ovary (CHO) cells, but typically omit the glycerol or DMSO shock for HeLa, NIH3T3, COS-7, and 293 cells.

3.5.1. Glycerol Shock

The glycerol shock step may be performed any time between 4 and 16 h after the start of the transfection (*see* **Note 16**). The maximum time cells should be exposed to the glycerol solution is 2 min. Volumes are provided for a 60-mm tissue culture dish.

1. Warm the freshly prepared glycerol shock solution, growth medium and PBS or other wash solution to 37°C.
2. Remove the transfection solution and wash the cells once with 5 mL PBS. Remove the wash solution.
3. Add 2 mL of the glycerol shock solution. Incubate for up to 2 min at room temperature.
4. Remove the glycerol shock solution and wash the cells twice with 5 mL of PBS.
5. Remove the final wash and add regular growth medium. Return the cells to a 37°C incubator.

3.5.2. DMSO Shock

DMSO, like glycerol, is toxic to cells and the concentration and exposure times require careful optimization for each cell type. Volumes provided below are for a 60-mm tissue culture plate.

1. Prepare a fresh DMSO shock solution (2 mL per 60-mm plate) and warm the solution to 37°C.
2. Remove the medium from the cells.
3. Add 2 mL of the DMSO shock solution to the cells and incubate for up to 2.5 min at room temperature.
4. Remove the DMSO shock solution and add 5 mL of growth medium. Return the cells to a 37°C incubator.

4. Notes

1. Transfection efficiency may be affected by different batches of serum used in growth medium. It is advisable to compare the prospective new lot of serum with the old one in a transfection experiment.
2. The quality of the DNA is important. The method chosen for DNA purification should produce high quality DNA that works well in transfections. The DNA should be free from contaminants (e.g., an A_{260}/A_{280} ratio of 1.8–1.9; it should look "clean" by agarose gel analysis). The DNA should be sterilized by ethanol precipitation; this is especially critical for long-term, stable experiments. Several satisfactory methods exist for producing transfection-grade plasmid DNA. The plasmid DNA may be purified by CsCl gradient centrifugation ***(19,26)***. Care must be taken to remove the cesium and ethidium bromide from the final preparation. Some anion exchange chromatography procedures (e.g., Qiagen, Chatsworth, CA) work well for producing high-quality plasmid DNA. Some commercial purification systems may leave contaminants in the DNA preparation that interfere with transfection.

3. The concentration of DNA used for transfection should be approx 1 mg/mL in TE buffer (10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA) or nuclease-free water. Addition of a large volume of DNA in Tris buffer could change the pH of the mixture, which is a critical parameter for calcium phosphate methods.
4. In cotransfection experiments where two or more types of DNA are being transfected, it is important to keep the total amount of DNA constant for each treatment that is to be compared. DNA amounts outside an optimal range will affect calcium phosphate precipitation formation and reduce transfection efficiency. In situations where a smaller total amount of experimental DNA is being transfected, a neutral plasmid DNA such as the pGem®-3zf+ vector (Promega Corp.) may be added to increase the total amount of DNA in the transfection.
5. Commercial suppliers offer kits containing premade, pretested reagents for use in transfecting cells with the calcium phosphate method (e.g., Promega Corp. ProFection® Mammalian Transfection System-Calcium Phosphate System, cat. no. E1200).
6. Rigorous practice of sterile techniques is essential for optimal transfection results. General procedures are discussed in ***(20,27,28)***. Mycoplasma or other contaminating organisms can adversely affect the results of a transfection experiment. Obtain cells from a reliable source, and discard cells that have been passaged excessively (i.e., more than 10 times since the freezer stock). Culture cells to provide optimal growth patterns prior to transfection. For most cells, this entails diluting or subculturing the cells every 3–4 d to prevent them from becoming confluent or overly dense.
7. Parameters critical for successful transfections have been correlated with precipitate size. Smaller, less aggregated precipitates have been associated with higher transfection efficacy ***(16)***. Factors that affect the precipitate size include calcium and DNA concentrations, buffer, pH, and temperature. Several modifications to the standard calcium phosphate coprecipitation protocol are published, such as increasing the calcium concentration and decreasing the precipitation interval ***(16)***.
8. Modifications of calcium phosphate transfection techniques for suspension cells are described in ***(26)***.
9. Treating cells with 100 µ*M* chloroquine during the initial 3–5 h of the calcium phosphate coprecipitation method may increase the uptake of the DNA. A 100 m*M* stock solution of chloroquine diphosphate (e.g., Sigma cat. no. C6628; 60 mg/mL in water; filter sterilized and protected from light) may be diluted 1:100 directly into the medium either before or after the addition of the calcium phosphate–DNA coprecipitate to the cells. The cells must be washed after the chloroquine treatment prior to adding complete growth medium to avoid excessive toxicity.
10. The method of addition of the DNA to the buffered saline (for example, fast vs dropwise, with or without concomitant bubbling) together with precipitation time affect the transfection efficiency ***(22,29)***.
11. If a reporter gene was used for transient transfections, prepare cell extracts if necessary, and assay for reporter activity. Otherwise, assay for the marker or protein of interest.

12. The BES-buffered saline solution, at pH 6.95, allows finer calcium phosphate/DNA precipitates to form gradually during the overnight incubation period with the cells when the CO_2 level is maintained near 3% in the incubator ***(17,18)***.
13. A Fyrite® gas analyzer (Bacharach, Pittsburgh, PA) may be used to monitor the CO_2 concentration of the incubator.
14. It may be difficult generating stable transfection efficiencies for certain cell types (e.g., CHO cells) that detach and resettle elsewhere on the plate during the selection period. In this case, the observed number of colonies will not be consistent from one cell dilution to another.
15. Glycerol and DMSO may enhance uptake of DNA by altering the cell membrane or by osmotic effects ***(23)***. Glycerol and DMSO solutions (usually 10–20%) are added to the cells for a precisely timed interval of 30 s to a few minutes, and then are removed from the cells.
16. In general, the calcium phosphate solution should be left on the cells for as long as possible, and the glycerol shock step performed 16 h after the start of the transfection. However, not all cells will tolerate a longer calcium phosphate exposure, and for these cells it would be best to perform the glycerol shock step earlier, for example, 4–6 h after the initial DNA exposure. A recent report indicated that lower concentrations of glycerol (4–6%) present on the cells with the calcium phosphate/DNA precipitate for 3–4 h, was an effective treatment in the cell lines tested ***(22)***. The optimum conditions for the glycerol shock should be determined empirically for each cell line.

References

1. Vaheri, A. and Pagano, J.S. (1965) Infectious poliovirus RNA: a sensitive method of assay. *Virology* **27**, 434–436.
2. McCutchan, J. H. and Pagano, J. S. (1968) Enhancement of the infectivity of Simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* **41**, 351–357.
3. Szybalska, E. H. and Szybalski, W. (1962) Genetics of human cell lines IV. DNA-mediated heritable transformation of a biochemical trait. *Proc. Natl. Acad. Sci. USA* **48**, 2026–2034.
4. Graham, F. L. and van der Eb, A. J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–467.
5. Fraley, R., Subramani, S., Berg, P., and Papahadjopoulos, D. (1980) Introduction of liposome-encapsulated SV40 DNA into cells. *J. Biol. Chem.* **255**, 10,431–10,435.
6. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M., and Danielsen, M. (1987) Lipofection: A highly efficient, lipid mediated DNA transfection procedure. *Proc. Natl. Acad. Sci. USA* **84,** 7413–7417.
7. Haensler, J. and Szoka, F. C. (1993) Polyamidoamine cascade polymers mediate efficient transfection of cells in culture. *Bioconj. Chem.* **4**, 372–379.
8. Kukowska-Latallo, J. F., Bielinska,A. U., Johnson, J., Spindler, R., Tomalia, D. A., and Baker, J. R. (1996) Efficient transfer of genetic material into mammalian

cells using Starburst polyamidoamine dendrimers. *Proc. Natl. Acad. Sci. USA* **93**, 4897–4902.
9. Farber, R. E., Melnick, J. L., and Butel, J. S. (1975) Optimal conditions for uptake of exogenous DNA by chinese hamster lung cells deficient in hypoxanthine-guanosine phosphoribosyl transferase. *Biochim. Biophys. Acta* **390**, 298–311.
10. Boussif, O., Lezoualch, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., and Behr, J. P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. USA* **92**, 7297–7301.
11. Wong, T. K. and Neumann, E. (1982) Electric field mediated gene transfer. *Biochem. Biophys. Res. Commun.* **107**, 584–587.
12. Ye, G. N., Danielle, H., and Sanford, J. C. (1990) Optimization of delivery of foreign DNA into higher-plant chloroplasts. *Plant Molec. Biol.* **15**, 809– 819.
13. Wigler, M., Silverstein, S., Lee, L. S., Pellicer, A., Cheng, Y., and Axel R. (1977) Transfer of purified Herpes Virus thymidine kinase gene to cultured mouse cells. *Cell* **11**, 223–232.
14. Loyter, A., Scangos, G. A., and Ruddle, F. H. (1982) Mechanisms of DNA uptake by mammalian cells: fate of exogenously added DNA monitored by the use of fluorescent dyes. *Proc. Natl. Acad. Sci. USA* **79**, 422–426.
15. Chu, G. and Sharp, P. A. (1981) SV40 DNA transfection of cells in suspension: analysis of efficiency of transcription and translation of T antigen. *Gene* **13**, 197–202.
16. Jordan, M., Schallhorn, A., and Wurm, F. M. (1996) Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucl. Acids Res.* **24,** 596–601.
17. Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752.
18. Chen, C. A. and Okayama, H. (1988) Calcium phosphate-mediated gene transfer: a highly efficient transfection system for stably transforming cells with plasmid DNA. *Biotechniques* **6**, 632–638.
19. Kingston, R. E. (1995) Introduction of DNA into mammalian cells, in *Current Protocols in Molecular Biology* (Ausubel, F., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G. Smith, J. A., and Struhl, K., eds.), Wiley, New York.
20. Jakoby, W. B. and Pastan, I. H., eds. (1979) *Methods in Enzymology* **58**, Academic, New York, NY.
21. Frost, E. and Williams, J. (1978) Mapping temperature-sensitive and host-range mutations of adenovirus type 5 by marker rescue. *Virology* **91**, 39–50.
22. Wilson, S. P., Liu, F. L., Wilson, R. E., and Housley, P. R. (1995) Optimization of calcium phosphate transfection for bovine chromaffin cells: relationship to calcium phosphate precipitate formation. *Anal. Biochem.* **226**, 212–220.
23. Zauner, W., Kichler, A., Schmidt, W., Sinski, A., and Wagner, E. (1996) Glycerol enhancement of ligand-polylysine / DNA transfection. *Biotechniques* **20**, 905–913.
24. Lowy, D. R., Rands, E., and Scolnick, E. M. (1978) Helper-independent transformation by unintegrated Harvey sarcoma virus DNA. *J. Virology* **26**, 291–298.

25. Lewis, W. H., Srinivasan, P. R., Stokoe, N., and Siminovitch, L. (1980) Parameters governing the transfer of the genes for thymidine kinase and dihydrofolate reductase into mouse cells using metaphase chromosomes or DNA. *Somat. Cell Genet.* **6**, 333–347.
26. Sambrook, J., Fritsch, E. F., and Maniatis, T., ed (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
27. Freshney, R. I., ed. (1986) *Animal Cell Culture*. IRL Press, Oxford, England, pp. 1–11.
28. Freshney, R. I., ed. (1987) *Culture of Animal Cells*. A.R. Liss, Inc., New York, pp. 51–59.
29. Seelos, C. (1997) A critical parameter determining the aging of DNA-calcium phosphate precipitates. *Anal. Biochem.* **245**, 109–111.

12

DEAE-Dextran Transfection of Mammalian Cultured Cells

Elaine T. Schenborn and Virginia Goiffon

1. Introduction

DEAE-dextran transfection is one of the oldest chemical, nonviral methods developed to transfer RNA or DNA to cultured mammalian cells *(1,2)*. Early transfection studies used viral RNA *(1)* and DNA *(2)*, which were easy to propagate and purify, and allowed phenotypic discrimination of the transfected mammalian cells. The DEAE-dextran method is generally used only for transient expression studies, because long-term stable transfections are less successful using this reagent *(3)*. The standard transfection protocol involves exposing the cells to a DEAE-dextran and DNA solution. An alternative procedure is to expose the cells first to DEAE-dextran, wash the cells, and then add DNA *(4)*.

In the standard protocol, positively charged DEAE-dextran polymer binds with the negatively charged nucleic acids. The complex adsorbs to negatively charged cellular membranes and is taken into the cell, presumably by endocytosis *(5,6)*. The mechanism of action is less clear for the DEAE-dextran pretreatment protocol. Once the nucleic acids are internalized in the cell, they can become trapped within lysosomes. The addition of chloroquine may enhance transfection in certain cells by reducing degradation of the DNA in the lysosome *(6)*. Because chloroquine is cytotoxic, the concentration and exposure time must be tested and closely monitored. Transfection efficacy may also be increased by adding glycerol *(7)* or dimethyl sulfoxide (DMSO) *(7)* to the cells following exposure to DEAE-dextran and DNA. These compounds may enhance uptake of DNA by altering the cell membrane or by osmotic effects *(8)*.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

2. Materials

All solutions should be prepared with deionized, nuclease-free water and tissue-culture-grade reagents. Solutions should be sterilized by filtration through a 0.2 µm filter.

1. Growth media: suitable for the cell type (*see* **Note 1**).
2. Transfection-quality DNA (*see* **Notes 2–5**): containing the reporter, selectable marker, or other genetic element of interest.
3. Transfection-quality water (deionized, nuclease-free, sterilized).
4. Phosphate-buffered saline (PBS; 1X): 137 m*M* NaCl, 2.7 m*M* KCl, 4.3 m*M* Na_2HPO_4, and 1.47 m*M* KH_2PO_4. Adjust the final pH to 7.1. Filter-sterilize and store at room temperature. A 10X stock may be prepared for convenience and stored at room temperature.
5. 10 mg/mL DEAE-dextran in water: The molecular weight of the DEAE-dextran should be approximately 500,000 Dalton (e.g., Amersham Pharmacia Biotech, Piscataway, NJ, cat. no. 17-0350-01). Sterilize using a 0.2 µm filter. Aliquot and store at 4°C. Commercial suppliers of pre-tested transfection reagents are available (*see* **Note 6**).
6. 8 m*M* chloroquine diphosphate in PBS (e.g., Sigma, St. Louis, MO, cat. no. C6628). Filter sterilize and aliquot into amber vials, or cover with foil to protect from light. Store at 4°C.
7. Dimethyl sulfoxide (DMSO) Shock Solution: 1X PBS, 10% tissue culture grade DMSO (e.g., Sigma, cat. no. D2650); the DMSO should be from a freshly opened ampule.

3. Methods

3.1. Plating Adherent Cells for Transfection (*see* **Note 7**)

The following protocols are designed for 60-mm tissue culture dishes. The number of cells, reagent volumes, and DNA amounts may be scaled up or down proportionately for different sized culture dishes. To simplify calculations, **Table 1** lists the growth areas for standard sizes of tissue culture dishes relative to the area of a 60-mm dish.

1. The day before the transfection experiment, trypsinize exponentially growing cells.
2. Neutralize the trypsin with serum-containing medium and pipet the cells gently to obtain a single-cell suspension.
3. Count the cells using a hemacytometer and dispense approximately 3 x 10^5 cells per 60-mm culture dish.
4. Rock the plates gently from side to side to distribute the cells evenly over the growth surface of the dish. Avoid a swirling motion, which can concentrate the cells on the outer edge of the dish.
5. Incubate the cells overnight in growth medium in a humidified CO_2 incubator. The plating density may be adjusted depending on how rapidly the cells divide

Table 1
Relative Area for Standard Sized Tissue Culture Plates

Size of Plate	Growth Area (cm^2)[a]	Relative Area [b]
96 well	0.32	0.015
24 well	1.88	0.09
12 well	3.83	0.18
6 well	9.4	0.45
35 mm	8.0	0.38
60 mm	21	1
100 mm	55	2.6

[a] This information was calculated for Corning™ culture dishes.
[b] Relative area is expressed as a factor of the total growth area compared to a 60-mm plate. To determine the approximate cell number to plate, multiply 3×10^5 cells by this factor.

and how toxic the transfection treatment. In general, the cells should be 50–80% confluent the day of the transfection, and nearly confluent at the time of harvest or dilution into selective media.

3.2. Standard DEAE-Dextran

The DEAE-dextran protocol is useful for transient, but not stable, transfections *(3)*. In the standard protocol outlined in **Fig. 1**, the cells are exposed simultaneously to DNA and DEAE-dextran. Volumes and amounts are provided for 60-mm size tissue-culture dishes. The cell densities, reagent volumes, and DNA amounts may be scaled up or down proportionately for different sized dishes.

1. The day prior to transfection, plate the cells as described in **Subheading 3.1.**
2. Warm the PBS and DEAE-dextran solutions to 37°C. About 11 mL of PBS will be needed for each 60-mm plate.
3. In a sterile tube, dilute 2–6 µg DNA to a final volume of 326 µL in 1X PBS. Add 17 µL of 10 mg/mL DEAE-dextran and tap the tube gently several times to mix. The DEAE-dextran concentration is approximately 0.5 mg/mL in this solution.
4. Remove the growth medium and wash the cells twice with 5 mL PBS.
5. Disperse the DNA/DEAE-dextran mixture evenly over the cells. Return the cells to a humidified 37°C incubator for 30 min. Rock the plates once or twice during the incubation to keep the cells from drying out.
6. Gently overlay 3.5 mL of growth medium onto the cells. *Optional:* The growth medium may contain 35 µL of 8 m*M* chloroquine (*see* **Note 8**).
7. Return the cells to a humidified 37°C incubator for up to 2.5 h or until cytotoxicity is apparent.
8. Carefully replace the medium, *or* proceed with a DMSO shock (*see Method 3.4.*).
9. Harvest the cells approximately 48 h after the start of the transfection (*see* **Note 9**).

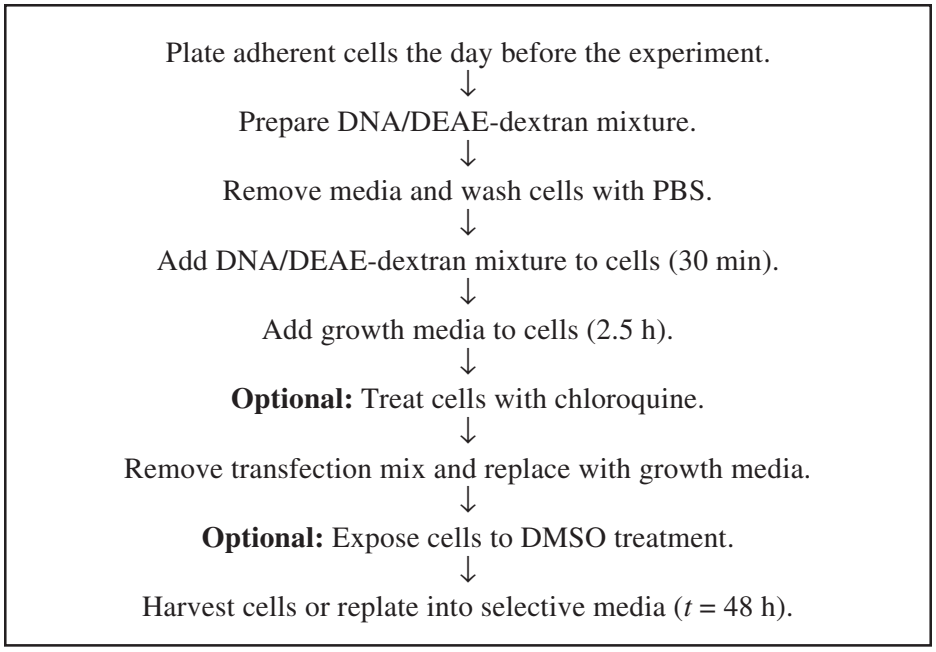


Fig. 1. The Standard DEAE-Dextran Protocol.

3.3. DEAE-Dextran Pretreatment

In this alternative protocol, cells are pretreated with DEAE-dextran prior to a longer DNA exposure (**Fig. 2**). The optimal standard or pretreatment DEAE-dextran protocol for any particular cell line needs to be empirically determined.

1. The day prior to transfection, plate the cells as described in **Subheading 3.1.**
2. Warm the PBS to 37°C. About 20 mL will be needed for each 60-mm plate.
3. Dilute the DEAE-dextran stock solution to 1 mg/mL in PBS. Prepare 2 mL of diluted DEAE-dextran for each 60-mm plate.
4. Dilute 2–6 µg DNA in PBS to a final volume of 325 µL.
5. Remove the growth medium from the cells. Add 5 mL of PBS. Incubate for 15 min at room temperature.
6. Aspirate the wash solution from the cells and add 2 mL of the diluted DEAE-dextran solution. Incubate at room temperature for 9 min.
7. Aspirate the DEAE-dextran solution. Gently wash the cells twice with 5 mL warmed PBS. Take care not to dislodge the cells, which may begin to detach after treatment with DEAE-dextran.
8. Aspirate the final wash. Pipet the diluted DNA evenly over the cells. Incubate for 30 min in a humidified 37°C incubator. Rock the plates once or twice during the incubation to keep the cells from drying out.

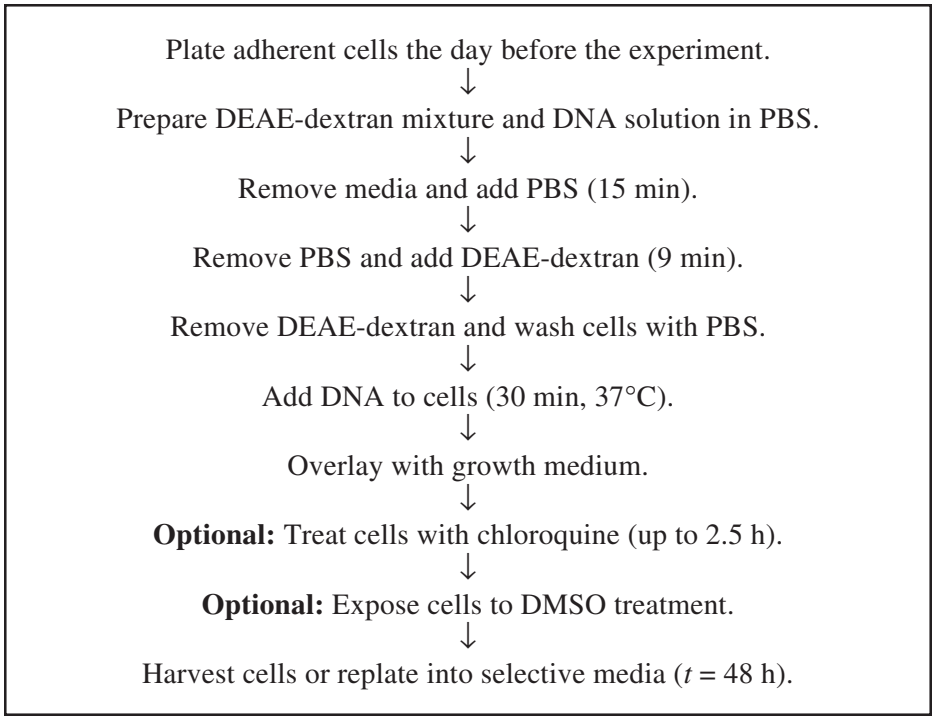


Fig. 2. The DEAE-Dextran Pretreatment Protocol.

9. Gently overlay 3.5 mL of growth medium onto the cells. *Optional:* The growth medium may contain 35 µL of 8 m*M* chloroquine (*see* **Note 8**). Return the cells to a humidified 37°C incubator. If chloroquine was used, replace the transfection solution with complete growth medium after 4 h, or earlier if toxicity is apparent.
10. Harvest the cells approximately 48 h after the start of the transfection (*see* **Note 9**).

3.4. Dimethyl Sulfoxide (DMSO) Shock

The utility of performing a DMSO shock step should be determined for each cell line (*see* **Note 10**). A typical protocol for a DMSO shock is provided below. Volumes provided are for a 60-mm tissue culture plate.

1. Prepare a fresh DMSO shock solution (2 mL per 60-mm plate) and warm to 37°C.
2. Remove the medium from the cells.
3. Add 2 mL of the DMSO shock solution to the cells and incubate for up to 2.5 min at room temperature.
4. Remove the DMSO shock solution and add 5 mL of growth medium. Return the cells to a 37°C incubator.

4. Notes

1. Transfection efficiency may be affected by different batches of serum used in growth medium. It is advisable to compare the prospective new lot of serum with the old one in a transfection experiment.
2. The quality of the DNA is important. The method chosen for DNA purification should produce high quality DNA that works well in transfections. The DNA should be free from contaminants (e.g., an A_{260}/A_{280} ratio of 1.8–1.9; it should look "clean" by agarose gel analysis). The DNA should be sterilized by ethanol precipitation. Several satisfactory methods exist for producing transfection grade plasmid DNA:
   a. The plasmid DNA may be purified by CsCl gradient centrifugation ***(9,10)***. Care must be taken to remove the cesium and ethidium bromide from the final preparation.
   b. Some anion-exchange chromatography procedures (e.g., Qiagen, Chatsworth, CA) work well for producing high-quality plasmid DNA. Some commercial purification systems may leave contaminants in the DNA preparation which interfere with transfection.
3. The concentration of DNA used for transfection should be approximately 1 mg/mL in TE buffer (10 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA) or nuclease-free water. Addition of a large volume of DNA in Tris buffer could change the pH of the mixture.
4. In cotransfection experiments where two or more types of DNA are being transfected, it is important to keep the total amount of DNA constant for each treatment that is to be compared. In situations where additional DNA is required, a neutral plasmid DNA such as the pGem®-3zf+ vector (Promega Corp.) may be added to increase the total amount of DNA in the transfection.
5. For some cell lines, using less DNA than the amounts recommended in the DEAE protocol may yield better results. A dose-response curve can be determined for optimal transfection conditions.
6. Commercial suppliers offer kits containing premade, pretested reagents for use in transfecting cells with the DEAE-dextran method (e.g., Promega Corp., Madison, WI, ProFection® Mammalian Transfection System—DEAE-Dextran System, cat. no. E1210).
7. Rigorous practice of sterile techniques is essential for optimal transfection results. General procedures are discussed in refs. ***11–13***. Mycoplasma or other contaminating organisms can adversely affect the results of a transfection experiment. Obtain cells from a reliable source, and discard cells that have been passaged excessively (i.e., more than 10 times since the freezer stock). Culture cells to provide optimal growth patterns prior to transfection. For most cells, this entails diluting or subculturing the cells every 3–4 d to prevent them from becoming confluent or overly dense.
8. With some cell lines, transfection efficacy is enhanced by the addition of 80 µM chloroquine along with the DNA when using either the standard DEAE-dextran

or the DEAE-dextran pretreatment protocols. For other cells, the use of chloroquine has little effect or causes unacceptable levels of cytotoxicity.

9. If low transfection efficiency occurs along with excessive cell death, several modifications may be made to the DEAE-dextran procedures. The concentration of the DEAE-dextran added to the cells may be decreased, or the incubation time may be shortened. The exposure time to chloroquine may be reduced or omitted. Alternatively, an increase in cell density may help to offset toxicity.
10. DMSO is toxic to cells and the concentration and exposure times require careful optimization for each cell type. Cells are exposed to DMSO for precisely timed intervals, up to 2.5 min.

References

1. Vaheri, A. and Pagano, J. S. (1965) Infectious poliovirus RNA: a sensitive method of assay. *Virology* **27**, 434–436.
2. McCutchan, J. H. and Pagano, J. S. (1968) Enhancement of the infectivity of Simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* **41**, 351–357.
3. Gluzman, Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175–182.
4. Al-Moslih, M. I. and Dubes,G. R. (1973) The kinetics of DEAE-dextran induced cell sensitization to transfection. *J. Gen. Virol.* **18**, 189–193.
5. Yang, Y. W. and Yang, J. C. (1997) Studies of DEAE-dextran-mediated gene transfer. *Biotechnol. Appli. Biochem.* **25**, 47–51.
6. Luthman, H. and Magnusson, G. (1983) High efficiency polyoma DNA transfection of chloroquine treated cells. *Nucl. Acids Res.* **11,** 1295–1308.
7. Lopata, M. A., Cleveland, D. W., and Sollner-Webb, B. (1984) High level transient expression of a chloramphenicol acetyltransferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucl. Acids Res.* **12**, 5707–5717.
8. Zauner, W., Kichler, A., Schmidt, W., Sinski, A., and Wagner, E. (1996) Glycerol enhancement of ligand-polylysine / DNA transfection. *Biotechniques* **20**, 905–913.
9. Kingston, R. E. (1995) Introduction of DNA into mammalian cells, in *Current Protocols in Molecular Biology* (Ausubel, F., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G. Smith, J. A., and Struhl, K., eds.), Wiley, New York.
10. Sambrook, J., Fritsch, E. F., and Maniatis, T., ed. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY.
11. Jakoby, W. B. and Pastan, I. H., eds. (1979) *Methods in Enzymology,* vol. 58, Academic, New York, pp. 3–119.
12. Freshney, R. I., ed. (1986) *Animal Cell Culture*, IRL Press, Oxford, England, pp. 1–11.
13. Freshney, R. I., ed. (1987) *Culture of Animal Cells*, A.R. Liss, New York, pp. 51–59.

13

Liposome-Mediated Transfection of Mammalian Cells

Elaine T. Schenborn and Jennifer Oler

1. Introduction

The term "liposome" originated approximately 30 years ago to refer to lipid bilayers that form colloidal particles in an aqueous medium ***(1)***. Artificially developed liposomes were used effectively by 1980 to deliver DNA to cultured cells ***(2)***. The next advancement in liposome delivery vehicles was the development of synthetic cationic lipids by Felgner and colleagues ***(3)***. Subsequently, many different types of lipids have been synthesized and formulated explicitly for facilitating delivery of DNA or RNA entry into mammalian cells.

Liposome-mediated transfection of mammalian cells offers several advantages over other transfection methods. For example, liposome-mediated delivery can achieve relatively high efficiency of nucleic acid transfer into cultured cells. This method is particularly suited to cell types, such as primary cells, that are very sensitive to toxic effects of other traditional chemical transfer reagents, such as calcium phosphate or DEAE-dextran ***(4)***. Efficient liposome-mediated DNA transfer occurs to cells that grow attached to culture plates or that grow in suspension. Liposome-mediated gene delivery is appropriate for transient or stable paradigms that incorporate DNA into the chromosome. When mouse embryonic stem (ES) cells are the recipients of gene delivery, the end result of stable transfection can be transgenic animals. Yeast artificial chromosomes have been successfully transferred with liposomes to ES cells to produce transgenic mice ***(5)***. In addition to transfection of cells in vitro, liposomes are also in vivo gene delivery reagents, in contrast to traditional chemical methods or electroporation.

Several different types of macromolecules, such as nucleic acids, proteins, and cytotoxic drugs, have been delivered to cells using liposomes, making this

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

method very versatile *(6)*. In the literature, DNA is the most common macromolecule transferred to cells using liposomes. There is no apparent size limitation for the DNA, with successful reports of gene transfer accomplished with oligonucleotides, plasmids, and even yeast artificial chromosomes *(7–9)*. Another type of nucleic acid, RNA, has also been successfully delivered to cultured mammalian cells *(10–12)*. Protein delivery to cultured cells has been reported *(13,14)*, although this type of application is not nearly as common or successful as nucleic acid delivery with liposomes.

A cationic lipid is the most common synthetic lipid component of liposomes developed for delivery of nucleic acids to cells. The cationic region of the molecule associates with negative charges present in the nucleic acids. The synthetic cationic lipid is often formulated into liposomes with other lipids, such as dioleoylphosphatidylethanolamine (DOPE). Several hypotheses have been proposed to explain how the DNA is delivered to the nucleus, but the exact mechanisms of action are still unclear.

Figure 1A depicts a generic cationic lipid structure in liposome formulations. The synthetic lipid is amphiphilic, with a hypdrophilic domain, or "cationic head" region, that often contains a quaternary nitrogen. Other effective variations of the cationic head region include polyamines, such as spermine in the cationic lipid DOGS (dioctadecyldimethylammonium chloride), also known commercially as Transfectam™ *(15)*. The head region associates with negatively charged phosphate groups of nucleic acids, thereby neutralizing the electrostatic repulsion between negatively charged cell membranes and nucleic acids. The "link" region connects the cationic head region to the lipid portion of the molecule. The chemical linkages generally consist of ester bonds that are biodegradable within cells *(16,17)*. Earlier versions of cationic lipids, such as DOTMA *(3)*, contained ether linkages, which are stable chemically and less susceptible to oxidation and bio-degradation than the ester linkages *(3,17)*. Other types of chemical "links" have been incorporated, such as the glycyl linkage present in the cationic lipid DOGS *(15)*. The hydrophobic domain, or lipid region, of the synthetic lipids are diacyl in nature and comprise saturated (e.g., DOGS) or unsaturated (dioleoyl in DOTMA) carbon chains. Acyl chain length is generally between 12 and 18 carbons, with smaller or larger lipids outside this range being less effective for gene transfer applications *(17,18)*.

Liposome formulations may, or may not, contain the lipid DOPE, shown in **Fig. 1B**. This particular lipid increases the gene transfer ability of certain synthetic cationic lipids when included in liposome formulations *(17,19)*. The DOPE may facilitate gene transfer to the nucleus by destabilizing the endosomal compartment, and causing release of lipid/nucleic acid complexes that are trapped in endosomes after cellular uptake *(20)*. The facilitation of gene delivery by DOPE may also be correlated with its fusogenic activity with

A

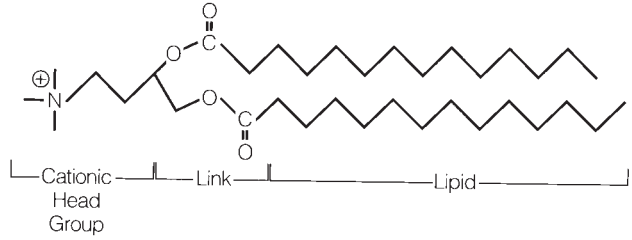


Fig. 1A. Structure of cationic lipids.

B

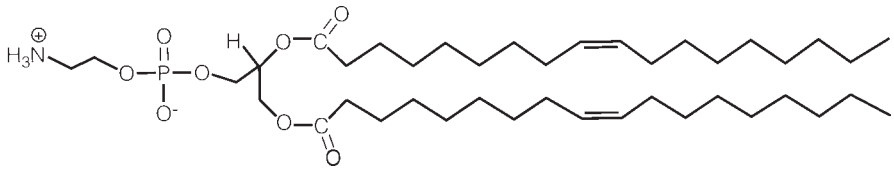


Fig. 1B. DOPE (dioleoylphosphalitidylethanolamine).

cellular membranes, which results in increased cellular uptake of the liposome/nucleic acid complexes ***(21)***.

Liposomes designed for gene transfer are generally prepared in an aqueous solvent and form bilayer structures. Unilamellar vesicles are generally prepared by sonication ***(3,22)***, and the resultant liposome vesicles have sizes ranging from approx 50 to 100 nm ***(17)***. Multilamellar vesicles, 300–700 nm in diameter, can be prepared by vortex action of lipids in an aqueous solvent. Unilamellar and multilamellar vesicles have been both successfully used as gene delivery reagents ***(17)***.

Reporter vectors are recommended for optimization studies. Several reporter systems are designed for easy, sensitive detection of the reporter gene specifically transcribed and translated in transfected cells. The most frequently used reporter gene systems today are luciferase, beta-galactosidase, green fluorescent protein, and chloramphenicol acetyltransferase (CAT) (*see* **ref. *23***, for review). Plasmid DNA vectors are available commercially that are especially designed for high-level expression of the reporter gene in mammalian cells. The reporter systems allow optimization strategies to test multiple variables in a relatively short time frame, particularly with use of multiwell plates. The optimization studies rely upon transient transfection paradigms, and the reporter gene product is generally analyzed two days posttransfection.

2. Materials

The three essential components for in vitro transfection are cultured cells, nucleic acid, and transfection reagent. Selection of the "best" available cationic liposome transfection reagent for a particular cell line must be determined empirically.

1. Growth media with serum: Suitable for the cell type.
2. Serum-free media.
3. Transfection quality DNA (*see* **Note 1**).
4. Liposome preparation: available commercially from suppliers such as Promega Corporation (Madison, WI) or Life Technologies (Gaithersburg, MD).
5. Phosphate-buffered saline (PBS; 1X): 137 m*M* NaCl, 2.7 m*M* KCl, 4.3 m*M* Na_2HPO_4, and 1.47 m*M* KH_2PO_4. Adjust the final pH to 7.1. Filter-sterilize and store at room temperature. A 10X stock may be prepared for convenience and stored at room temperature.
6. Staining solution: 2% methylene blue, 60% methanol.

3. Methods

The following specific protocol is provided using NIH3T3 cells and TransFast™ Reagent (Promega cat. no. E2431). For other cell lines, the optimal amount of reagent and DNA may be different. Optimization for the appropriate amount of lipid, DNA, and transfection interval, presence or absence of serum will need to be determined for any liposome transfection reagent (*see* **Note 2**). Optimization is best performed with a reporter vector. A flowchart for the procedure is shown in **Fig. 2**. The following protocol provides details for the use of TransFast™ Reagent with NIH3T3 cells. Suggestions are provided in the **Notes** for procedural modifications using different reagents and cell lines.

3.1. Protocol for Adherent NIH3T3 Cells

1. The day before transfection, plate the cells into the desired multiwell plates or dishes approximately 24 h prior to performing the transfection. For a 24-well plate, this corresponds to approx 5×10^4 cells such that the cells will be between approx 60% and 80% confluent the day of the transfection (*see* **Note 3**).
2. The day before transfection, prepare the liposome reagent. Suspend the vial of TransFast™ Reagent lipid film with 400 µL nuclease-free water, vortex, and freeze the reagent (*see* **Note 4**).
3. On the day of the transfection, prepare the liposome/DNA complexes. Warm the TransFast™ Reagent to room temperature. To a sterile tube, add serum-free media to a final volume of 200 µL, 1 µg DNA and 3 µL TransFast™ Reagent. Vortex. (These are the volumes for a single well of a 24-well plate. For a 60-mm dish, add media to a final volume of 2 mL, 10 µg DNA, and 30 µL TransFast™ Reagent.) The charge ratio of cationic lipid to anionic DNA is 1:1.

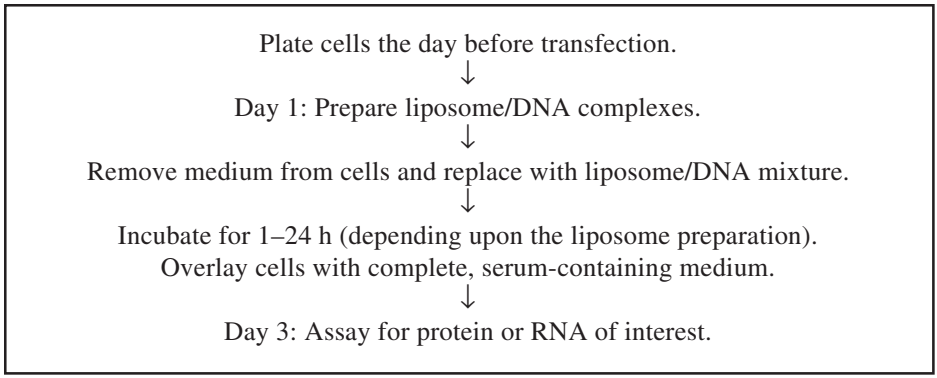


Fig. 2. Flowchart for liposome-mediated transient transfection of adherent cells.

4. Allow the liposome/DNA complexes to form for 10–15 min at room temperature (*see* **Note 5**).
5. Remove the growth media from the cells and add the liposome/DNA complexes to the cells (*see* **Note 6**).
6. Incubate the cells for 1 h in the incubator and overlay with complete, serum-containing media (*see* **Note 7**).
7. Return the cells to the incubator.
8. Harvest protein or RNA from the cells approximately 48 h post-transfection. Carefully wash the cells with PBS to remove media and prepare cell extracts (*see* **Note 8**).

3.2. Protocol for Cells Grown in Suspension

1. On the day of the transfection, calculate the concentration of cells with the aid of a hemacytometer. Determine the number of cells required for the transfection experiments. (1×10^6 cells per transfection is usually sufficient.) Centrifuge the cells for 5 min at 300*g* in a swinging bucket rotor.
2. Remove the media by aspiration, or pipeting, and resuspend the cell pellet in serum-free media to a concentration of 2×10^6 cells/mL. Re-count the cells using a hemacytometer and adjust the volume if necessary.
3. Aliquot 0.5 mL of cells (1×10^6 cells) to each well of a six-well plate.
4. Prepare the liposome/DNA complexes in a sterile tube and add to the cells. For example, with K562 cells, use 4 µg DNA, 12 µL TransFast™ Reagent and serum-free media to a final volume of 0.5 mL per well of the six-well plate. Allow the DNA/lipid complex to form at room temperature for 10–15 min (*see* **Note 2**).
5. Add 0.5 mL of the transfection mixture to each well. Incubate the cells for 1 h and overlay the cells in each well with 5 mL of complete, serum-containing media.
6. Return cells to incubator.
7. Harvest protein or RNA from the cells approximately 48 h post-transfection. First centrifuge the cells as described above in **step 1**. Then remove the media and lyse the cells.

3.3. Modification for Stable Transfectants

1. The cells should be transfected with a plasmid containing a gene for drug resistance, such as neomycin phosphotransferase (*see* **Note 9**). As a negative control, include transfection of cells using DNA that does not contain the gene for the drug-resistance marker.
2. At 48 h post-transfection, adherent cells are trypsinized and replated at several different dilutions (for example, 1:50, 1:100, 1:500) in media containing the appropriate antibiotic.
3. For the next 14 days, replace the drug-containing media every 3–4 days.
4. During the second week monitor the cells for distinct "islands" of surviving cells. Cell death should occur from the negative control transfection using DNA without the drug resistance gene marker.
5. Transfer individual clones by standard means such as cloning rings. Guidelines are provided in ***(24)***.

3.3.1. Procedure to Calculate Stable Transfection Efficiency

The following procedure may be used to determine the percentage of stable transfectants produced after a transfection experiment.

1. After approx 14 days of selection in the appropriate drug, monitor the cultures microscopically for presence of viable cell clones. When distinct "islands" of surviving cells are visible and nontransfected cells have died out, proceed with **step 2**.
2. Prepare the staining solution.
3. Remove the growth media from the cells by aspiration.
4. Add stain to the cells, sufficient to cover the bottom of the dish.
5. Incubate for 5 min.
6. Remove the stain and rinse gently under running cold water. Shake off excess moisture.
7. Allow the plates to air dry. The plates can be stored at room temperature.
8. Count the number of blue colonies and calculate the percentage of transfectants based on the cell dilution and original cell number.

4. Notes

1. The quality of the DNA can affect the transfection efficiency. Impurities such as high EDTA concentration, RNA contamination, and presence of endotoxins can interfere with complex formation between the liposomes and DNA.
2. Optimization is critical for achieving high-efficiency lipid-mediated transfection ***(21,25)***. Certain liposomes work better than others in combination with cultured cells. For any given liposome and cell type, it is important to test variables, such as DNA concentration, ratio of cationic lipid to DNA, and transfection interval. A convenient strategy for optimization is to use reporter genes and multiwell

plating to analyze multiple parameters in systematic testing procedures. Parameters that are important to optimize for all liposomes developed to date include the following.

a. Optimal DNA amount: With many lipids there is often a characteristic "window" for optimal amount of DNA within a given cell type. It is common NOT to see a dose-dependent response with respect to DNA load. For 24-well plates, a range of 0.2–1 μg of DNA per well is usually adequate for determining the optimal amount of DNA.
b. Ratio of cationic lipid to DNA: In general, excess positive charge contributed by the cationic lipid performs better with cultured cells. However, recommended ratios vary for commercially available lipids depending upon supplier protocols. It may be necessary to perform a lipid titration to determine the optimal liposome/DNA ratio for a particular cell line.
c. Transfection interval: Transfection intervals vary widely among the different commercially available lipids from as short as 1 h to as long as 24 h. For any given lipid and cell combination, it may be necessary to run a time course, using the manufacturer's suggested ranges to determine the optimal balance between transfection and toxicity.
d. Presence or absence of serum: The commercially available lipids generally perform best in the absence of serum. Some perform poorly or not at all in the presence of serum. If the desired cell line is sensitive to serum starvation, it may be necessary to find a lipid that performs well in the presence of serum.
e. Selection of lipid: Most commercially available lipids perform differently compared against other lipids in particular cells. If high-transfection efficiency is important, it may be necessary to obtain and optimize several different lipids in the desired cultured cells before selecting a lipid to use in subsequent experiments.

3. Cultured cells used in transfections should be maintained in the appropriate media and at subconfluent densities. Growth characteristics of certain cell lines, such as NIH3T3, can change if cultures are allowed to become confluent. Cells should be subcultured every 3–4 days in media without antibiotics, and proper sterile technique should be practiced at all times to prevent contamination with bacteria, mycoplasma, and fungi. Mycoplasma contamination of cultures will decrease transfection efficiency. It is also best to use cells that have not been passaged for many generations ***(26)***; less than 10 passages is preferable because the transfection properties of the cells change with age in culture.
4. Liposome preparations supplied as sonicated suspensions should not be frozen.
5. Gene transfer protocols include a 10–45 min incubation step of liposomes with nucleic acids, depending upon manufacturer's recommendations. Condensation or compaction of the DNA occurs during this time with concurrent formation of liposome/nucleic acid complexes ***(27)***.
6. The manufacturer's protocol for other cationic liposomes may specify removing the transfection reagents before adding complete medium to the cells.

7. Some manufacturers recommend washing the cells with phosphate-buffered saline (PBS) or serum-free media before applying the liposome/DNA complexes. Perform this step carefully to avoid washing cells off the plate.
8. Liposome/nucleic acid complexes can be toxic to cells. To minimize toxicity, test the effects of decreasing the liposome to DNA ratio, decreasing the amount of DNA, and decreasing the transfection interval. Higher toxicity may also be observed with lower cell density at the time of the transfection. Paradoxically, higher transfection efficiencies may occasionally be associated with higher toxicity.
9. Selection of stable transfected cells is facilitated by transferring DNA that contains a selectable gene marker for drugs such as neomycin or hygromycin **(28–30)**.

References

1. Sessa, G. and Weissmann, G. (1968) Phospholipid spherules (liposomes) as a model for biological membranes. *J. Lipid Res*. **9**, 310–318.
2. Fraley, R., Subramani, S., Berg, P., and Papahadjopoulos, D. (1980) Introduction of liposome-encapsulated SV40 DNA into cells. *J. Biol. Chem.* **255**, 10,431–10,435.
3. Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J. P., Ringold, G. M., and Danielsen, M. (1987) Lipofection: A highly efficient, lipid mediated DNA transfection procedure. *Proc. Natl. Acad. Sci. USA* **84,** 7413–7417.
4. Loeffler, J. P., Batthel, F., Feltz, P., Behr, J. P., Sassone-Corsi, P., and Feltz, A. (1990) Lipopolyamine-mediated transfection allows gene expression studies in primary neuronal cells. *J. Neurochem.* **54,** 1812–1815.
5. Lamb, B.T. and Gearhart, J. D. (1995) YAC transgenics and the study of genetics and human disease. *Current Opinion in Genetics and Development* **5,** 342–348.
6. Gregoriadis, G. (1995) Engineering liposomes for drug delivery: Progress and problems. *Trends Biotechnol.* **13,** 527–537.
7. Capaccioi, S., DiPasquale, G., Mini, E., Mazzei, T., and Quattrone, A. (1993) Cationic lipids improve antisense oligonucleotide uptake and prevent degradation in cultured cells and in human serum. *Bioch. Biophys. Research Communications* **197,** 818–825.
8. Felgner, J. Bennett, F., and Felgner, P. L. (1993) Cationic lipid mediated delivery of polynucleotides. *Methods* **5,** 67–75.
9. Lee, J. T. and Jaenisch, R. (1996) A method for high efficiency YAC lipofection into murine embryonic stem cells. *Nucl. Acids Res.* **24,** 5054–5055.
10. Wilson, T., Papahadjopoulos, D., and Taber, R. (1979) The introduction of poliovirus RNA into cells via lipid vesicles (liposomes). *Cell* **17,** 77–84.
11. Malone, R.W., Felgner, P. L., and Verma, I. M. (1989) Cationic liposome-mediated RNA transfection. *Proc. Natl. Acad. Sci. USA* **86,** 6077–6081.
12. Leopold, L. H., Shore, S. K., Newkirk, T. A., Reddy, R. M., and Reddy, E. P. (1995) Multi-unit ribozyme-mediated cleavage of bcr-abl mRNA in myeloid leukemias. *Blood* **85,** 2162–2170.
13. Debs, R. J., Freedman, L. P., Edmunds, S., Gaensler, K. L., Duzgunes, N., and Yamamoto, K. (1990) Regulation of gene expression *in vivo* by liposome-mediated delivery of a purified transcription factor. *J. Biol Chem.* **265,** 10,189–10,192.

14. Gao, S. and Huang, L. (1993) Cytoplasmic expression of a reporter gene by codelivery of T7 RNA polymerase and T7 promoter sequence with cationic liposomes. *Nucleic Acids Res.* **21**, 2867–2872.
15. Behr, J. P., Demeneix, B., Loeffler, J. P., and Mutul, J. P (1989) Efficient gene transfer into mammalian primary endocrine cells with lipopolyamine-coated DNA. *Proc. Natl. Acad. Sci. USA* **86,** 6982–6986.
16. Leventis, R. and Silvius, J. R. (1990) Interactions of mammalian cells with lipid dispersions containing novel metabolizable cationic amphiphiles. *Biochem. Biophys. Acta* **1023,** 124–132.
17. Felgner, J. H., Kumar, R., Sridhar, C. N., Wheeler, C. J., Tsai, Y. J., Border, R., Ramsey, P., Martin, M., and Felgner, P. L. (1994) Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. *J. Biol. Chem.* **269,** 2550–2561.
18. Balasubramaniam, R. P., Bennett, M. J., Aberle, A. M., Malone, J. G., Nantz, M. H., and Malone, R. W. (1996) Structural and functional analysis of cationic transfection lipids: the hydrophobic domain. *Gene Therapy* **3,** 163–172.
19. Wheeler, C. J., Sukhu, L., Yang, G., Tsai, Y., Bustamente, C., Felgner, P., Norman, J., and Manthorpe, M. (1996) Converting an alcohol to an amine in a cationic lipid dramatically alters the co-lipid requirement, cellular transfection activity and the ultrastructure of DNA-cytofectin complexes. *Biochim. Biophys. Acta* **1280,** 1–11.
20. Farhood, H., Servina, N. S., and Huang, L. (1995) The role of dioleoylphosphatidylethanoloamine in cationic liposome mediated gene transfer. *Biochim. Biophys. Acta* **1235,** 289–295.
21. Litzinger, D. C. and Huang, L. (1992) Phosphatidyethanolamine liposomes: drug delivery, gene transfer and immunodiagnostic applications. *Biochim. Biophys. Acta* **1113,** 201–227.
22. Papahadjopoulos, D., Hui, S., Vail, W. J., and Poste, G. (1976) Studies on membrane fusion. Interactions of pure phopholipid membranes and the effect of myristic acid, lysolecithin, proteins and dimethysulfoxide. *Biochim. Biophys. Acta* **448,** 254–264.
23. Groskreutz, D. and Schenborn, E. (1997) Reporter systems, in *Recombinant protein protocols: Detection and isolation* (Tuan, R., ed.), Humana Press, Totowa, NJ, pp. 11–28.
24. Jakoby, W. B. and Pastan, I.H., ed. (1979) *Methods in Enzymology* **58**. Academic Press, New York, NY.
25. Caplen, N. J., Kinrade, E., Sorgi, F., Gao, X., Gruenert, D., Geddes, D., Coutelle, C. Huang, L., Alton, E., and Wiliamson, R. (1995) *In vitro* liposome-mediated DNA transfection of epithelial cell lines using the cationic liposome DC-Chol/DOPE. *Gene Therapy* **2,** 603–613.
26. Jarnagin, W. R., Debs, R.J., Wang, S-S., and Bissell, D. M. (1992) Cationic lipid-mediated transfection of liver cells in primary culture. *Nucleic Acids Res.* **20,** 4205–4211.
27. Kabanov, A. V. and Kabanov, V. A. (1995) DNA complexes with polycations for the delivery of genetic material into cells. *Bioconjugate Chem.* **6,** 7–20.

28. Stuhlmann, H., Cone, R. D., Mulligan, R. C., and Jaenisch, R. (1984) Introduction of a selectable gene into different animal tissue by a retrovirus recombinant vector. *Proc. Natl. Acad. Sci. USA* **81,** 7151–7155.
29. Southern, P. J. and Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Gen.* **1,** 327–341.
30. Gritz, L. and Davies, J. (1983) Plasmid-encoded hygromycin-B resistance: the sequence of hygromycin-B-phosphotransferase gene and its expression in *E. coli* and *S. cerevisiae*. *Gene* **25,** 179–188.

14

Assays for Transcriptional Activity Based on the Luciferase Reporter Gene

S. Roy Himes and M. Frances Shannon

1. Introduction

Reporter genes provide easy and efficient methods for the indirect measurement of relative rates of transcription. Utilizing common DNA cloning methods, a putative regulatory region can be coupled to the coding sequence of a reporter gene such that expression of the reporter protein product varies according to the regulatory potential of the DNA tested. The assays for reporter enzymes have the advantage of high sensitivity with low background, and, although an indirect measure, the amount of protein product is usually directly proportional to the level of transcriptional activation. Alternatives to reporter gene assays such as the direct measurement of the level of specific mRNAs for the endogenous gene can be influenced by RNA stability changes as well as transcription rates in response to stimulation. Assays for mRNA are also more labor intensive and difficult to quantify.

The use of reporter systems allows the fine mapping of transcription control regions into specific enhancer, promoter, and silencer regulatory elements. The region of DNA containing the full potential for transcriptional activation can be sequentially deleted and tested in the reporter system to identify the types of elements present. Lost or enhanced activation of reporter activity is easily measured and corresponds to active elements within the DNA. These individual elements can be further characterized by inserting one or multiple copies of the element immediately upstream of a heterologous basal promoter in the reporter construct (*see* **Note 1**). Weak, constitutive promoters, such as the minimally active fragment of the herpes simplex virus thymidine kinase promoter ***(1)***, are usually used. An enhancer region, for example, should confer a stimulus-specific or cell-specific transcriptional response on the basal promoter with the

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

DNA fragment inserted in either orientation. A silencer element, on the other hand, when used with a strong promoter such as a viral promoter ***(1)***, would result in an inhibition of reporter expression.

Reporter genes are again useful in the analysis of the transcription factors that bind DNA at specific *cis*-acting elements and control the function of these elements. Mutations can be made in specific regulatory elements within enhancer/promoter regions, which prevent protein binding and the effect on reporter expression tested. Gain or lack of function experiments for these transcription factors can also be quantified by reporter assay. For example, transcription factors can be overexpressed or their expression inhibited with antisense methods to determine the role of individual factors in the activity of a specific element. Finally, reporter systems can provide a useful endpoint measurement for receptor structure-function studies and the delineation of components within signal transduction pathways.

Numerous reporter systems have been developed to accurately measure transcription rates. The two most commonly used systems are based on the chloramphenicol acetyltransferase (CAT) gene and the firefly (*Photinus pyralis*) luciferase gene ***(2,3)***. This chapter, however, will only provide techniques for using the luciferase reporter system. Although both assay systems can provide an accurate and linear quantitation of transcription rates, they differ in sensitivity, ease of performance, quantifiable range, and the stability of proteins within the cell and of the enzyme reaction itself. The main advantages of the luciferase system are the nonisotopic assay of bioluminescence, the high degree of sensitivity, the absence of endogenous enzyme activity in eukaryotic cells, the wide linear range of quantitation, and the ease and speed of performance. The production of advanced instrumentation such as luminometers that read 96-well plates allows an even more rapid and extensive collection of data. Techniques used for the analysis of cytokine gene promoters in the Jurkat T-cell line will be used as an example throughout the chapter.

2. Materials

2.1. Cell Culture and Transfection

1. Jurkat T-cells: A subline of the cell type available from the American Type Culture Collection.
2. RPMI medium: RPMI 1640 (Gibco-BRL, Grand Island, NY) supplemented with 100 U/mL each of penicillin and streptomycin antibiotics, 10 m*M* Hepes**,** 0.25% sodium bicarbonate, and 10% fetal calf serum (FCS) (heat inactivated at 56°C).
3. Phorbol-12-myristate-13-acetate (PMA; Sigma, St. Louis MO): Dissolve in DMSO to a concentration of 1 mg/mL, aliquot, and store at –70°C.
4. Calcium ionophore (A23187, Boehringer Mannheim, Germany): Dissolve in DMSO to a concentration of 10 m*M*, aliquot, and store at –70°C.

2.2 Stock Solutions and Reagents

1. 1 *M* potassium phosphate buffer: Combine four parts 1 *M* K_2HPO_4 with one part 1 *M* KH_2PO_4, bring to pH 7.8, and autoclave.
2. 1 *M* $MgSO_4$: Make using Milli-Q purified water and autoclave.
3. 5 m*M* coenzyme A: Dissolve free acid (Boehringer Mannheim) in Milli-Q water, aliquot, and store at –20°C.
4. 100 m*M* ATP: Dissolve ATP (Boehringer Mannheim) in Milli-Q water, aliquot, and store at –20°C.
5. 1 m*M* D-luciferin: Dissolve D-luciferin (Boehringer Mannheim) in 5 m*M* potassium phosphate buffer (*see* **Note 2**).
6. 1 *M* DTT: dissolve dithiotheitol in Milli-Q water, aliquot, and store at –70°C.
7. Phosphate-buffered saline (PBS): Dissolve 0.2 g KCl, 8.0 g NaCl, 0.2 g K_2HPO_4 and 2.9 g $Na_2PO_412H_2O$ in 800 mL water, adjust to pH 7.4, make to a final volume of 1 L and autoclave.
8. Cell lysis solution: 100 m*M* potassium phosphate buffer, 2 m*M* DTT, and 1 m*M* EDTA.
9. Detergent lysis solution: 50 m*M* potassium phosphate buffer, 2% Triton X-100, 20% glycerol, and 4 m*M* DTT.
10. Firefly luciferase assay buffer: 100 m*M* potassium phosphate buffer, 2 m*M* DTT, 8 m*M* $MgSO_4$, 175 µ*M* coenzyme A, and 750 µ*M* ATP.
11. Neomycin (Gibco-BRL).

2.3. Equipment

1. Bio Rad (Hercules, CA) Gene Pulser electroporator or equivalent and electroporation cuvets.
2. Liquid scintillation counter.
3. Packard (Meredin, CT) Top Count scintillation/luminescence counter or equivalent.
4. Table top centrifuge.
5. Microcentrifuge.
6. Tissue culture incubator with 5% CO_2.
7. 37°C water bath.

3. Methods

3.1. Cell Transfection

3.1.1. Transient Transfection of Reporter Constructs (*see* **Note 3**)

1. Passage Jurkat T-cells 24 h prior to transfection so that the cell density will reach $5–6 \times 10^5$ cells/mL of culture at the time of transfection.
2. Place RPMI medium and an aliquot of FCS in a 37°C water bath to warm.
3. Pellet cells at 400*g* for 5 min in a bench-top centrifuge.
4. Remove the supernatant by aspiration and resuspend the cell pellet to approx 3×10^7 cells per mL in RPMI medium.
5. Determine the total number of cells by counting with a hemocytometer.

6. Dilute cells to a concentration of 5×10^6 cells per 300 µL with RPMI medium.
7. Pipet the appropriate amount of purified plasmid DNA onto the side wall of an electroporation cuvet (*see* **Notes 4** and **5**).
8. Add 300 µL of cells, pipeting against the cuvet side to mix with the DNA. Cells in stock tube should be mixed by shaking before addition to each cuvet to prevent settling in the tube.
9. Add 30 µL of FCS to each cuvet.
10. Electroporate each cuvet at 270 V with a capacitance of 960 µF (*see* **Note 5**). Resuspend cells by shaking cuvet immediately before electroporation.
11. Rest the cells for 5–10 min at room temperature.
12. Remove cells from the cuvet with a sterile pastuer pipet and place in a tissue culture flask with 5 mL of RPMI medium.
13. Allow cells appproximately 24 h to recover before further processing (*see* **Note 6**).

3.1.2. Stable Transfection of Reporter Constructs (*see* **Notes 7** and **8**)

1. Transfect Jurkat T-cells according to the procedure described in **Subheading 3.1.1.** using 10 µg of luciferase reporter plasmid and 1 µg of a plasmid containing the neomycin or other selectable marker.
2. Add neomycin to give a final concentration of 500 µg/mL to the medium when cells have doubled twice posttransfection (*see* **Note 9**).
3. Culture the cells in neomycin until the control transfected cells all die.
4. Clone neomycin-resistant cells by a suitable procedure for obtaining individual cells (*see* **Note 10**). A cell sorter or limiting dilution plating in 96-well plates can be used.
5. The number of distinct integration events must be determined by Southern blot analysis of individual clones (***4,5***).

3.2. Treatment of Transfected Cells

Twenty-four hours after transient transfection or after selection of stable cell lines, cells are stimulated to activate transcription from the promoter being investigated.

1. Thaw stocks of PMA and calcium ionophore immediately before use.
2. Add sufficient PMA and calcium ionophore to cells to have a final concentration of 20 ng/mL and 1 µ*M* respectively.
3. Harvest the cells for luciferase assay as described below (**Subheading 3.4.**) at a range of time points following stimulus addition (*see* **Note 10**).
4. Choose the time-point with the highest level of luciferase activity to carry out all subsequent experiments.

3.3. Assay for Luciferase Activity in Cell Lysates (*see* **Notes 10–12**)

3.3.1. Preparation of Cell Lysates

1. Centrifuge the cells in a conical centrifuge tube at 400*g* for 5 min at 4°C.

2. Resuspend the cell pellet in 10 mL of ice cold PBS, repellet, and wash a second time with PBS.
3. After the final wash, use a micropipet to remove any residual PBS from the cell pellet.
4. Resuspend the cell pellet in 100 µL of cell lysis buffer and transfer to a 1.5 mL microcentrifuge tube.
5. Lyse cells with three rounds of freeze–thaw by immersing tubes in liquid nitrogen or in ethanol/dry ice to freeze and subsequent immersion in a room-temperature water bath to thaw (*see* **Note 13**).
6. Spin tubes in a microcentrifuge at 14,000*g* for 5 min at 4°C to pellet cell debris.
7. Transfer the supernatant to a new tube and place on ice for assay of protein concentration.
8. Determine the concentration of protein in each cell lysate by a suitable procedure such as the Bradford assay ***(6)***.
9. Store lysates at 4°C until the assay for luciferase activity is performed.

3.3.2. Assay for Luciferase Activity Using a Scintillation Counter

1. Dilute samples to be assayed to approximately the same protein concentration with cell lysis buffer and keep at 4°C.
2. Assemble the assay buffer from stock components at 4°C.
3. Add 400 µL of assay buffer to a second set of tubes and allow them to equilibrate to room temperature (*see* **Note 14**).
4. Thaw frozen aliquots or prepare fresh luciferin immediately before assay is performed (*see* **Note 15**).
5. The sample to be tested should be kept at 4°C and brought to room temperature immediately before assay.
6. Set the scintillation counter on manual mode with the coincidence circuit switched off to allow single-photon counting (*see* **Note 16**).
7. Add a defined amount of protein from the cell lysates to the assay buffer tube.
8. Add 40 µL of 1 m*M* D-luciferin.
9. Mix and immediately place the tube into a scintillation vial for measurement (*see* **Note 17**).
10. Repeat **steps 6–8** for each sample individually to avoid decay of the signal (*see* **Note 18**).

3.4. Assay for Luciferase Acitivity in 96-Well Tissue Culture Plates (*see* Note 19)

1. Transfect Jurkat T-cells as described in **Subheading 3.1.1.** and allow to recover in tissue-culture flasks for 24 h.
2. After 24 h determine the number of viable cells by removing a small aliquot (~500 µL).
3. Add an equal volume of trypan blue to the small aliquot of cells and let sit for 5 min.
4. Count the number of viable cells (cells that have not taken up the blue dye) using a hemocytometer.

5. Pellet the remaining cells at 400*g* for 5 min and remove all the medium by aspiration.
6. Resuspend the cells in sufficient RPMI medium without phenol red to give a concentration of 1×10^6 viable cells per mL (*see* **Note 20**).
7. Add 100 µL of cells (1×10^5 cells) to each well in the 96 well plate.
8. Add the stimulus to the cells in a volume of 10 µL or less.
9. At the appropriate time poststimulation, lyse the cells by adding 50 µL of detergent lysis buffer to each well and mix by gently tapping and shaking the plate.
10. Leave cells to lyse at room temperature for 15 min (*see* **Note 21**).
11. Assemble luciferase assay buffer from stock solutions at 4°C in a sufficient amount to add 50 µL of solution to each well.
12. Allow the assay buffer to equilibrate to room temperature.
13. Immediately before adding buffer to the plate, thaw 1 m*M* D-luciferin and add 100 µL per mL of assay buffer.
14. Add 50 µL assay buffer/luciferin to each well using a multichannel micropipetor and mix by gently tapping and shaking the plate.
15. Place in the 96-well plate luminometer for light measurement.

4. Notes

1. Plasmids containing multiple copies of the same insert are susceptible to recombination, particularly if all the inserts are not in the same orientation. The plasmid DNA should be amplified in a RecA-strain of bacteria, e.g., JM109.
2. D-luciferin is not soluble at acidic pH and will racemize more rapidly at basic pH. Although the potassium phosphate buffer is slightly basic, it can be used to dissolve the luciferin. The luciferin powder is acidic and will bring down the pH of the buffer. If all of the luciferin does not dissolve, add small amounts of the stock buffer until all powder is dissolved. The solution should be aliquoted and frozen at –70°C immediately.
3. The method of DNA transfection will vary according to the cell type. The efficiency of the transfection should be optimized, however, and be stable and reproducible. Because transfection efficiency can depend on the immediate viability of the cells, highly consistent transfection efficiency over time can be difficult to maintain. It is often necessary to thaw out a new aliquot of cells for transfection every 6–8 wk. Individual experiments should contain sufficient cells to transfect all reporter plasmids being analyzed in the study with several replicate transfections for each plasmid. The use of an internal control in each transfection can allow normalization of transfection efficiency within each experiment and provide a more reliable quantitation of transcription rates. The use of internal controls will be discussed below.
4. The amount of DNA to be transfected will depend on a number of variables, the most important of which are the transfection efficiency for the cell type and the level of transcriptional activity from the DNA being tested. Amounts ranging from 1–30 µg should be tested and the amount required to detect activity measured empirically. Plasmids whose transcriptional activity is to be compared

directly should be carefully quantitated at the same time with both spectrophotometric analysis and agarose gel electrophoresis.

5. A second reporter gene can be used to minimize experimental variability that may arise during transfection and sample processing. The second reporter system is used as an internal control to which the firefly luciferase reporter is normalized. The second reporter gene is usually placed under the control of a constitutive promoter. The CAT gene as a coreporter is suitable only for measurement of luciferase in cell lysates since the CAT assay needs to be performed seperately. The second method using the Renilla luciferase gene can be used with 96 well plates. The Renilla Dual Luciferase™ assay system is available from the Promega Corporation (Madison, WI), and the buffers and reagents used are proprietary.
6. The electroporation conditions will differ for each cell type. A death curve must be performed for each cell type plotting increasing voltage vs cell viability posttransfection. A plasmid that will allow measurement of the transfection efficiency should be used. A useful plasmid for this purpose is a CMV- or RSV-driven β-galactosidase reporter, where individual cells can be assayed by blue color reaction in a β-galactosidase assay *(7)*. At least 50% of the cells will be killed by electroporation to achieve acceptable levels of transfection. Transfection efficiency will vary from 10% to 60% depending on the cell type.
7. The length of time a particular cell line requires for recovery after transfection can be roughly determined by the time-point at which the cells double at the same rate as nontransfected cells.
8. Analysis of gene regulation sometimes requires stable integration of reporter constructs into the cell genome. Several problems can occur in stable transfectants associated with random integration of plasmid DNA into the genome. The activation potential of regulatory regions can change dramatically depending on the site of integration into the chromatin. This variability occurs because of influence from regulatory elements in the chromatin flanking the inserted DNA. Insertion of the reporter construct within a locus control region containing boundary elements can overcome this problem, but is impractical for most studies. Variable expression can be minimized by using pools rather than clones of transfected cells. If stimulus-specific induction is to be analyzed, then clones that contain inappropriate, constitutive reporter expression should be removed. The CAT reporter system is often inappropriate for these studies because significant levels of protein can accumulate within the cell even with weak constitutive activity. The luciferase enzyme has a short half-life in most cells and for this reason may give clearer simulus-specific responses and is preferable for most studies using stable transfectants.
9. A suitable neomycin selectable plasmid is pRCCMV (Invitrogen Inc., San Diego, CA). Other selectable markers such as puromycin resistance can also be used. It is not suitable to generate a reporter plasmid containing a second promoter/enhancer to drive the expression of neomycin resistance because this strong promoter may interfere with the measurement of activity from the test promoter. The plasmid with the selectable marker should be cotransfected at a 1:10 ratio with the promoter reporter plasmid.

10. A suitable time-course for cytokine gene promoters in T cells is every 2 h for up to 12 h and then 16 h and 24 h.
11. Jurkat cells are highly resistant to G418 used to select for neomycin expression and hence high doses must be used in the selection. The appropriate dose must be determined empirically for each cell type by titration of G418 with non-transfected cells. The minimum level of G418 that kills all the cells within a 7–14 d period is used for selection.
12. Because of the low solubility in aqueous solutions of these compounds and in order to avoid toxicity of DMSO to the cells, the stock solutions are first diluted 1:1000 in RPMI media before addition to the cells.
13. The luciferase enzyme is unstable with an approximate half-life of 3 h in most mammalian cell lines ***(8)***. This instability is due to the targeting of the luciferase protein to peroxisomes ***(9)***. Reporter vectors that encode forms of luciferase which have peroxisomal localization sequences removed have been engineered and are available from the Promega Corporation (Madison, WI). Because the level of stability is dependent on cell type, a time-course must be performed for the reporter construct with each cell type and stimulus to be used in order to determine the time-point of maximum luciferase expression.
14. The luciferase enzyme catalyses a bioluminescent reaction that requires the substrate D-luciferin, ATP, Mg^{2+}, and O_2. The basic bioluminescence reaction results in a flash of light that rapidly decays because of products from the reaction feeding back to inhibit catalysis by luciferase. The luciferase reaction can occur with minimal feedback inhibition if coenzyme A is used as a substrate, owing to the formation of the intermediate luciferyl-CoA ***(10)***. The use of coenzyme A results in a light signal with a longer duration and increased intensity. Although the reaction is still unstable, the light signal can be sustained from 5 to 10 min, compared to less than a minute in the basic reaction. This time frame permits light measurement in machines that do not have an automated injection device.
15. Cells should be lysed by freeze/thaw instead of detergent lysis. The luciferase enzyme is more unstable in extracts from detergent lysis and suitable only for assay immediately after cell lysis.
16. The temperature optimum for the bioluminescence reaction is approximately room temperature, from 20 to 25°C.
17. Scintillation counters are equipped with a coincidence circuit to minimize interference from chemiluminescence in radioisotope measurements. This circuit must be switched off in the programming before doing a bioluminescence assay. If the coincidence circuit cannot be switched off, a formula can be used to establish a linear relationship between the concentration of luciferase enzyme and the counts per minute (CPM) measured. A calculation of the square root of measured CPM minus background CPM should be made, $(\text{sample} - \text{background})^{1/2}$ ***(11)***.
18. The microcentrifuge tube does not usually interfere with light detection in the counter.
19. Measurement with a luminometer is identical to the above procedure except the reaction is initiated by addition of luciferin with an autoinjection device.

20. Transfected cells can be directly assayed for luciferase activity from tissue-culture wells. The processing of samples on 96-well plates allows the efficient measurement of multiple samples. This method can also be used to evaluate large sets of compounds, such as drug libraries, using reporter activity as an end-point measurement. The assay is not as sensitive as the procedure above and should not be used if the level of reporter expression is low. The assay also requires a plate luminometer such as a TopCount scintillation counter/luminometer (Canberra-Packard, Meriden, CT). The cells are cultured in either white or black 96 well plates (Canberra-Packard). White plates allow more sensitive measurement of light but have a high background and crosstalk between wells. Most of the background in white plates is due to plate fluorescence from exposure to light. This fluorescence can be minimized by placing pates in the dark during cell lysis and adding reagents under low light. The leakage of light from highly reactive wells into more negative wells can be minimized by not plating strong positives next to weak signals. If the level of luciferase activity is high enough, then black plates should be used to avoid these problems.
21. The RPMI medium should not contain phenol red because the dye quenches the luciferase bioluminescence.
22. Prepare reagents for assay while cells are being lysed because the assay should be performed as soon as possible after lysis.

References

1. Wegner, R. H., Moreau, H., and Nielsen, P. J. (1994) A comparison of different promoter, enhancer and cell type combinations in transient transfections. *Anal. Biochem.* **221,** 416–418.
2. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Recombinant genomes which express chloroamphenicol acetyltransferase in mammalian cells. *Mol. Cell Biol.* **2,** 1044–1051.
3. Wood, K.V. (1991) *Chemiluminescence: Current Status* (Stanley, P. and Cricka, L., eds.), Wiley, New York, p. 543.
4. Southern E. M. (1975) Detection of specific sequences among DNA fragments seperated by gel electrophoresis. *J. Mol. Biol.* **98,** 503–517.
5. Current protocols in Molecular Biology, vol 1 (Ausubel, F. M., et al., eds) Wiley, New York.
6. Bradford, M. M. (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72,** 248–254.
7. Rihs, H.-P. and Peters, R. (1989) Nuclear transport kinetics depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the SV40 T antigen. EMBO J. 8, 1479–1484.
8. Thompson, J. F., Hayes, L. S., and Lloyd, D. B. (1993) Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene* **103,** 171–177.
9. Gould, S. J. and Subramani, S. (1988) Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.* **7,** 5–13.

10. Bronstein, I., Fortin, J., Stanley, P. E., Stewart, G. S., and Kricka, L. J. (1994) Chemiluminescent and bioluminescent reporter gene assays. *Anal. Biochem.* **219,** 169–181.
11. Nguyen, V. T., Morange, M., and Bensuade, O. (1988) Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Anal. Biochem.* **171,** 404–408.

15

Transient Transfection of Schneider Cells in the Study of Transcription Factors

Guntram Suske

1. Introduction

Transfections into Schneider's *Drosophila* line 2 (abbreviated SL2 or S2) derived from *Drosophila* embryos ***(1)*** have been used to analyze activation properties of mammalian transcription factors (**Table 1**), to identify activation and inhibitory domains, as well as to investigate specific protein-protein interactions in vivo. SL2 cells are particularly suited to this task because they are devoid of many ubiquitous mammalian transcription factor activities and thus their transcriptional properties can be investigated in the absence of interference by endogenous factors. The usage of SL2 cells, as a host for studying the structure and function of heterologous, mammalian transcription factors, was originally reported by the group of R. Tjian to identify functional domains of AP-2 ***(2)***, Sp1 ***(3,4)*** and CTF/NF-I ***(5)***.

Figure 1 illustrates the assay in its simplest conception. A reporter plasmid containing a natural or artificial promoter with binding sites for a given transcription factor fused to a reporter gene (commonly, the bacterial chloramphenicol acetyltransferase [CAT] gene or the firefly luciferase [Luc] gene) is cotransfected along with an expression vector for the transcription factor (TF) to be analyzed. Expression of the transcription factor is driven by a strong *Drosophila* promoter, for instance, the actin 5C or the alcohol dehydrogenase promoter. The transcription factor will be transcribed and translated in the cells and subsequently bound to its recognition sequence present on the reporter plasmid, thus controlling the expression of chloramphenicol acetyltransferase, luciferase, or any other reporter system. At a given time following transfection, the cells are lysed and assayed for enzymatic activity. Enzymatic activity correlates directly with transcription factor activity. As a control, the DNA

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

Table 1
Analysis of Heterologous Transcription Factors in SL2 Cells[a]

Transcription Factor	Vector	Promoter	Reference
Zinc finger proteins			
EKLF	pPac	Actin 5C	***(12)***, ***(11)***
c-Krox	pRmHa3	Metallothionein	***(13)***
Krox-20	pPac	Actin 5C	***(14)***
Sp1	pPac	Actin 5C	***(3)***
	pAdh	Alcohol dehydrogenase	***(3)***
Sp3	pPac	Actin 5C	***(15)***
Sp4	pPac	Actin 5C	***(16)***
Sp1(rat)	pGEMAct	Actin 5C	***(17)***
YY1	pPac	Actin 5C	***(18)***
Nuclear receptors			
GR	pPac	Actin 5C	***(10)***
T3R	pPac	Actin 5C	***(19)***
RXR	pPac	Actin 5C	***(19)***
AhR	pGEMAct	Actin 5C	***(17)***
Arnt	pGEMAct	Actin 5C	***(17)***
Others			
AP2	pAdh	Alcohol dehydrogenase	***(2)***
CTF/NF1	pAdh	Alcohol dehydrogenase	***(5)***
NF-1	pPac	Actin 5C	***(20)***
NF1-C2	pPac	Actin 5C	***(21)***
Ets1	pPac	Actin 5C	***(22)***
p53	pPac	Actin 5C	***(23)***
c-rel	pPac	Actin 5C	***(24)***
GATA-1	pPac	Actin 5C	***(21)***, ***(12)***
	pmt	Metallothionein	***(11)***
SREBP	pPac	Actin 5C	***(25)***
NFI-X	pAdh	Alcohol dehydrogenase	***(26)***
E2F1	pPac	Actin 5C	***(27)***
C/EBPα	pGac	Actin 5C	***(28)***
C/EBPß	pGac	Actin 5C	***(28)***

[a]The table does not include transcription factors from insects.

binding activity in the lysates is monitored by an electrophoretic mobility shift assay (EMSA).

A general strategy for the analysis of the transcriptional properties of a mammalian transcription factor in SL2 cells is as follows.

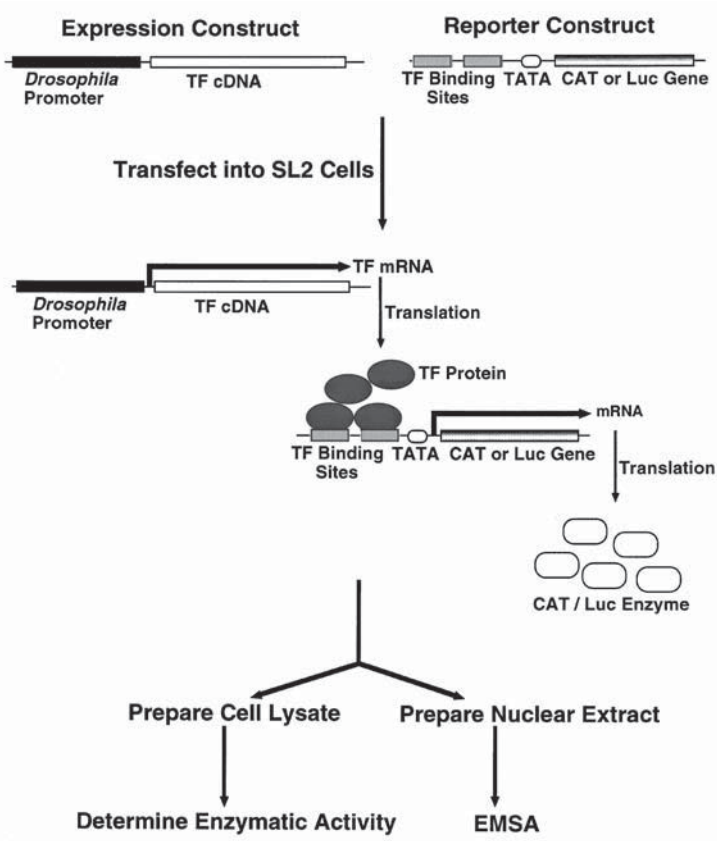


Fig. 1. Diagram outlining the use of transient transfections into SL2 cells to study transcription factors.

1. The DNA recognition sequence, or at least a promoter which contains binding sites for the transcription factor, should be known. If the binding site is unknown, it is determined by techniques such as EMSA, DNaseI footprinting, and dimethyl sulfate methylation protection.
2. A reporter plasmid has to be constructed. The design of the reporter plasmid includes the fusion of a responsive promoter to a reporter gene (*see* **Note 1**). The CAT gene or the luciferase gene from the firefly *Photinus pyralis* are commonly used for this task (*see* **Note 2**).
3. An expression vector for the complete transcription factor has to be constructed. The expression plasmid used for transcription/translation of the recombinant transcription factor in SL2 cells typically contains the transcription factor cDNA flanked by a strong constitutive enhancer/promoter region of *Drosophila* origin and a polyadenylation signal (*see* **Note 3**).

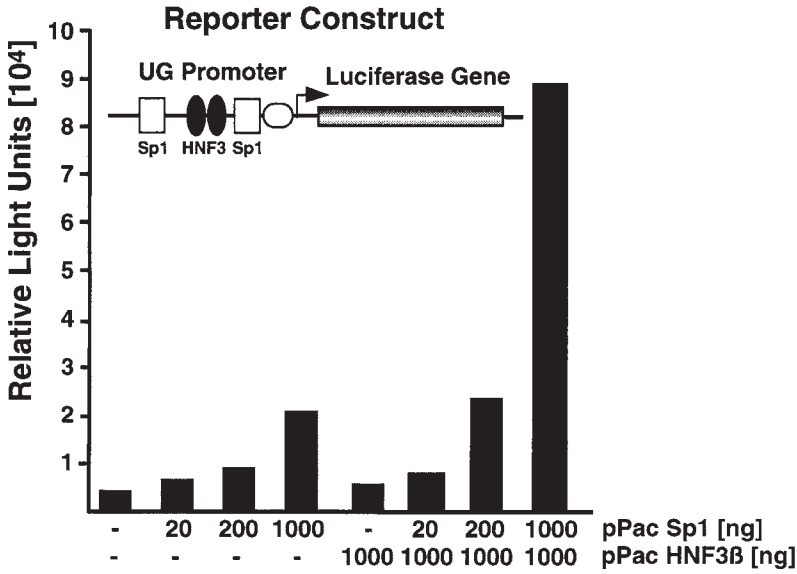


Fig. 2. Sample transfection showing synergistic activation of the rabbit uteroglobin/CC10 (UG) promoter by the transcription factors HNF3ß and Sp1 in transfected SL2 cells. Four micrograms of the UG-luciferase reporter plasmid containing binding sites for HNF3 and Sp1 were transfected into SL2 cells, along with variable amounts of the plasmid pPacSp1 (20, 200, and 1000 ng) in the presence or absence of 1 µg of an expression plasmid for HNF3ß (pPacHNF3ß). The cells were subsequently lysed and luciferase activities determined. HNF3ß alone did not activate transcription from the uteroglobin/CC10 promoter. However, the modest activation by Sp1 alone was enhanced in the presence of HNF3ß.

4. Following DNA transfer into SL2 cells: i) the expression of the transcription factor, ii) its binding to the recognition sequence, and iii) its transcriptional activity need to be determined in the transiently transfected SL2 cells. Expression of the transcription factor can be tested by Western blot analysis (*see* **Note 4**). DNA binding (and expression) is analyzed by EMSA with nuclear extracts prepared from transiently transfected SL2 cells. Transcriptional activity is tested by cotransfection of the expression vector along with an appropriate reporter construct (see above). Usually, a series of parallel transfections is set up using a constant amount of reporter plasmid and a variable amount of expression plasmid.
5. The action (additive, synergistic, or competitive) of different transcription factors on a promoter containing binding sites for multiple transcription factors is analyzed by cotransfection of the reporter plasmid along with two or more appropriate transcription factor expression plasmids as illustriated in **Fig. 2**.

6. To identify individual DNA binding, activation, and inhibitory domains of a given transcription factor, a series of N-terminal, C-terminal, or internal deletion mutants have to be generated and tested in transient transfections (*see* **Note 5**). To avoid incorrect interpretation of the transcriptional abilities of these mutants, the expression level of every single construct is tested by Western blotting and/or EMSA of transiently transfected SL2 cells (*see* **Note 6**).

2. Materials

All solutions should be prepared with tissue-culture-grade water.

1. Tissue culture incubator at 25°C.
2. Microcentrifuge and centrifuge for 50 mL conical sterile plastic tubes.
3. Laminar flow hood.
4. Tissue culture plastic ware including pipets, 25-cm^2 flasks and 60-mm sterile Petri dishes.
5. Sterile Pasteur pipets.
6. SL2 cells: These can be obtained from the American Type Culture Collection (ATCC CRL-1963).
7. Schneider Medium: Schneider Medium (Gibco BRL, Grand Island, NY) containing 10% fetal calf serum tested for growth of insect cells (Gibco BRL), 2 m*M* L-glutamate and antibiotics (100 U/mL penicillin, 100 µg/mL streptomycin) (*see* **Note 7**).
8. 2X HEPES-buffered saline (2X HeBS): 42 m*M* HEPES, pH 7.1, 274 m*M* NaCl, 9.4 m*M* KCl, 2.8 m*M* Na_2HPO_4, 0.2% dextrose. Dissolve 16 g NaCl, 0.7 g KCl, 0.5 g $Na_2HPO_4 \times 2H_2O$, 2 g dextrose, and 10 g HEPES (*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid) in 800 mL water. Adjust the pH to 7.1 with 5 *M* NaOH and then the volume to 1 L with water. Filter-sterilize and store in 50-mL aliquots at 4 or –20°C.
9. 0.25 *M* $CaCl_2$: Dissolve 36.75 g of $CaCl_2 \times 2H_20$ in 1 L water, filter-sterilize and store at 4°C.
10. 0.25 *M* Tris-HCl, pH 7.8: Autoclave.
11. Phosphate-buffered saline (PBS): 137 m*M* NaCl, 2.7 m*M* KCl, 4.3 m*M* Na_2HPO_4, 1.4 m*M* KH_2PO_4. One liter of a 10X concentrated stock solution can be prepared by dissolving 80 g NaCl, 2 g KCl, 7.65 g $Na_2HPO_4 \times 2H_2O$, and 1.9 g KH_2PO_4. The stock solution is autoclaved and stored at room temperature.
12. Dimethyl sulfoxide (DMSO).
13. Plasmid DNA: CsCl-purified (two gradients) supercoiled plasmid DNA or other high quality plasmid DNA.
14. Hypotonic buffer: 10 m*M* HEPES-KOH, pH 7.9 at 4°C, 1.5 m*M* $MgCl_2$, 10 m*M* KCl, 0.5 m*M* DTT (dithiothreitol), 0.2 m*M* PMSF (phenylmethylsulfonyl fluoride). The buffer can be stored at 4°C for several weeks. However, DTT and PMSF have to be added freshly immediately before use. DTT is stored as a 0.5-M stock solution in water, PMSF is dissolved in isopropanol as a 0.2-M stock solution.
15. High salt buffer: 20 m*M* HEPES-KOH, pH 7.9 at 4°C, 1.5 m*M* $MgCl_2$, 420 m*M* NaCl, 0.2 m*M* EDTA, 25% glycerol, 0.5 m*M* DTT, 0.2 m*M* PMSF.

3. Methods

3.1. Propagation and Storage of SL2 Cells

This protocol describes the maintenance and monolayer subculture of Schneider's *Drosophila* line 2 (SL2 cells). SL2 cells are grown in closed tissue culture flasks in the absence of CO_2 at 22–26°C. They grow in a loose monolayer with little tendency to pile up at central foci. They attach quite strongly to the plastic flask during the logarithmic growth phase. Upon confluence, they are only weakly attached. Cultures are maintained by subculturing, and aliquots of these cultures are frozen for long-term storage.

3.1.1. Culture from Frozen Cells

1. Place 10 mL of complete Schneider medium/10% FCS in a sterile 50-mL conical plastic tube.
2. Thaw a frozen aliquot of SL2 cells rapidly in a 25°C water bath by shaking it back and forth by hand. When the sample is almost thawed, rinse the tube with 70% ethanol to sterilize its outside.
3. Transfer content of the tube to the 50-mL conical plastic tube (from **step 1**).
4. Pellet cells by centrifugation at 1000*g* for 10 min at room temperature.
5. Remove medium and resuspend cells in 5 mL of fresh complete medium/10% FCS.
6. Transfer cell suspension to a 25-cm^2 flask. Close flask tightly, rock it gently by hand to distribute the cells evenly, and incubate at 25°C; after 2 to 3 h, cells have attached.
7. The next day, aspirate old medium and replace with 5 mL of fresh medium/10% FCS (*see* **Note 8**). Continue incubating at 25°C by replacing every second day old medium with fresh medium until cells reach confluency.

3.1.2. Maintenance and Subculturing SL2 Cells

1. For subculturing SL2 cells prepare a new flask by adding 4.5 mL complete medium/10% FCS to a 25-cm^2 flask.
2. Aspirate the medium from a confluent culture of SL2 cells (from **step 7**). Add 5 mL of fresh complete medium/10% FCS, close the flask tightly and resuspend the cells by gently knocking the flask with the palm of the hand until most of the cells flush from the bottom of the flask (*see* **Note 9**).
3. Seed 0.5 mL of the suspension into the 25-cm^2 flask from **step 8** and rock it evenly to distribute the cells. Incubate cells at 25°C and replace the medium every second or third day until they reach confluency.

3.1.3. Freezing SL2 Cells for Long-Term Storage

1. Transfer the cell suspension from a confluent flask (*see* **step 9**) into a 50-mL conical plastic tube. Centrifuge cells for 10 min at 1000*g* and discard supernatant.

2. Resuspend cell pellet in 5-mL cold complete medium/10% FCS containing 10% DMSO. Place on wet ice and dispense 1 mL aliquots into freezing vials and incubate overnight in a –80°C freezer.
3. The next day, transfer frozen cells to a –135°C freezer or to liquid nitrogen for long-term storage.

3.2. Transfection of SL2 Cells

SL2 cells are usually transfected by the calcium phosphate method essentially as described by DiNocera and Dawid *(6)* (*see* **Note 10**). Each 60-mm plate receives a constant amount of DNA (up to 20 µg) including the reporter plasmid, the transcription factor expression plasmid, and a control reporter plasmid whose expression is independent of the transcription factor to be analyzed. Variable amounts of expression plasmids are compensated for by empty expression plasmid.

3.2.1. Plating of SL2 Cells

For transfections, SL2 cells are plated on sterile 60-mm cell culture plastic dishes (*see* **Note 14**).

1. Resuspend cells from a confluent flask (*see* **Subheading 3.1.2, step 9**) in 10-mL complete Schneider medium/10% FCS and transfer to a sterile 50-mL conical plastic tube.
2. Count the cells using a hemacytometer.
3. Adjust cell concentration to 1.4×10^6 cells/mL by adding complete medium/10% FCS to the cell suspension. Mix the suspension thoroughly by pipeting up and down.
4. Plate 3 mL (4.2×10^6 cells) of the cell suspension on sterile 60-mm cell culture plastic dishes. Rock gently to equally distribute the cells.
5. Incubate cells at 25°C overnight.

3.2.2. Calcium Phosphate Transfection of SL2 Cells

1. Prepare and purify RNA-free supercoiled plasmid DNA according to standard procedures (*see* **Note 11**). Dissolve DNA in 1/10 TE and determine the concentration photometrically. If necessary, adjust DNA concentration between 0.5 and 2 µg/µL.
2. For a series of cotransfections, add to a sterile microcentrifuge tube 4 µg reporter plasmid, 0.01–2 µg expression plasmid for the transcription factor of choice (*see* **Note 12**), variable amounts of empty expression plasmid (to make up a total of 2 µg with the transcription factor expression plasmid), and 4 µg of an internal standard plasmid (*see* **Note 13**). At this stage, the DNA can be stored at –20°C.
3. Add 0.25 mL 0.25 M $CaCl_2$ solution to the DNA.
4. Add the DNA/$CaCl_2$ solution dropwise to a tube containing an equal volume of 2X HeBS while bubbling air with a Pasteur pipet attached to an electrical pipetor continously into the tube to mix.

5. Let the precipitate stand for 30 min at room temperature.
6. Distribute the precipitate dropwise onto the cells and gently agitate to mix precipitate and medium (*see* **Note 15**).
7. Incubate cells in the absence of CO_2 at 22–26°C.
8. Following 24 h after addition of DNA, remove medium and add 3 mL of fresh complete medium/10% FCS at the border of the plate to avoid dislodging of the cells (*see* **Note 16**).
9. Forty-eight hours after addition of the DNA precipitate, remove medium.
10. Rinse the cells carefully twice with 5 mL PBS.
11. Add 1.5 mL cold PBS and dislodge cells by pipeting up and down or by scraping with a rubber policeman.
12. Transfer cell suspension in PBS into a 1.5-mL microcentrifuge tube. Cells can now be lysed for CAT, luciferase, and ß-galactosidase assays, or nuclear extract preparation.

3.3. Prepariation of Lysates from Transfected SL2 Cells for CAT, Luciferase, and ß-galactosidase Assays

1. Transfer the 1.5 mL of SL2 cell suspension in PBS from **step 11** (*see* **Subheading 3.2.2.**) to a microcentrifuge.
2. Pellet cells by a full speed 10 s centrifugation and discard supernatant.
3. Aspirate PBS and resuspend pellet in 200 µL of 0.25 *M* Tris-HCl, pH 7.8.
4. Freeze cell suspension in liquid nitrogen for 3 min.
5. Thaw cell suspension for 3 min at 37°C.
6. Repeat freeze–thaw cycle twice.
7. Centrifuge lysed cells at top speed in a microfuge for 5 min.
8. Transfer supernatant to a fresh microtube. At this stage, the lysate can be stored at –20°C.
9. Determine protein concentration and perform appropriate enzymatic assays (ß-galactosidase-, CAT-, or Luc-assay) according to standard protocols *(7)*.

3.4. Preparing Nuclear Extracts from Transfected SL2 Cells

This protocol describes the nuclear extract preparation of transfected SL2 cells to analyze the expressed transcription factor(s) for DNA-binding in EMSAs. The protocol is adapted from a procedure described for mammalian cell lines *(8)*.

1. Transfer tubes with 1.5 mL of SL2 cell suspension in PBS from **step 12** (*see* **Subheading 3.2.2.**) to a microcentrifuge (*see* **Note 17**).
2. Pellet cells by a 10-s centrifugation at top speed and 4°C and discard supernatant.
3. Resuspend cells in 400 µL cold hypotonic buffer by flicking the tube.
4. Incubate cells for 10 min on ice to allow swelling.
5. Vortex cell suspension for 5 s.
6. Centrifuge cell suspension at top speed in a microcentrifuge for 10 s at 4°C and discard supernatant leaving the pellet as dry as possible.

7. Resuspend cells in 40 µL cold high salt buffer by pipeting up and down.
8. Incubate tube for 20 min on ice for high-salt extraction of proteins.
9. Centrifuge for 2 min at top speed and 4°C.
10. Transfer supernatant fraction (containing DNA binding proteins) to a cold microcentrifuge tube.
11. Aliquot into four 10-µL fractions for protein determination, electrophoretic mobility studies, and Western blotting. The total protein yield is approximately 20 µg per 10^6 transfected SL2 cells.
12. Freeze and store aliquots at –80°C.
13. Perform EMSA or Western blotting according to standard procedures ***(7)***.

4. Notes

1. In the case of transcription factors with well-defined recognition sequences, fusions of single or multiple copies of the binding site to a strong core promoter region that is active in SL2 cells can be considered for analysis. Widely used core promoter regions, which are active in SL2 cells, are derived from the herpes simplex virus thymidine kinase (HSV-tk) gene ***(3,4)*** and the viral E1b gene ***(9)***, respectively. Alternatively, a TATA box region is obtained from a *Drosophila* gene like the alcohol dehydrogenase gene ***(10)***.
2. Other reporter systems like ß-galactosidase from *Escherichia coli*, human-placental-secreted alkaline phosphatase (SEAP), human growth hormone (hGH), or bacterial ß-glucoronidase (GUS) may also be useful.
3. Owing to the absence of common mammalian transcription factor homologs in SL2 cells, mammalian enhancer/promoter units used for the expression of recombinant genes in mammalian tissue-culture cells like the cytomegalovirus (CMV), Rous sarcoma virus (RSV), or the simian virus 40 (SV40) enhancers/promoters are not or only poorly active in SL2 cells. Thus, the promoter unit for the expression of transcription factors in SL2 cells has to be of *Drosophila* origin. The most widely used promoters of *Drosophila* origin are the distal promoter from the actin 5C gene, the alcohol dehydrogenase promoter, the copia transposable element long-terminal repeat (copia LTR), the heat-shock protein 70 gene promoter, and the inducible metallothionein promoter (*see* **Tables 1** and **2**).
4. If specific antibodies against the transcription factor are not available, it is possible to tag the transcription factor with an epitope at the most N- or C-terminal end for Western-blot analysis. Hemagglutinin, FLAG or *myc* epitopes are commonly used for this purpose because antibodies against these epitopes are commercially available.
5. When C-terminal deletions are generated, it is important to provide a stop codon immediately downstream of the coding sequence. This is achieved by using linkers or adaptors that contain stop codons in all three reading frames. In N-terminal deletion mutants, expression of the mutant protein is achieved, for instance, by using the *Drosophila* alcohol dehydrogenase initiator region that contains as convenient cloning sites a *BamHI* and a *XhoI* recognition sequence immediately

Table 2
Plasmids Useful for Normalizing Transfection Efficiency Variations in SL2 Cells

Vector Enhancer unit	Promoter	Reporter Gene	Reference
copia-CAT	copia LTR	CAT	***(6)***
copia-LacZ	copia LTR	ß-galactosidase	***(29)***
pAdh-lacZ	Alcohol dehydrogenase	ß-galactosidase	***(14)***
pAdh-Luc	Alcohol dehydrogenase	Luciferase	***(30)***
pPac-CAT	Actin 5C	CAT	***(10)***
phsp82lacZ	Heat shock protein 82	ß-galactosidase	***(12)***
pH1-ßgal	Histone H1	ß-galactosidase	***(11)***
HSP70-CAT	Heat shock protein 70	CAT	***(6)***

downstream of the start codon ***(3)***. When using the *Drosophila* actin 5C promoter, a 700 bp untranslated leader sequence plus the first eight codons obtained from the *Drosophila* ultrabithorax cDNA guarantees translation of the truncated protein ***(3)***.

6. Small protein fragments are usually expressed at higher levels compared to the complete protein.
7. Although growth of SL2 cells is very sensitive to the serum batch, there is no absolute need to purchase insect cell tested fetal calf serum (FCS). Many batches of sera that are used for the propagation of mammalian cell lines are also appropriate for the propagation of SL2 cells. However, each individual serum has to be tested. The use of antibiotics in the complete medium is optional.
8. When fresh medium is added to the flask, place it upright and add medium opposite to the cell monolayer to avoid dislodging of the cells.
9. SL2 cells are sensitive to trypsin. Trypsinization, therefore, has to be avoided. For subculturing, simple knocking of the tissue-culture flask as described or alternatively gentle up and down pipeting of the medium is adequate to flush the cells from the bottom of the flask.
10. An alternative protocol using liposome-mediated transfection has also been described ***(11)***.
11. Supercoiled DNA to be transfected should be purified twice by centrifugation to equilibrium in CsCl-ethidium bromide gradients. DNA can also prepared using commercial column procedures. According to our own experience CsCl gradient purified DNA yields better transfection efficiencies.
12. Very high amounts of expression plasmid can lead to squelching. For the first series of transfections, the amount of transfected expression plasmid DNA should be varied by three orders of magnitude. For 60-mm plates, 2, 50, 200, 500, and 2000 ng of plasmid DNA are usually sufficient to determine the linear range of transcriptional activity.

13. Internal standard plasmids are used to normalize plate-to-plate variations in transfection efficiency. An internal standard plasmid for SL2 cell transfections contains a reporter gene driven by a strong *Drosophila* promoter. The activity of the promoter has to be independent of the cotransfected transcription factor and the reporter gene has to be different from the reporter gene whose expression is controlled by the cotransfected transcription factor. Plasmids that have been used to monitor variations in transfection efficiency in SL2 cells are listed in **Table 2**.
14. Smaller dishes or six-well microtiter plates can be used also. In this case, the number of cells to be plated has to be decreased accordingly.
15. The medium becomes very turbid after the addition of the DNA/$CaCl_2$ solution. This is because of the high concentration of phosphate in the Schneider medium leading to additional calcium phosphate precipitates after addition of the DNA/calcium phosphate precipitate, which still contains soluble calcium ions. The appearance of a very turbid calcium phosphate precipitate may be a reason why several investigators have used Shields and Sang M3 insect medium (Sigma, St. Louis, MO) or D-22 insect medium (Sigma) instead of Schneider medium to propagate and transfect SL2 cells. When D-22 or M3 medium is used instead of Schneider medium, the precipitate is only visible in the microscope. Our own investigations suggest that the use of D-22 insect medium for transfection yields slightly higher efficiencies. The usage of Shields and Sang M3 insect medium in transfections resulted in lower efficiencies.
16. Replacement of the medium/calcium precipitate at this stage is optional. SL2 cells are extremely insensitive to long calcium phosphate exposure. Our own investigestions suggest that addition of fresh complete medium/10% FCS increases transfection efficiencies slightly.
17. The amount of cells that have to be used for the preparation of nuclear extracts to obtain satisfying results in band-shift assays is strongly dependent on the transient expression level and the stability of the transcription factor in SL2 cells, as well as on the transfection efficiency. In optimal cases it is sufficient to use only half of the cells from a 60-mm plate for nuclear extract preparation. In this case, the other half could be used in parallel for reporter gene assays. However, it might be necessary to transfect a 60-mm plate or even a 10-cm plate with up to 20 µg of the expression plasmid for the transcription factor to obtain sufficient transcription factor protein for band-shift assays.

Acknowledgments

Dr. M. Kalff-Suske, Dr. J. Klug, J. Dennig, and H. Braun are gratefully acknowledged for critically reading the manuscript.

References

1. Schneider, I. (1972) Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J. Embryol. Exp. Morphol.* **27,** 353–365.

2. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. *Genes Dev.* **2,** 1557–1569.
3. Courey, A. J. and Tjian, R. (1988) Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55,** 887–898.
4. Courey, A. J., Holtzman, D. A., Jackson, S. P., and Tjian, R. (1989) Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell* **59,** 827–836.
5. Mermod, N., O'Neill, E. A., Kelly, T. J., and Tjian, R. (1989) The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell* **58,** 741–753.
6. DiNocera, P. P. and Dawid, I. B. (1983) Transient expression of genes introduced into cultured cells of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **80,** 7095–7098.
7. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds. (1997) *Current Protocols in Molecular Biology*, Wiley, New York.
8. Andrews, N. C. and Faller, D. V. (1991) A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19,** 2499.
9. Pascal, E. and Tjian, R. (1991) Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. *Genes Dev.* **5,** 1646–1656.
10. Yoshinaga, S. K. and Yamamoto, K. R. (1991) Signaling and regulation by a mammalian glucocorticoid receptor in *Drosophila* cells. *Mol. Endocrinol.* **5,** 844–853.
11. Gregory, R. C., Taxman, D. J., Seshasayee, D., Kensinger, M. H., Bieker, J. J., and Wojchowski, D. M. (1996) Functional interaction of GATA1 with erythroid Krüppel-like factor and Sp1 at defined erythroid promoters. *Blood* **87,** 1793–1801.
12. Merika, M. and Orkin, S. H. (1995) Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Mol. Cell. Biol.* **15,** 2437–2447.
13. Galera, P., Musso, M., Ducy, P., and Karsenty, G. (1994) c-Krox, a transcriptional regulator of type I collagen gene expression, is preferentially expressed in skin. *Proc. Natl. Acad. Sci. USA* **91,** 9372–9376.
14. Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dollé, P., Duboule, D., and Charnay, P. (1990) The segment-specific gene *Krox-20* encodes a transcription factor with binding sites in the promoter region of the *Hox-1,4* gene. *EMBO J.* **9,** 1209–1218.
15. Hagen, G., Müller, S., Beato, M., and Suske, G. (1994) Sp1-mediated transcriptional activation is repressed by Sp3. *EMBO J.* **13,** 3843-3851.
16. Hagen, G., Dennig, J., Preiß, A., Beato, M., and Suske, G. (1995) Functional analyses of the transcription factor Sp4 reveal properties distinct from Sp1 and Sp3. *J. Biol. Chem.* **270,** 24,989–24,994.
17. Kobayashi, A., Sogawa, K., and Fujii, K. Y. (1996) Cooperative interaction between AhR.Arnt and Sp1 for the drug-inducible expression of CYP1A1 gene. *J. Biol. Chem.* **271,** 12,310–12,316.

18. Seto, E., Lewis, B., and Shenk, T. (1993) Interaction between transcription factors Sp1 and YY1. *Nature* **365,** 462–464.
19. Xu, J., Thompson, K. L., Shephard, L. B., Hudson, L. G., and Gill, G. N. (1993) T3 receptor suppression of Sp1-dependent transcription from the epidermal growth factor receptor promoter via overlapping DNA-binding sites. *J. Biol. Chem.* **268,** 16,065–16,073.
20. Nehls, M. C., Grapilon, M. L., and Brenner, D. A. (1992) NF-I/Sp1 switch elements regulate collagen alpha 1(I) gene expression. *DNA Cell. Biol.* **11,** 443–452.
21. Fischer, K.-D., Haese, A., and Nowock, J. (1993) Cooperation of GATA-1 and Sp1 can result in synergistic transcriptional activation or interference. *J. Biol. Chem.* **268,** 23,915–23,923.
22. Gegonne, A., Bosselut, R., Bailly, R. A., and Ghysdael, J. (1993) Synergistic activation of the HTLV1 LTR ets-responsive region by transcription factors ets1 and Sp1. *EMBO J.* **12,** 1169–1178.
23. Chen, X., Farmer, G., Zhu, H., Prywes, R., and Prives, C. (1993) Cooperative DNA binding of p53 with TFIID (TBP): a possible mechanism for transcriptional activation. *Genes Dev.* **7,** 1837–1849.
24. Kerr, L. D., Ransone, L. J., Wamsley, P., Schmitt, M. J., Boyer, T. G., Zhou, Q., Berk, A. J., and Verma, I. M. (1993) Association between proto-oncoprotein Rel and TATA-binding protein mediates transcriptional activation by NF-kappa B. *Nature* **365,** 412–419.
25. Yieh, L., Sanchez, H. B., and Osborne, T. F. (1995) Domains of transcription factor Sp1 required for synergistic activation with sterol regulatory element binding protein 1 of low density lipoprotein receptor promoter. *Proc. Natl. Acad. Sci. USA* **92,** 6102–6106.
26. Roulet, E., Armentero, M. T., Krey, G., Corthesy, B., Dreyer, C., Mermod, N., and Wahli, W. (1995) Regulation of the DNA-binding and transcriptional activities of *Xenopus* laevis NFI-X by a novel C-terminal domain. *Mol. Cell. Biol.* **15,** 5552–5562.
27. Lin, S.-Y., Black, A. R., Kostic, D., Pajovic, S., Hoover, C. N., and Azizkhan, J. C. (1996) Cell cycle-regulated association of E2F and Sp1 is related to their functional interaction. *Mol. Cell. Biol.* **16,** 1668–1675.
28. Lee, Y. H., Williams, S. C., Baer, M., Sterneck, E., Gonzalez, F. J., and Johnson, P. F. (1997) The ability of C/EBPß but not C/EBPα to synergize with an Sp1 protein is specified by the leucine zipper and activation domain. *Mol. Cell. Biol.* **17,** 2038–2047.
29. Farmer, G., Friedlander, P., Colgan, J., Manley, J. L., and Prives, C. (1996) Transcriptional repression by p53 involves molecular interactions distinct from those with the TATA box binding protein. *Nucl. Acids Res.* **24,** 4281–4288.
30. Trivedi, A., Vilalta, A., Gopalan, S., and Johnson, D. L. (1996) TATA-binding protein is limiting for both TATA-containing and TATA-lacking RNA polymerase III promoters in *Drosophila* cells. *Mol. Cell. Biol.* **16,** 6909–6916.

16

Triplex-Forming Oligonucleotides and Their Use in the Analysis of Gene Transcription

Marina Kochetkova and M. Frances Shannon

1. Introduction

The ability of oligonucleotides to bind double-stranded DNA in a sequence specific manner, forming a DNA triple helix structure, has led to a number of novel approaches in the study of various structural and functional properties of DNA in vitro and in vivo including transcription inhibition ***(1–3)***, site-directed mutagenesis ***(4)*** and detection of DNA polymorphism ***(5)***.

Conventional approaches to the functional analysis of gene promoters utilizes mutations or deletions in the promoter in question that drives reporter gene activity. This technique cannot be applied to study transcription of endogenous genes in the context of the natural chromatin environment.

The triplex approach, on the other hand, may allow an analysis of transcription regulation of endogenous genes using well-established and noncomplicated experimental techniques. It is also highly specific, as binding of a third strand as long as 17 bases will discriminate between 1×10^5 genes in a human cell. Binding of a triplex-forming oligonucleotide (TFO) in the major groove of the duplex DNA target prevents formation of DNA–protein complexes, thus allowing the evaluation of the significance of the transcription factor binding site in the gene promoter of interest using reporter or endogenous gene assays. Different recognition patterns, or motifs, have been described for the third DNA strand (*see ref.* ***6***). The most commonly used in recent times is R^*R:Y type triplex. This DNA complex involves G^*G:C, A^*A:T and T^*A:T triads with an antiparallel orientation of the third strand and is stable at physiological pH. Also, the orientation of the G and T triplex-forming oligonucleotides is dependent upon the exact sequence of the target site. Another described motif is Y^*R:Y, when homopyrimidine oligonucleotides bind parallel to the duplex

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

homopurine strand via formation of $C^{+*}G$:C and $T^{*}A$:T triplets, but only at acidic pH because it requires protonation of the cytosines. Several modified bases have been substituted for cytosine in order to extend formation of homopyrimidine oligonucleotide-directed triple helices to the physiological pH range. α-oligonucleotides that are resistant to cellular nucleases also have been shown to form DNA triple helices. Predominantly polypurine:polypyrimidine sequences have been shown to be suitable targets for triple helix formation. This requirement represents the main limitation of this approach, because it restricts its application to the promoters that contain relatively long (≈15 bp, *see* **Note 1**) polypurine:polypyrimidine sites which span or overlap the transcription factor binding site ***(6)***. To overcome this restricton some TFOs have been designed to "switch" strands when blocks of homopurine sequences alternate between strands at the duplex DNA target site. Derivatization of TFOs by covalently attached intercalating groups or a cholesterol residue has been shown to increase the stability of the triplex complex.

There are also other constraints such as intracellular and nuclear delivery of oligonucleotides that vary depending on the cell type and oligonucleotide structure, as well as the intracellular stability of the TFO (reviewed in ***[7]***, and also see the notes for the relevant section in this chapter). Currently, a growing number of studies are focused on overcoming these potential problems by introducing modifications to the natural oligonucleotides and applying new intracellular delivery methods, and, recently, TFOs have been shown to inhibit transcription from endogenous IL-2 receptor α ***(8)***, aldehyde dehydrogenase ***(9)***, and GM-CSF ***(10,11)*** genes.

In this chapter, we describe the application of TFOs in the analysis of gene promoter function in vitro and in cells in culture. The first and critical step for the investigator using this approach is to select a target for triplex formation within the promoter of interest and to design specific TFOs, as well as control oligonucleotides. A choice needs to be made between the different motifs available, so we recommend that recent available literature is consulted and the motif most suitable for the particular case chosen. The control oligonucleotide usually represents a scrambled sequence of the specific TFO.

2. Materials

2.1. Cells and Growth Media

1. Jurkat T-cells: A subline of the cell type available from the American Type Culture Collection.
2. RPMI media: RPMI 1640 (Gibco-BRL, Grand Island, NY), supplemented with 10 m*M* HEPES, 0.25% sodium bicarbonate, 100 U/mL penicillin, 100 U/mL streptomycin, and 10% fetal calf serum (FCS) (heat inactivated at 56°C for 30 min).

3. Phorbol ester: Phorbol 12-miristate 13-acetate (PMA) (Sigma, St. Louis, MO) Prepare as a 1 mg/mL stock solution in dimethyl sulfoxide (DMSO) and store at –70°C.
4. Calcium ionophore: A23187 (Boehringer Mannheim, Germany) Prepare as a 10 m*M* stock solution in DMSO and store at –70°C.

2.2. Stock Solutions and Reagents

High-quality deionized Milli-Q water should be used to prepare stock solutions, buffers, and reagents.

1. 1 *M* DTT: Dissolve dithiothreitol powder in Milli-Q water, aliquot and store at –70°C.
2. TE buffer:10 m*M* Tris-HCl, pH 8.0, 0.1 m*M* EDTA.
3. DNase 1 dilution buffer: 10 m*M* Tris-HCl, pH 7.4, 10 m*M* NaCl, 3 m*M* $MgCl_2$, and 1 m*M* $CaCl_2$.
4. 10X T4 polynucleotide kinase buffer: 700 m*M* Tris-HCl, pH 7.6, 100 m*M* $MgCl_2$ and 50 m*M* DTT.
5. 5X Acrylamide load buffer: 500 µL glycerol, 50 µL 100 m*M* EDTA, 100 µL 0.5% xylene cyanole, 100 µL 0.5% bromophenol blue, and water to total volume of 1 mL.
6. 100 ng/µL Poly[d(IC)]: Dissolve Poly[d(IC)]·Poly[d(IC)] (Pharmacia Biotech, Uppsala, Sweden) in water, store at –20°C.
7. Formamide load buffer: 1 mL deionized formamide, 40 µL 100 m*M* EDTA, 100 µL 0.5% xylene cyanole, and 100 µL 0.5% bromophenol blue.
8. 10X TBE buffer: Dissolve 108 g Tris, free base (ultrapure grade) and 55 g boric acid in 800 mL water, add 40 mL 0.5 *M* EDTA and the adjust volume to 1 L.
9. Bradford reagent: Dissolve 10 mg Coomassie blue in 5 mL of 95% ethanol, add 10 mL of 85% phosphoric acid, make up to 100 mL total with water, and filter through Whatman paper size No. 1. Store at 4°C.
10. Phosphate-buffered saline (PBS): Dissolve 0.2 g KCl, 8.0 g NaCl, 0.2 g K_2HPO_4 and 2.9 g $Na_2PO_4{\cdot}12H_2O$ in 800 mL water, adjust the pH to 7.4, and make up to a total volume of 1 L and autoclave.
11. DEPC-treated water: To Milli-Q water add 0.1% (vol/vol) diethylpyrocarbonate, shake well for 1 min, and autoclave.
12. 5X Transcription buffer: 200 m*M* Tris-HCl, pH 7.5, 30 m*M* $MgCl_2$, and 10 m*M* spermidine (HCl) made up in DEPC-treated water. store at –20°C.
13. 1 *M* potassium phosphate buffer: Combine four parts 1 *M* K_2HPO_4 with one part 1 *M* KH_2PO_4, bring to pH 7.8, and autoclave.
14. 100 m*M* ATP: Dissolve ATP (Boehringer Mannheim, Germany) in Milli-Q water and store at –20°C.
15. 1 m*M* D-luciferin: Dissolve D-luciferin (Boehringer Mannheim, Germany) in 5 m*M* potassium phosphate buffer (*see* **Note 2**). Store at –70°C.
16. 10 mg/mL tRNA: Dissolve tRNA from baker's yeast (Boehringer Mannheim, Germany) in water, phenol/chloroform extract, ethanol precipitate, redissolve in DEPC-treated water, and store at –20°C.

17. 10 mg/mL Ribonuclease A (RNase A) solution: Dissolve RNase A (from bovine pancreas, Boehringer Mannheim, Germany) in water and boil for 15 min. Store at –20°C.
18. Luciferase lysis buffer: 100 m*M* potassium phosphate buffer, 2 m*M* DTT, and 10 m*M* EDTA.
19. Luciferase assay buffer: 100 m*M* potassium phosphate buffer, 2 m*M* DTT, 10 m*M* $MgSO_4$, 320 µ*M* coenzyme A, and 500 µ*M* ATP.
20. Guanidinium thiocyanate buffer: 4 *M* guanidinium thiocyanate, 25 m*M* Na citrate, pH 7.0, 0.5% sarcosyl, and 0.1 *M* ß-mercaptoethanol. Add ß-mercaptoethanol last just before use.
21. RNA hybridization buffer: 40 m*M* PIPES, pH 6.7, 0.4 *M* NaCl, 1 m*M* EDTA, 80% formamide. Make 5X solution of the first three components, add 0.1% vol/vol DEPC, autoclave, and mix with formamide in a proportion 1:4.
22. Acrylamide (6%)/ Urea [8 *M*] solution: Dissolve 114 g acrylamide, 6 g bis-acrylamide (both from Bio-Rad, Hercules, CA) and 960 g urea in 800 mL of Milli-Q water and deionize by stiring with 50 g of mixed bed resin (for example, AG 501-X8, Bio-Rad). Filter, add 200 mL 10X TBE, and adjust to 2 L with Milli-Q water. Store at 4°C.
23. Binding buffer: 10% sucrose, 5 m*M* $MgCl_2$, 20 m*M* Tris-HCl, pH 7.5.
24. Footprinting stop solution: 10 m*M* EDTA, 2% SDS, 0.3 *M* NaOAc.
25. Phenol: equilibrated in Milli-Q water.
26. Phenol/chloroform/isoamyl alcohol (25:24:1).
27. 5'-Fluorescently tagged and 3'-amino modified Triplex-forming oligonucleotides (*see* **Note 16**).
28. 0.4 *M* acetic acid, 0.4 *M* NaCl, pH 3.
29. 2 *M* sodium acetate, pH 4.2.
30. 3 *M* sodium acetate, pH 4.6.
31. 2.5 m*M* CTP, ATP, GTP.
32. 100 µ*M* UTP.
33. 1 mg/mL bovine serum albumin (BSA).
34. 1 m*M* DTT.
35. RNasain (Promega, Madison, WI).
36. RQ1 DNase (Promega).
37. 5 *M* ammonium acetate.
38. Digestion buffer: 10 mL Tris-HCl, pH 7.5, 5 m*M* EDTA, 200 m*M* NaCl.
39. 10% SDS.
40. Proteinase K (10 mg/mL).
41. 100% ethanol.
42. 70% ethanol.
43. Isopropanol.
44. RNA polymerase (*see* **Note 25**).
45. Calf intestinal alkaline phosphatase (CIAP, Boehringer Mannheim, Germany).
46. T4 polynucleotide kinase (T4 PNK).
47. Geneclean purification kit (BIO101).

48. [^{32}P]-α-UTP.
49. [^{32}P]-γ-ATP.
50. X-ray film.

2.3. Equipment

1. Microcentrifuge and table top centrifuge.
2. High-voltage vertical electrophoresis system.
3. Gel drier.
4. Liquid scintillation counter.
5. Bio Rad Gene Pulser, electroporater or equivalent, and electroporation cuvettes.
6. Tissue culture incubator with humidified atmosphere 5% CO_2.
7. Vacuum evaporator centrifuge for Eppendorf tubes (Speed Vac).

3. Methods

3.1. DNase 1 Footprinting Analysis

DNase 1 footprinting is performed to estimate the ability of a triplex forming oligonucleotide (TFO) to bind to its target site within the gene promoter and to assess the affinity and specificity of this interaction.

3.1.1. Preparation of One-End Labeled DNA Probe

1. Select two preferably unique restriction sites within the chosen plasmid that will release the DNA fragment containing the element of interest (*see* **Note 3**).
2. Digest 15 µg with 10–20 U of the first restriction enzyme for 1 h at 37°C in a 100 µL of appropriate buffer.
3. Add 5 U of calf intestinal alkaline phosphatase (CIAP, Boehringer Mannheim, Germany) to the same tube and incubate further for 1 h.
4. Extract the mixture with phenol/chloroform/isoamyl alcohol (25:24:1) twice.
5. Transfer aqueous phase into new tube, add 0.1 vol 3 *M* sodium acetate (pH 4.6) and 3 vol of ethanol. Mix well and leave at –20°C for 30 min.
6. Centrifuge at 13,000*g* for 10 min at 4°C to precipitate DNA.
7. Remove the supernatant carefully, add 200 µL of cold 70% ethanol, vortex, and recentrifuge.
8. Remove the supernatant and briefly dry pellet at room temperature (RT).
9. Redissolve DNA pellet in 60 µL of TE buffer (*see* **Note 4**).
10. Check the digest efficiency and DNA recovery by running 1 µL of the solution on a 1% agarose gel.
11. Transfer 20 µL of the solution into new microcentrifuge tube.
12. Add 3 µL of T4 polynucleotide kinase (PNK) buffer, 3 µL of [^{32}P]-γATP, 1 µL (10 U) of T4 PNK, and 3 µL of TE to make up to 30 µL total.
13. Incubate 1 h at 37°C and then heat at 65°C for another 20 min to inactivate the enzyme.
14. Cool to room temperature and digest DNA with second selected restriction endonuclease.

15. Run whole mixture on agarose gel, excise desired band, and purify DNA from gel using Geneclean purification kit following the manufacturer's instructions (*see* **Note 5**).
16. Count Cherenkov radiation of 1 µL of final DNA solution using scintillation ß-counter (*see* **Note 6**).

3.1.2. DNase 1 Footprinting

1. Mix ~20,000 cpm of prepared one-end labeled DNA probe with increasing amounts of TFO (*see* **Note 7**) in 20 µL of binding buffer in microcentrifuge tubes (*see* **Note 8**).
2. Incubate at room temperature for 1 h.
3. To each sample add 0.25 µg poly[d(IC)] in water and DNase 1 to 0.1 U/µL final. Use freshly prepared DNase I stock solution in DNase 1 dilution buffer (*see* **Note 9**).
4. After 30 s at room temperature stop digest by adding equal volumes of footprinting stop solution and 10 µg t-RNA.
5. Extract with 50 µL of phenol.
6. Transfer upper aqueous phase into new tubes, extract again with 50 µL of phenol/chloroform/isoamyl alcohol (25:24:1).
7. Precipitate DNA from aqueous phase with 130 µL of cold ethanol (no salt required).
8. Leave at –20°C for 30 min, followed by centrifugation at 13,000*g* at 4°C for 10 min. Remove the supernatant and wash pellets with 100 µL of cold 70% ethanol twice (*see* **Note 10**).
9. Dry DNA pellets briefly at RT and resuspend in 4 µL of formamide load buffer.
10. Vortex tubes for 5 s, centrifuge quickly to bring sample to the bottom of the tube, heat at 95°C for 3 min, and then place immediately on ice.
11. Run samples on 6% acrylamide/7 *M* urea/1X TBE gel, dry the gel, and expose to X-ray film to visualize DNase 1 digestion pattern.

3.2. Electrophoretic Mobility Shift Assay (EMSA)

Retardation assays with recombinant or nuclear proteins are used to establish the ability of TFOs to compete with the transcription factor for its DNA binding site and the TFO concentration sufficient to inhibit DNA-protein complex formation (*see* **Note 11**).

3.2.1. Preparation of the Radiolabeled DNA Probe

1. Label in separate tubes 20 ng of each strand of chosen promoter fragment and 20 ng of each strand of control DNA fragment (if using one), by adding 2 µL of 10X kinase buffer, 1 µL of [^{32}P]-γATP, water to make final volume 20 µL, and 1 µL (10U) of T4 PNK followed by incubation for 1 h at 37°C.
2. Combine two strands of the same fragment together, heat for 3 min at 95°C and cool down slowly to room temperature.
3. Add 2 µL of 5X acrylamide load buffer, load samples on 10% native acrylamide gel, and run for 1 h at 300 V.

4. Expose gel briefly to X-ray film (1–2 min), excise top (double stranded fragment) bands, and elute DNA from the gel pieces in 200 µL of TE overnight at RT.
5. Centrifuge at 13,000*g* for 3 min and transfer supernatant carefully to a fresh tube, avoiding small acrylamide pieces.
6. Store probes frozen at –20°C until further use (*see* **Note 12**).

3.2.2. Electrophoretic Mobility Shift Assay

1. Mix 1 µL (~0.2 ng) of prepared double stranded DNA probes with different concentrations (*see* **Note 13**) of specific and control TFOs in 10 µL total of binding buffer. Samples containing only probes in the same buffer should also be included.
2. Incubate 1 h at room temperature to allow formation of triplex complex.
3. Add 4 µL of 5X stock solution of the appropriate protein binding buffer (*see* **Note 14**) recombinant or nuclear protein containing the transcription factor of interest and water to a final volume of 20 µL, following by incubation for 30 min at room temperature.
4. Add to each sample 4 µL of 5X acrylamide load buffer and run on a native acrylamide gel.
5. Dry the gel in vacuo at 80°C and autoradiograph to visualize the DNA protein complexes.

3.3. Introduction of Oligos into Cells

Having shown the ability of TFOs to bind to its target site and to block protein–DNA complex formation, it is necessary to check that sufficient concentrations of the TFOs can be delivered inside the cells of the selected type (*see* **Note 15**).

1. Synthesize 5'-fluorescently tagged and 3'-aminomodified (*see* **Note 16**) TFOs (specific and control) at 1 m*M* scale.
2. Prepare 1 m*M* stock solutions of TFOs in water and filter-sterilize using 0.45 µm filter.
3. Place 1×10^5cells per well in 0.5 mL RPMI culture medium supplemented with 10% FCS into 24-well tissue culture plate.
4. Add TFOs, specific and control, to the final concentration at 1 µM under sterile conditions.
5. Incubate for 0 min, 30 min, 1 h, 2 h, and 3 h in the tissue culture incubator at 37°C.
6. Transfer cell suspensions into conical centrifuge tubes and spin at 450*g* for 5 min at 4°C.
7. Remove supernatant and resuspend cell pellets in 2 mL of ice-cold PBS, respin, discard supernatant, and repeat PBS wash once more.
8. Wash cell pellets in 1 mL of 0.4 *M* acetic acid, 0.4 *M* NaCl, pH 3.0 (*see* **Note 17**), respin, and remove supernatant.
9. Resuspend cell pellets again in 0.5 mL of PBS for analysis by FACS or confocal microscopy according to standard procedures.

3.4. Reporter Gene Studies

Reporter gene constructs, driven by the promoter of the gene of interest, are used to evaluate the role of the transcription factor binding site in the promoter activity (*see* **Note 18**).

1. Synthesize 3'-aminomodified TFOs, (specific and control) at 10 m*M* scale.
2. Mix 4×10^6 Jurkat T cells in 300 µL of RPMI culture media supplemented with 20% FCS and 5 µg of the reporter plasmids, including control construct, in electroporation cuvets under sterile conditions by gentle flicking.
3. Stand at room temperature for 10 min.
4. Electroporate cells using a Bio Rad Gene Pulser at 270 V and a capacitance 960 µF.
5. Stand at room temperature for 10 min.
6. Transfer electroporated cells using Paster pipets into tissue culture flasks containing 5 mL of RPMI media supplemented with 10% FCS.
7. Leave cells in a tissue culture incubator for 20–24 h at 37°C.
8. Add TFOs at the desired concentrations (*see* **Note 19**) to the cells transfected with specific reporter construct and control promoter construct. Transfected cell samples untreated with oligonucleotides should also be included as a references for uninduced and stimulated reporter gene activities.
9. Leave cells in a tissue culture incubator for 2 h at 37°C.
10. Stimulate cells by adding to the culture medium PMA to a final concentration of 20 ng/mL and calcium ionophore to a final concentration of 1 µ*M*, leaving untreated cell samples as a reference.
11. Mix flask contents by gentle shaking and leave cells in a tissue culture incubator for 8 h at 37°C.
12. Transfer cells to 10 mL conical centrifuge tubes and spin at 450*g* for 5 min at 4°C.
13. From this point until processed keep cells at 4°C.
14. Remove supernatant by careful aspiration, and resuspend cell pellets in 5 mL ice-cold PBS.
15. Spin cells at 450*g* for 5 min at 4°C, remove supernatant, respin for 1 min more, and remove last traces of the supernatant, carefully avoiding cell pellets.
16. Resuspend cells in 50 µL of cold luciferase lysis buffer and lyse by three cycles of freeze/thaw.
17. Centrifuge lysed cells at 4300*g* for 5 min at 4°C and quickly transfer supernatants to fresh microcentrifuge tubes, avoiding cell debris.
18. Keep cell extracts cold at all times and freeze at –20°C to store longer than1h.
19. Estimate total protein concentrations in the cell extracts using Bradford reagent in 96-well round bottom microtiter plate following standard procedure ***(12)***.
20. Mix the amount corresponding to 10 µg of total protein of cell extracts from each sample with 400 µL of luciferase assay buffer.
21. Bring cell extracts and luciferin to room temperature.
22. Add 60 µL of luciferin solution to samples and measure the light emission of each one immediately after addition in a scintillation ß-counter, including samples

containing 1 ng of purified luciferase in 400 µL of luciferase assay buffer as a positive control or of luciferase assay buffer only as a negative control.

3.5. Endogenous Gene Studies

The ribonuclease (RNase) protection assay, a highly sensitive and specific method for the detection and quantitation of mRNA species, in conjunction with the direct measurement of the protein levels in the cell medium is employed to estimate the effect of the triplex blocking the given transcription factor binding site within the promoter of the endogenous gene (*see* **Note 20**).

3.5.1. Preparation of Total Cellular RNA

1. Place 5×10^6 Jurkat T-cells in 10 mL of RPMI media with 10% FCS per sample into separate flasks.
2. Add appropriate amounts of TFOs, specific and control, from 1 m*M* stock solutions under sterile conditions, leaving one sample untreated as a reference.
3. Incubate at 37°C for 2 h.
4. Activate and process cells, as described in **Subheading 3.4., steps 9–14,** reserving 1.5 mL of cell culture medium for the protein measurements by ELISA (*see* **Note 21**).
5. Resuspend final cell pellet in 800 µL of ice cold guanidinium thiocyanate buffer and mix vigorously (*see* **Note 22**).
6. Combine in a 2 mL Eppendorf tube 800 µL of cell lysates; 200 µL of NaAcetate (2 *M* at pH 4.2); 800 µL of phenol (equilibrated in Milli-Q water), and 200 µL of chloroform/isoamyl alcohol (24:1). Vortex for 30 s and leave on ice for 15 min.
7. Microcentrifuge for 5 min at 4°C.
8. Transfer upper phase avoiding interface into new tubes, add 1 mL of isopropanol, and precipitate RNA on ice for 30 min.
9. Microcentrifuge for 10 min at 4°C, remove supernatant and resuspend RNA pellet with 400 µL of guanidinium thiocyanate buffer.
10. Reprecipitate with 1 mL ethanol on ice for 1h.
11. Microcentrifuge for 10 min at 4°C, remove ethanol supernatant and add 500 µL of 70% cold ethanol, rinse pellet well, and respin for 3 min. Remove all the supernatant carefully with a pipetman tip. Pellet should be clear or translucent (*see* **Note 23**).
12. Resuspend pellet in 200 µL of DEPC-treated water.
13. Read the UV absorbance of 2 µL of this solution at 260 and 280 nm (*see* **Note 24**) and determine the concentration of RNAs using conversion $1A_{260}$=40 µg/mL and store RNA at –70°C until further use.

3.5.2. RNA Probe Synthesis

1. Mix in RNase-free Eppendorf tubes, 2 µL of 5X transcription buffer (usually included by manufacturer of the RNA polymerase), 2 µL of CTP, ATP and GTP pool at 2.5 m*M* each, 0.5 µL of 100 µ*M* UTP, 2 µL of 2000–4000 Ci/mmol $\alpha[^{32}P]$-

UTP, 1 µL of 1 mg/mL BSA, 0.5 µL of 1 m*M* DTT, 0.5 µL of 40U/µL RNasin, 0.5 µL of 5–20 U/µL RNA polymerase (*see* **Note 25**).
2. Add 1.5 µL of cDNA template (*see* **Note 26**) and incubate for 1 h at 37°C.
3. Add 1 µL of RQ1 DNase and incubate for further 15 min.
4. Add 100 µL TE and precipitate probes with 2 µL of 100 µg/µL yeast t-RNA, 67 µL of 5 *M* NH_4 Acetate, and 420 µL ethanol, vortex, and stand at –20°C for 30 min.
5. Microcentrifuge at 4°C for 10 min, wash pellet in 70% ethanol, and dissolve in 5 µL formamide loading buffer.
6. After heating at 95°C for 3 min, load probes on 6% acrylamide/7 *M* urea/1X TBE gel and run gel at 750 V until the xylene cyanol dye reaches approx two-thirds to the bottom of the gel.
7. Expose gel briefly (1–2 min) to X-ray film, excise bands corresponding to the probes, and elute RNA from the gel fragments in 300 µL of 0.5 *M* NH_4 Acetate, 0.1 m*M* EDTA, and 0.1% SDS by shaking for 1 h at RT.
8. Microcentrifuge for 5 min and transfer 200 µL to fresh tube, avoiding acrylamide fragments.
9. Precipitate probes again by adding 2 µL of 10 µg/µL yeast t-RNA, 84 µL of 5 *M* ammonium acetate, and 700 µL ethanol, vortex, and stand at –20°C for 30 min.
10. Spin, dry briefly at room temperature and dissolve in 100 µL of hybridization buffer.
11. Count Cherenkov radiation of 1 µL of final solution using ß-counter (*see* **Note 6**).

3.5.3. Hybridization and RNase Treatment

1. Remove aliquots corresponding to 10 µg (*see* **Note 27**) of the total cellular RNA samples to be analyzed, freeze at –70°C for 15 min and dry completely in a SpeedVac (no heat).
2. Dissolve in 10 µL of hybridiziation buffer, add 1 µL of each specific and control probes, diluted to ~20,000 cpm/µL.
3. Heat reactions to 80°C for 5 min and incubate at 45°C for 16–20 h. It is critical to maintain a constant temperature at this stage.
4. Add RNase A at 10 µg/mL (*see* **Note 28**) in 100 µL of digestion buffer, incubate for 30 min at 30°C.
5. Add 2 µL 10% SDS and 3 µL 10 mg/mL proteinase K, incubate 15 min at 37°C.
6. Add 2 µL of 10 µg/µL t-RNA and extract mixture with 100 µL phenol/chloroform/isoamyl alcohol (25:24:1), vortex, spin briefly and transfer aqueous phase to 250 µL ethanol and leave at –20°C for 15 min.
7. Microcentrifuge for 10 min at 4°C, dry pellet briefly, and resuspend in 4 µL of formamide loading solution.
8. After heating at 95°C for 3 min, load samples on 6% acrylamide/7 *M* urea/1X TBE gel, and run gel in the same manner as for the probes preparation.
9. Dry gel and visualize bands by autoradiography or phosphoimager followed by quantitation of the RNA's levels of interest relative to the concentration of internal RNA reference.

4. Notes

1. Some authors believe that a triple helix forming sequence should contain at least 20 bases in order to bind its target site with sufficient affinity so as to achieve biochemical effects, when no chemical modifications are employed to increase DNA triplex stability.
2. If all of the luciferin powder does not go in the solution after addition of buffer volume needed, add more of the stock solution dropwise until all is dissolved.
3. The optimal DNA fragment size is between 100 and 500 bp, with the triplex target site positioned approx 50–150 bp from the labeled end. A restriction enzyme generating a 5' overhanging end is preferable for the first cut site, because this provides the most labeling options.
4. This CIP-treated unlabeled probe can be stored at –20°C for future use.
5. Glass milk DNA purification kits can be obtained from a number of companies. We used Bresa-Clean (Bresatec, Adelaide, Australia).
6. Probe should be labeled to a high-specific activity of 10,000–20,000 cpm per 1–10 pmol of DNA. Prepared probe can be stored at –20°C for no longer than two weeks. ^{33}P radioisotope can be used instead of ^{32}P. Although it has a half-life that is twice as long as ^{32}P, it gives a weaker signal.
7. It is necessary to titrate the TFO against the probe to estimate the *KD* of the binding and to ensure the authenticity of the footprint by being dose dependent.
8. The highest concentration of control oligonucleotide and a reference sample containing probe alone in the same buffer should also be included.
9. The concentration of DNase 1 required to obtain a good footprint must be determined empirically by digestion of the probe alone with the increasing concentrations of the enzyme usually in the range from 0.1 to 10 U/μL.
10. Salts should be thoroughly removed from DNA pellets at this stage, otherwise bands in the ladder can become compressed.
11. A promoter fragment, containing the transcription factor binding site and triplex target sequence with at least 5 bp of flanking sequences should be used as a DNA probe. It is advisable to include in the assay as a control for the specificity of the triplex inhibitory effect, a DNA fragment with a consensus binding site for the protein in question but without the triplex target sequence
12. Prepared probe can be stored at –20°C for no longer than 2 wk.
13. The range of specific TFO concentrations should include relative *KD*, estimated in the footprinting assay, with two or three steps as a factor of 2, going above and below this value. The control TFO may be used only at the highest concentration of specific TFO.
14. Optimal DNA binding conditions, such as ions and salt concentrations, should be used for each transcription factor at this step. The percentage of the acrylamide gel used for EMSA depends on the size of the protein being analyzed.
15. Although the method suggested here is not quantitative, the idea is to establish the fact that oligos are being taken up by the cells and localize in the nuclei. The quantitative methods have also been described, using radiolabeled oligonucle-

otides (***13*** and references therein). It has been shown previously that if oligonucleotides are not trapped in the cytoplasm, they concentrate mainly in the nucleus of the cell, quickly equilibrating the concentration within and outside the cell membrane ***(13)***. Jurkat T-cells are given as an example in this and subsequent sections.

16. It is necessary to use modified oligonucleotides for the experiments in living cells to prevent them from degradation by nucleases. The 3'-amino derivatized analogs have been shown to be significantly more stable than unmodified oligonucleotides in experiments in cells. Other modifications that increase the oligo stability, such as conformationally restricted sugar analogs, oligomers bearing achiral internucleoside linkages, N3'-P5' phosphoramidate oligonucleotides have been described (reviewed in ***14***).
17. Low pH washing has been reported to remove almost 95% of cell surface bound ligands.
18. Luciferase reporter constructs containing inducible gene promoters are given here as an example, but other reporter systems, for example, chloramphenicol acetyltransferase (CAT), can also be used. Unrelated gene promoter reporter construct that functions in selected cell type, but does not contain the TFO binding site, should be included in the study as a negative control for the inhibitory effect of triplex formation.
19. Because of the slow kinetics of triplex formation, oligonucleotide concentrations well above the value of the dissociation constant may be required to drive the complex formation within the cells.
20. A cDNA fragment of the gene of interest can be subcloned into a plasmid that contains bacteriophage promoters and that construct is used as a template for synthesis of radiolabeled antisense RNA probes. It is necessary to include in the assay one or two other genes, whose promoters are affected by the transcription factor in question, but do not contain target site for the TFO. One of the housekeeping genes, GAPDH, for example, should also be used as an internal reference to normalize RNA levels. Successful RNase protection assays are dependent to a very high degree on all solutions and materials being RNase free.
21. ELISA is a very sensitive and convenient method for protein concentration analysis and ELISA kits for the majority of cytokines are commercially available. Other methods, for example, radioimmunoassay, can also be used. Keep cell culture media samples frozen at –20°C until assayed. Other intracellular proteins may need different forms of measurement.
22. Cell lysates may be kept at –20°C until RNA can be processed.
23. A white pellet may indicate salts, especially guanidinium precipitating with RNA.
24. A low (less than 2) 260/280 ratio indicates the presence of guanidinium.
25. All components except for the radiolabeled nucleotide may be purchased together in the in vitro transcription kit from Promega Corporation (Madison,WI).
26. Use DNA miniprep from 1.5 mL bacterial culture, phenol/chloroform extracted, ethanol precipitated and dissolved in 40 µL TE.
27. 10–20 µg is typically used, but as little as 2 µg can be assayed.

28. The optimal RNase concentration needs to be determined for the amount of RNA per assay depends on the availability of RNA and level of the signal for each probe. Initial digestion is usually performed for each new probe using 2.5, 10, and 40 µg/mL RNase A.

References

1. Helene, C., Thuong, N. T., and Harrel-Bellan, A. (1992) Control of gene expression by triple helix-forming oligonucleotides. The antigene strategy. *Ann. N.Y. Acad. Sci.* **660,** 27–36.
2. Ing, N. H., Beekman, J. M., Kessle,r D. J., Murphy, M., Jayaraman, K., Zendegui, J. G., Hogan, M. E., O'Malley, B. W., and Tsai, M.-J. (1993) *In vivo* transcription of a progesterone-responsive gene is inhibited by a triplex forming oligonucleotide. *Nucl. Acid Res.* **21,** 2789–2796.
3. Cooney, M., Chernuszewicz, C., Postel, E. H., Flint, S. J., and Hogan, M. E. (1988) Site-specific oligonucleotide binding represses transcription of the human *c-myc* gene *in vitro. Science* **241,** 456–459.
4. Wang, G., Levy, D. D., Seidman, M., and Glazer, P. M. (1995) Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation. *Mol. Cell. Biol.* **15,** 1759–1768.
5. Nishikawa, N., Oishi, M., and Kiyama, R. (1996) Consruction of a human genomic library of clones containing poly (dG-dA)**.**poly (dT-dC) tracts by Mg^{2+}-dependent triplex affinity capture. *J. Biol. Chem.* **270,** 9258–9264.
6. Stull, R. A. and Szoka, F. C. (19950 Antigene, ribozyme and aptamer nucleic acid drugs: progress and prospects. *Pharm. Res.* **12,** 465–483.
7. Maher, lll (1996) Prospects for the therapeutic use of antigene oligonucleotides. *Cancer Investigation*. **14,** 66–82.
8. Grigoriev, M., Praseuth, D., Guieysse, A. L., Robin, P., Thuong, N. T., Helene, C., and Harrel-Bellan, A. (1993) Inhibition of gene expression by triple helix-directed DNA cross-linking at specific sites. *Proc. Natl. Acad. Sci. USA* **90,** 3501–3505.
9. Tu, G.-C., Cao, Q.-N., and Israel, Y. (1996) Inhibition of gene expression by triple helix formation in hepatoma cells. *J. Biol. Chem.* **270,** 28,402–28,407.
10. Kochetkova, M. and M. F. Shannon (1996) DNA triplex formation selectively inhibits granulocyte-macrophage colony-stimulating factor gene expression in human T cells. *J. Biol. Chem.* **271,** 14,438–14,444.
11. Kochetkova, M., Iversen, P. O., Lopez, A. F., and Shannon, M. F. (1997) DNA triplex formation inhibits GM-CSF gene expression and supresses growth in Juvenile myelomonocytic leukemic cells. *J. Clin. Invest.* **99,** 3000–3008.
12. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory Press, New York, pp 10.01–10.10.
13. Akhtar, S. and Juliano, R. L. (1992) Cellular uptake and intracellular fate of antisense oligonucleotides. *Trends Cell. Biol.* **2,** 139–144.
14. Goodchild, J. (1990) Conjugates of oligonucleotides and modified oligonucleotides: A review of their synthesis and properties. *Bioconj. Chem.* **1,** 165–187.

17

Expression and Purification of Histidine-Tagged Transcription Factors

Donna R. Cohen

1. Introduction

1.1 Background

The analysis of transcription factors in terms of their structure, function, and mode of interaction with DNA and other components of the transcriptional machinery has been greatly facilitated in recent years by the identification of genes encoding these proteins. Not only has this allowed elucidation of primary sequence and structure, and identification of transcription factor subfamilies and functional domains, but it has also permitted synthesis of large quantities of recombinant proteins for use in DNA binding, protein–protein interaction, and in vitro transcription assays.

Although classical purification protocols may be used to recover recombinant proteins following expression, recombinant DNA technology also permits the engineering of fusion proteins bearing specific affinity tags. These tags greatly simplify the purification of the recombinant protein, which usually involves a single chromatography step with an appropriate affinity resin.

Metal chelate affinity chromatography has been known to be a useful technique for protein purification since 1975 ***(1)***, but over the last 10 yr the development of improved chelating agents for immobilizing metal ions on solid supports ***(2)*** has resulted in wide-spread application of this technology. The method is based on the interaction between histidine residues and electropositive transition metals, such as Ni^{2+}, Co^{2+}, and Zn^{2+} (reviewed in ***[3]***). Six consecutive histidine residues will bind very tightly to these metals under conditions of physiological pH, even in the presence of strong denaturing agents. Since very few naturally occurring proteins contain multiple neighbor-

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

ing histidine residues, a single purification step removes most contaminants from recombinant proteins bearing a six-histidine (6xHis) tag.

Histidine-tagged proteins can be purified following synthesis in prokaryotic or eukaryotic systems. Since the metal–histidine interaction is conformation-independent, the purification can be carried out under native or denaturing conditions. The tag itself is small, uncharged at physiological pH, and poorly immunogenic, so it does not interfere with structure, function, or secretion of the recombinant protein and does not have to be removed by cleavage.

This chapter details methods for purifying histidine-tagged transcription factors that have been synthesized in *Escherichia coli*. A cloning strategy is presented that can be adapted for a variety of vectors and target proteins, and screening methods for identifying expressing clones in bacterial systems are described. The suggested protocols for optimization of conditions can be adapted for use with any expression system, and the purification methods are universally applicable.

1.2. Transcription Factors Successfully Purified by this Method

Table 1 lists some examples of transcription factors that have been purified by metal affinity chromatography.

1.3. Preliminary Considerations

1.3.1. Choice of Expression System

Recombinant proteins synthesized in prokaryotic or eukaryotic systems may be purified using metal chelate chromatography. Both systems have advantages and disadvantages.

Synthesis of histidine-tagged proteins in yeast, insect, plant, or mammalian cells has the advantage that the resulting protein is likely to be correctly modified (for example, by phosphorylation) or processed. However, the establishment and use of these systems can be costly and time-consuming.

Bacterial expression systems have the advantage of being easily manipulated and readily available. Any laboratory able to perform the basic molecular biology required to generate an expression construct will be set up to grow bacterial cultures for protein expression. Major disadvantages, apart from lack of protein processing or modification, include poor expression of the protein due to toxicity or instability and potential problems with codon usage ***(18)***.

In practice, correct posttranslational modification of the recombinant proteins does not appear to be a major problem with transcription factors. Indeed, unmodified proteins recovered from bacteria are useful substrates for experiments designed to elucidate posttranslational regulation of transcription factor function.

Table 1
Transcription Factors Purified by Metal Chelate Chromatography

Protein	Expression system	Reference
c-Jun	Mammalian cells	Papavassiliou et al., 1992 ***(4)***
Serum response factor (SRF)	Mammalian cells	Jenknecht et al., 1991 ***(5)***
RXRa	Mammalian cells	Bugge et al., 1992 ***(6)***
C2 protein of tomato yellow leaf curl geminivirus	*E. coli*	Noris et al., 1996 ***(7)***
c-Fos (full length)	*E. coli*	Abate et al., 1990 ***(8)***
wbFos (truncated protein)	*E. coli*	Abate et al., 1990 ***(9)***
HNF3a	*E. coli*	Zaret and Stevens, 1995 ***(10)***
c-Jun (full length)	*E. coli*	Abate et al., 1990 ***(8)***
wbJun (truncated protein)	*E. coli*	Abate et al., 1990 ***(9)***
RNA polymerase *(E. coli)*	*E. coli*	Kashlev et al., 1993 ***(11)***
SRY (testis determining factor)	*E. coli*	Cohen et al., 1994 ***(12)***
TATA binding protein (TBP)	*E. coli*	Kato et al., 1994 ***(13)***
TFIID	*E. coli*	Hoffmann and Roeder, 1991 ***(14)***
WT1 (Wilms tumor protein)	*E. coli*	Rauscher III et al., 1990 ***(15)***
c-Fos	Baculovirus	Corvello et al., 1995 ***(16)***
histone H2B (yeast)	Yeast	Lorch and Kornberg, 1994 ***(17)***

1.3.2. Preparation of an Expression Construct

The generation of a construct for expressing a histidine-tagged protein usually involves ligation of a fragment encoding the selected protein into an appropriate plasmid vector that will permit expression in the cells of choice. There are a number of considerations that will influence the choice of vector, the positioning and method of introduction of the histidine tag, and the mechanism of generation of the insert fragment.

Protein Size

If the protein of interest is very small, it may be desirable to generate a construct in which the protein is expressed as a fusion with a larger protein, to enhance stability during synthesis. Protein size is unlikely to be a problem in the case of most full-length transcription factors; proteins as small as mol wt 11,000 have been successfully produced as histidine-tagged proteins without the assistance of any protein fusion (*see* **Fig. 6**).

Codon Usage

Transcription factors often have domains rich in basic amino acids, and these can pose a problem in the synthesis of full-length proteins in bacteria, since

two of the six codons specifying arginine residues (namely AGG and AGA) are the least frequently used codons in *E. coli*. One solution to this problem is to position the histidine tag at the C-terminus of the protein to ensure that only full-length, correctly translated products are purified. A second solution is to reengineer highly basic domains of the protein to replace potentially troublesome codons with more frequently used *E. coli* codons. Such an approach was used to generate the bZip domain of the Fos transcription factor (wbFos; ***[9]***). Codon usage is not generally a problem if proteins are synthesized in eukaryotic cells.

Functionally Important Domains

In the majority of cases, the 6xHis tag does not interfere in the structure or function of the tagged protein. There may be instances, however, where a functionally important domain is located at one or the other terminus of the protein, and this may influence the positioning of the 6xHis tag. In some cases, it may be desirable to include a cleavage site between the protein and the tag to facilitate its removal.

Vectors

A number of commercial companies have vectors available for the generation of 6xHis-tagged proteins, including Qiagen (Chatsworth, CA), Invitrogen (San Diego, CA), and Clontech (Palo Alto, CA).

The Qiagen pQE vectors are designed for the expression of proteins in *E. coli*. They belong to the pDS family of plasmids ***(19)*** and have several common features (such as an IPTG-inducible promoter) as well as features unique to particular subclasses of vectors (6xHis tag at N- or C-terminus; mouse DHFR sequence to create fusion proteins; various multiple cloning site enzyme combinations; ATG codon supplied by insert or by vector). All classes of vector have versions for cloning in each of the three reading frames. The bacterial strain used for expression of proteins from the pQE vectors is *E. coli* M15, which harbors a plasmid (variously referred to as pREP4 or pDMI.1) that encodes the *lac* repressor ***(20)***.

Invitrogen also has an extensive selection of vectors for expression of 6xHis-tagged proteins, and their range includes vectors for use in bacterial, insect, or mammalian cells. All vectors have versions for cloning in each of the three reading frames. The prokaryotic expression vectors have a choice of either high-level constitutive expression or IPTG-inducible expression. The tag may be N- or C-terminal, and cleavage sites and other epitopes for detection of the expressed protein are variously included. The vectors for use in mammalian cells make use of the high-level expression driven by the CMV enhancer–promoter (or the RSV LTR promoter in one case), offer N- or C-terminal tag-

ging options, include other epitopes to assist in detection, and in one vector allow for secretion of the synthesized protein.

Clontech has only two vectors specifically designed for the generation of 6xHis-tagged proteins. One of these is for use in baculovirus expression systems, and the second plasmid, p6xHis-GFP, is intended as a control vector to allow easy monitoring of the purification steps.

As an alternative to this range of commercially available vectors, the 6xHis tag can be easily incorporated into any expression vector by PCR or insertion of a linker fragment.

The vector in the constructs used for figures in this chapter, pDS56/RBSII/SphI, was obtained from the investigators who originally developed the system (*see **[19]***; kind gift of R. Gentz of Hoffmann-La Roche, Basel). Although it is related to the vectors available from Qiagen, this plasmid does not contain the 6xHis tag sequence, which is added by PCR to each of the inserts before ligation into the vector. This can be more convenient than attempting to match reading frames of insert and vector. Because it is often necessary to add appropriate restriction sites to cDNA inserts to facilitate cloning, the one PCR step is used to add restriction sites and an in-frame 6xHis tag. Obviously, any suitable expression vector can be used if the cloning strategy involves incorporation of the 6xHis tag into the insert fragment.

Inserts

In the majority of circumstances, an expression construct will be made using the coding region of a cDNA. For the purposes of this discussion, it is assumed that the cDNA of interest is available in another plasmid and has been thoroughly characterized in terms of sequence and open reading frame. Broadly speaking, insert fragments may be either resected from one plasmid by various combinations of restriction enzymes and then recloned into the expression vector of choice, or amplified by PCR and then ligated into the vector.

The advantage of resecting the insert fragment by restriction digestion is that there is no risk of PCR-introduced errors disrupting the coding region. However, this strategy has some drawbacks, because it is necessary to identify convenient restriction sites to excise the cDNA from an existing construct that allow recloning of the insert in-frame with the 6xHis tag to generate the expression construct.

If a resection/recloning strategy is not feasible, it may be simpler to employ a PCR strategy, using primers containing suitable restriction sites to amplify the cDNA insert. Care should be taken when using PCR to generate fragments for cloning to use DNA polymerases that have some proofreading activity. In addition, it is advisable to sequence the PCR-generated inserts once they have been cloned into the expression vector, so as to ensure that the protein being

expressed is not mutated as a result of a PCR-introduced error. Clearly, if it is possible to avoid the use of PCR at any stage in the generation of the insert fragment, this will reduce the workload in characterization and verification of the final expression construct.

1.3.3. Purification

Purification of histidine-tagged proteins can be carried out under native or denaturing conditions. The use of denaturing conditions is far more straightforward, and one regimen can be applied to many very different transcription factors. The proteins have to be renatured after purification, but dialysis to remove the denaturant appears to allow sufficient renaturation to recover DNA-binding functions. If it is desirable or necessary to use native conditions, then the solubility, subcellular localization, and accessibility of the tag will have to be determined for each individual protein.

Immobilized metal affinity chromatography (IMAC) resins are currently available from four commercial sources.

Qiagen supplies the Ni-NTA Sepharose CL-4B resin originally developed at Hoffmann-La Roche *(2)*. The protocols given in this chapter make use of the Qiagen resin and may have to be modified for use with products from other companies.

Clontech supplies the TALON resin, which is described as a nonnickel affinity resin, although precisely which metal is used is not disclosed in their literature. This resin is purported to have superior purification properties for 6xHis-tagged proteins compared to nickel resins, but in a side-by-side comparison (*see* **Fig. 3**) no difference was found between Ni-NTA and TALON in the purification of the c-Jun protein. In large-scale purification of the SRY protein (*see* **Fig. 5**), Ni-NTA resin gave superior results, in terms of both quantity and function of the recovered protein. Thus, whereas Ni-NTA has proved suitable for purification of a range of transcription factors, trials with the TALON resin were not consistently successful.

Both Ni-NTA and TALON resins are also available in spin-column format, but trials with the Ni-NTA spin columns failed to identify any advantages with their use compared to the protein miniprep method (**Subheading 3.2.2.**); the spin column method is no faster and the protein recovered is much less concentrated. In a direct comparison of the two methods (*see* **Fig. 3**), there was relatively poor recovery of protein from the spin columns.

There are two other commercial sources of matrices. Pharmacia (Uppsala, Sweden) supplies a number of resins in their HiTrap and HisTrap range, which make use of Cu^{2+}, Zn^{2+}, or Ni^{2+} ions for purifying 6xHis-tagged proteins. Invitrogen supplies the ProBond Metal-Binding Resin, which is a nickel-charged agarose resin.

2. Materials

Use high quality deionised water to prepare all aqueous solutions.

1. Spin columns (Pharmacia Biotech).
2. T4 DNA ligase.
3. 10 m*M* dNTP (Pharmacia Biotech).
4. 10 m*M* ATP (Pharmacia Biotech).
5. *Pwo* polymerase and reaction buffer (Boehringer-Mannheim, Darmstadt, Germany).
6. Reagents for SDS-polyacrylamide gels (preweighed acrylamide:bis-acrylamide, TEMED, APS, Kaleidescope Prestained Standards, Coomassie Brilliant Blue dye) are from BioRad (Hercules, CA).
7. Ni-NTA resin (Qiagen).
8. PCI: phenol:chloroform:isoamyl alcohol (25:24:1).
9. CI: chloroform:isoamyl alcohol (24:1).
10. TE buffer: 10 m*M* Tris-HCl pH 8.0, 1 m*M* EDTA pH 8.0.
11. LB medium: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl; add 15 g/L agar for plates.
12. 10X Ligation buffer: 300 m*M* Tris-HCl pH 7.8, 100 m*M* $MgCl_2$, 100 m*M* DTT.
13. 25 mg/ml Kanamycin: prepare a stock solution in water; sterilize by filtration and store at –20°C. The working concentration is 25 µg/mL.
14. 50 mg/mL ampicillin: prepare in water; sterilize by filtration and store at –20°C; the working concentration is 50 µg/mL.
15. Dissociation buffer: 1x = 10% glycerol, 50 m*M* Tris-HCl pH 6.8, 5% ß-mercaptoethanol, 0.1% bromophenol blue.
16. Acrylamide stock solution: 30% acrylamide:bis-acrylamide 29:1.
17. SDS-polyacrylamide gels: see Sambrook et al ***(21)*** for composition of gels and buffers.
18. Buffers for purification of 6xHis-tagged proteins under denaturing conditions:
    Buffer A = 6 *M* Gu-HCl, 0.1 *M* NaH_2PO_4, 10 m*M* Tris-HCl. Adjust to pH 8.0.
    Buffer B = 8 *M* Urea, 0.1 *M* NaH_2PO_4, 10 m*M* Tris-HCl. Adjust to pH 8.0.
    Buffer C = 6 *M* Gu-HCl, 0.1 *M* NaH_2PO_4, 10 m*M* Tris-HCl. Adjust to pH 6.0.
    Buffer C^U = Buffer C, using 8 *M* Urea instead of 6 *M* Gu-HCl.
    Buffer D = 6 *M* Gu-HCl, 0.1 *M* NaH_2PO_4, 10 m*M* Tris-HCl. Adjust to pH 5.0.
    Buffer D^U = Buffer D, using 8 *M* Urea instead of 6 *M* Gu-HCl.
19. Coomassie Brilliant Blue dye solution: 0.125% Coomassie Brilliant Blue in 40% methanol,7% acetic acid.
20 Destaining solution: 30% methanol, 10% acetic acid.
21. Horizontal agarose minigel apparatus: Mini-Sub Cell GT from BioRad or similar.
22. Vertical mini-gel apparatus for SDS-polyacrylamide gels: Mini Protean II apparatus (BioRad) or similar equipment.
23. Empty columns: Poly-Prep Chromatography Columns (BioRad)
24. Incubators: an incubator at 37°C suitable for growing bacterial plates and liquid cultures, and water baths for incubations at 37°C, 42°C, and 100°C.
25. Thermal cycler.

3. Methods

3.1. Generation of Expression Construct

3.1.1. Preparation of Vector

1. Digest 20 µg of plasmid vector DNA in 150–200 µL with appropriate restriction enzyme(s) (*see* **Note 1**).
2. Extract the digestion reaction with an equal volume of PCI. Vortex for 1 min, spin for 2 min in a microcentrifuge, and transfer the aqueous layer to a fresh tube. Add 100 µL of TE buffer to the organic phase to "back-extract" (*see* **Note 2**). Combine the aqueous phases.
3. Repeat the extraction of the combined aqueous phases with an equal volume (now approx 300 µL) of PCI. It is not necessary to back-extract on this occasion.
4. Extract the aqueous phase with an equal volume of CI; vortex 1 min, centrifuge 2 min and transfer the aqueous phase (approx 300 µL) to a fresh tube.
5. Precipitate the DNA with 150 µL 7.5 *M* ammonium acetate and 900 µL ethanol at –70°C for at least 30 min (*see* **Note 3**).
6. Recover the DNA by centrifugation (15–30 min), wash the pellet, and dry under vacuum.
7. Resuspend the DNA in 50–100 µL TE buffer (*see* **Note 4**).

3.1.2. Preparation of the insert fragment (*see* **Note 5**)

This method employs a PCR strategy for introduction of 6xHis codons and restriction sites into the insert fragment, and has been successfully used to amplify tagged transcription factor-encoding cDNAs ranging in size from 300 to 1100 bp. **Fig. 1** shows an example of the primers that can be used to add histidine residues and restriction sites to any cDNA insert. General information on the PCR technique can be found elsewhere (*see **[21–23]***). Use of aerosol resistant tips for all pipetting steps is recommended.

1. Dilute both primers to 20 pmol/µL; dilute plasmid template to 1–5 ng/µL; dilute *Pwo* polymerase 1:4 in 1x *Pwo* reaction buffer to give 1.25 units/µL (*see* **Note 6**).
2. To a sterile 0.5 mL tube, add 10 µL 10X reaction buffer (supplied with enzyme); 2 µL each 10 m*M* dNTP stock solution; 1 µL diluted plasmid template DNA; 5 µL each diluted primer; 70 µL ddH_2O, and 1 µL diluted enzyme (final volume 100 µL). Overlay with 50 µL mineral oil.
3. Amplification conditions (*see* **Notes 7** and **8**): first cycle: 94°C, 5 min; 37°C, 30 s; 55°C, 1 min. Next 25 cycles: 94°C, 30 s; 55°C, 30 s; 72°C, 1 min. Final extension: 72°C, 10 min.
4. Remove the mineral oil by adding 100 µL $CHCl_3$, vortex 1 min, centrifuge 2 min, and transfer aqueous (upper) phase to a 1.5 mL microcentrifuge tube.
5. Extract the DNA as described above for vector preparation (*see* **steps 2–4**), except use only 50 µL TE buffer for the back-extraction step.
6. Remove unincorporated dNTPs and excess primers by spin column chromatography (*see* **Note 9**).

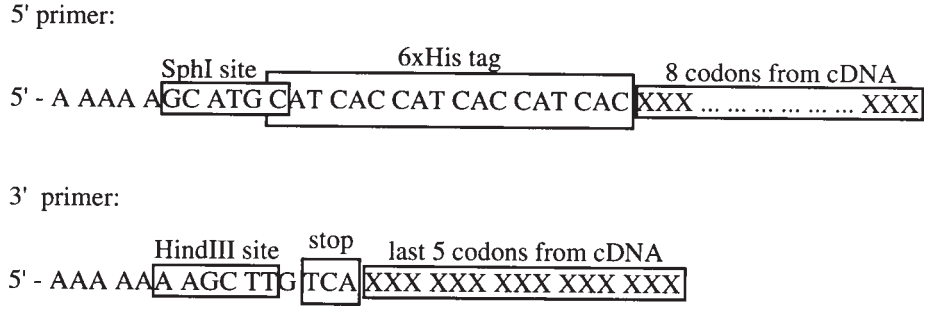


Fig. 1. Primers for PCR. The 5' primer contains an SphI site (which provides an ATG codon for the start of translation) followed by 6 histidine codons and then 24 nucleotides from the cDNA template (corresponding to the first 8 codons of the protein to be included in the tagged construct). The 3' primer incorporates the last 5 codons plus the termination codon of the cDNA, followed by a restriction site (e.g., HindIII).

7. Digest the purified PCR fragment with appropriate restriction enzymes (*see* **Note 10**).
8. Extract and precipitate the DNA as described for vector preparation (*see* **steps 2–6**).
9. Resuspend the DNA pellet in 20–30 µL TE buffer. Check the recovery of the insert fragment by agarose gel electrophoresis.

3.1.3. Ligation and Transformation (*see* **Note 11**)

1. To a sterile microcentrifuge tube, add 1 µL 10x ligation buffer, 1 µL 10 m*M* ATP, 150–200 ng vector DNA plus appropriate amount of insert DNA (combined vector + insert DNA volume not to exceed 7 µL), ddH_2O to bring the volume to 9 µL, and 1 µL T4 DNA ligase (5–7 Weiss units).
2. Incubate the ligation reaction at 14°C overnight or at room temperature (22°C) for 2–3 h.
3. Spin tube briefly to recover condensed liquid. Incubate on ice for 5–10 min to chill the tube.
4. Add 100 µL of competent M15 cells (*see* **Note 12**). Incubate on ice for 40 min.
5. Heat shock the transformation mixture at 42°C for 2 min.
6. Return the tube to ice for 5 min.
7. Add 1 mL LB medium and incubate at 37°C for 15–30 min.
8. Spin for 1 min at room temperature to pellet the cells. Remove the supernatant and resuspend the pellet in 110 µL LB medium. Divide between two LB agar plates containing kanamycin + ampicillin: spread 100 µL of the resuspended cells onto one plate and the remaining 10 µL (in a 100 µL drop of LB medium) onto the second plate. Incubate overnight at 37°C.

3.2. Screening for an Expressing Clone

Most commercial vectors for generating 6xHis-tagged proteins do not permit blue-white selection or possess any other convenient mechanism for dis-

tinguishing recombinant from parental clones. This will often not present a problem because the vector is double-digested. A few colonies, however, are usually obtained on the vector-only control, indicating that some vector is able to religate in the absence of insert. Therefore, it is necessary to screen several colonies to identify those bearing recombinant plasmids that express the tagged protein of interest. Either the crude lysis method (*see* **Subheading 3.2.1.**) or the protein miniprep method (*see* **Subheading 3.2.2.**) may be used for this purpose, although the former is faster, is less labor-intensive, and is preferred when large numbers of colonies are to be screened. The protein miniprep method can be used to confirm the identity of one or two expressing colonies identified by the crude lysis method, and to confirm that the protein can bind to the resin under denaturing conditions.

3.2.1. Crude Lysis of Small-Scale Cultures

1. Pick 17 colonies from the vector-insert ligation and one colony from the vector-only ligation with sterile toothpicks. Spot the bacterial cells to a labeled "master" plate (LB agar containing kanamycin + ampicillin) and then inoculate 3 mL LB (kan + amp) medium. Incubate with vigourous shaking overnight at 37°C.
2. The following day, set aside 1.5 mL of each overnight culture to use in miniscale preparation of plasmid DNA. Use 0.5 mL of overnight culture to inoculate 2.5 mL fresh, prewarmed LB (kan + amp) medium. Incubate with shaking at 37°C for 45 min (*see* **Note 13**).
3. Add 30 µL 100 m*M* IPTG solution to each culture to induce protein synthesis. Continue shaking at 37°C for a further 2 h.
4. Transfer 1.5 mL of each culture to a microcentrifuge tube. Pellet the bacterial cells by centrifugation at top speed for 1 min.
5. Resuspend the cell pellet in 100 µL 1x dissociation buffer.
6. Boil the samples for 5 min and then pellet the cell debris for 5 min.
7. Analyze 20 µL of each sample by SDS-polyacrylamide gel electrophoresis (*see* **Note 14**). Stain the gel with Coomassie Brilliant Blue dye solution. Destain in several changes of destain solution (*see* **Note 15**). **Fig. 2** shows the crude lysates obtained from clones expressing various transcription factors.

3.2.2. Rapid Protein Miniprep

1. Use 0.5 mL of overnight culture to inoculate 2.5 mL LB (kan + amp) medium as per protocol 3.2.1. Incubate cultures with vigourous shaking at 37°C for 45 min (*see* **Note 13**).
2. Add 30 µL 100 m*M* IPTG solution. Continue shaking at 37°C for a further 2 h.
3. Transfer 1.5 mL of culture to a microcentrifuge tube and spin at top speed for 1 min.
4. Resuspend cell pellet in 200 µL Buffer B and vortex gently to lyse cells (solution will clear when lysis is complete).

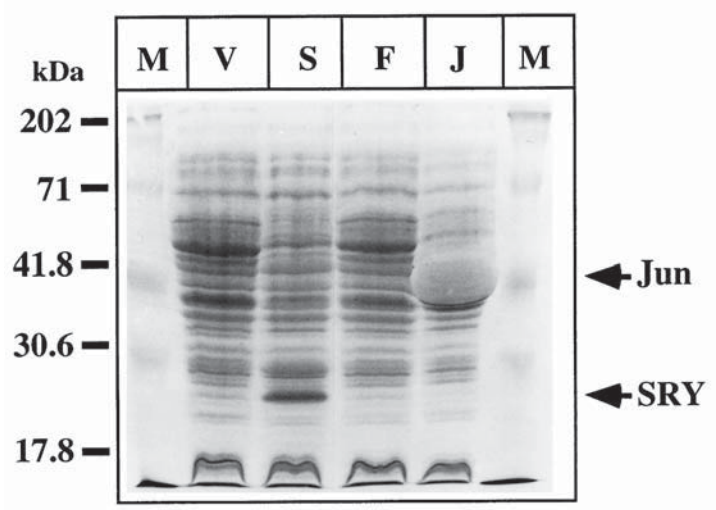


Fig. 2. Crude lysates from bacteria transformed with various expression plasmid constructs. Lanes contain 20 µL aliquots of crude lysates (prepared as given in **Subheading 3.2.1.**) from bacteria transformed with: V, pDS56/RBSII/SphI vector; S, pDS56-6xHis-SRY; F, pDS56-6xHis-Fra1; J, pDS56-6xHis-Jun. M, Kaleidescope Prestained Standards molecular weight markers. The position of the SRY and Jun proteins are indicated; the Fra-1 protein can not be distinguished from bacterial proteins. Samples were run on a 10% SDS-polyacrylamide gel and stained with Coomassie dye.

5. Centrifuge the lysate at top speed for 10 min. Transfer the supernatant to a fresh tube.
6. Add 50 µL of the 50% Ni-NTA resin slurry (*see* **Note 16**) to the clarified supernatant. Mix gently on a rotating wheel at room temperature for 30 min.
7. Pellet the resin by centrifugation at top speed for 20 s. Retain 20 µL of the supernatant for gel analysis (*see* **Note 17**) and remove the remainder by aspiration and discard.
8. Wash the resin three times with 1 mL of Buffer C^U, each time pelleting at top speed for 20 s. After the final wash, spin, and aspiration, spin the tube again briefly to bring final traces of liquid to the bottom of the tube and carefully remove by aspiration.
9. Add 25 µL of Buffer C^U/100 m*M* EDTA to the resin to elute the bound protein. Incubate at room temperature with gentle mixing for 2 min.
10. Pellet the resin by centrifugation for 20 s and remove 20 µL of the supernatant.
11. Add 5 µL of 5x dissociation buffer to each sample and boil for 5 min.
12. Analyze samples by SDS-PAGE and visualise the proteins by Coomassie staining. **Fig. 3** shows c-Jun protein purified by the protein miniprep method, comparing results obtained with Ni-NTA vs TALON resins. Recovery of c-Jun protein from a Ni-NTA spin column is also shown.

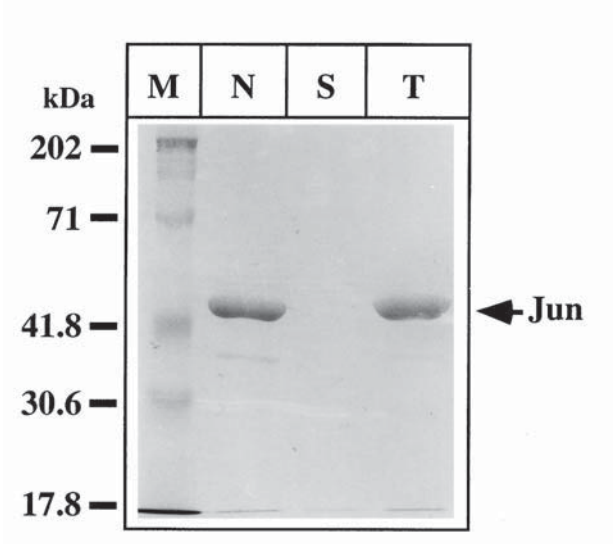


Fig. 3. Comparison of protein recoveries from various resins. A 6 mL culture of bacteria transformed with pDS56-6xHis-Jun was induced for protein synthesis as described in **Subheading 3.2.2.** After 2 h of induction, three 1.5 mL aliquots were removed, the cells were pelleted and resuspended in 200 µL (Ni-NTA miniprep [N]), 400 µL (Ni-NTA spin column [S]), or 500 µL (TALON mini-prep [T]) of Buffer B. Protein purification was carried out for each method according to **Subheading 3.2.2.** (N) or the manufacturer's instructions (S and T). Samples were run on a 10% SDS-polyacrylamide gel and stained with Coomassie dye. For "N" and "T" 20% of the recovered sample was analysed; for "S" 10% of the recovered sample was analyzed.

3.2.3. Optimization of Culture, Binding, and Elution Conditions

When a clone expressing the protein of interest has been identified, optimize conditions for synthesis and recovery of the protein before performing a large-scale purification exercise. There are three important criteria: length of protein induction step, ratio of culture volume to amount of resin, and binding and elution conditions. If denaturing conditions are used for the purification step, it is unnecessary to consider subcellular localization and protein solubility. For purification of proteins under native conditions, protocols for determination of protein solubility can be found elsewhere (for example, in the QIAexpressionist Handbook ***[24]***). Always use a single colony from a plate streak to inoculate the starter overnight culture.

3.2.3.1. Protein Induction

Perform a time-course to determine the optimal period of protein induction.

1. Use 5 mL of an overnight culture to inoculate 25 mL prewarmed LB (kan + amp) medium. Grow with vigourous shaking at 37°C for 45 min (*see* **Note 13**). Remove

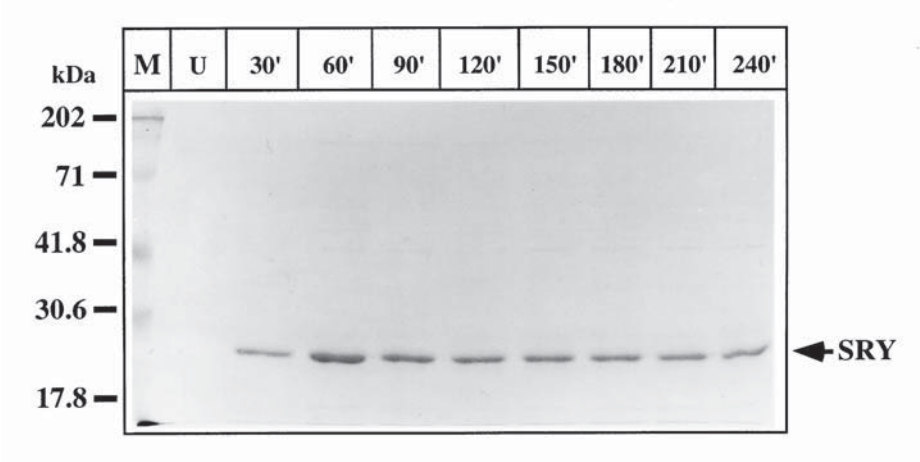


Fig. 4. Time-course of protein induction. A culture of bacteria transformed with pDS56-6xHis-SRY was set up and a time-course of induction was performed as described in **Subheading 3.2.3.1.** 20% of each recovered sample was analysed on a 10% SDS-polyacrylamide gel. U = uninduced culture. The gel was stained with Coomassie dye.

a 1.5 mL aliquot for the uninduced control sample; pellet the bacterial cells in a microcentrifuge tube at top speed for 1 min, remove the supernatant, and store the cell pellet at –20°C.

2. Add 300 µL 100 m*M* IPTG to the remaining culture to induce protein expression. Continue incubation at 37°C. Remove 1.5 mL samples at 30 min intervals and immediately pellet the cells and store at –20°C. Continue collecting samples for 4–5 h.
3. When all the induction time-course samples have been collected, add 200 µL Buffer B to each pellet and proceed with the protein miniprep method (*see* **Subheading 3.2.2.**).
4. Analyze the samples by SDS-PAGE and visualize the proteins by Coomassie staining. **Fig. 4** shows a time-course for induction of the SRY protein.

3.2.3.2. Amount of Resin

This pilot experiment will provide two pieces of information. First, it will indicate the level of synthesis of the protein of interest. Second, it will indicate the optimal amounts of resin and culture volume that will give the greatest purification of the recombinant protein relative to bacterial contaminants. It is very important when scaling up to larger volumes to maintain exactly the same ratio of lysate to resin as determined in this pilot experiment.

1. Use 40 mL of an overnight culture to inoculate 200 mL prewarmed LB (kan + amp) medium. Grow with vigourous shaking at 37°C for 45 min (*see* **Note 13**).

2. Add 2.4 mL 100 m*M* IPTG to induce protein synthesis and continue incubation at 37°C for the amount of time determined to be optimal (usually 1–2 h).
3. Pellet the bacterial cells (e.g., 6000 rpm, 5 min in a GSA rotor) and decant the supernatant.
4. Resuspend the pellet in 5 mL Buffer A (this corresponds to 20 mL Buffer A per 1 liter of culture, or approx 4–5 mL Buffer A per gram of cell pellet (*see* **Note 18**).
5. Mix the cell suspension slowly on a rotator wheel for 60 min at room temperature.
6. Centrifuge the lysate at 10,000*g* for 20 min to pellet debris (11,000 rpm, SS34 rotor).
7. Collect the supernatant and aliquot 2.5 mL, 1.25 mL, 600 µL, and 300 µL to separate tubes.
8. Add 1 mL of a 50% Ni-NTA slurry (equilibrate the resin in Buffer A as described in **Note 16**) to each aliquot of lysate. Mix gently on a rotator wheel for 60 min at room temperature.
9. Load each lysate–resin mixture into a separate empty column. Collect the flowthrough and retain for analysis (if necessary).
10. Wash columns with 2 × 4 mL Buffer C (or C^U). Retain washes for analysis (if necessary).
11. Elute the protein with Buffer D (or D^U), collecting 8 × 250 µL fractions.
12. Analyze aliquots of the fractions by SDS-PAGE (*see* **Note 19**) and visualise the proteins by Coomassie staining.

3.2.3.3. Binding and Elution Conditions

The above analysis will also indicate whether there is likely to be a significant level of contamination with bacterial proteins. If the expression of the recombinant protein is high, then contaminants will be minimal. However, if the protein is poorly expressed, then contaminants such as the bacterial histidine-rich protein can pose a significant problem during purification.

It may be possible to eliminate significant amounts of contaminating proteins by tactics such as:

a. Including ß-mercaptoethanol (10 m*M*) in Buffer A, to remove proteins that might be copurifying as a result of disulfide bridges formed between recombinant and contaminant proteins.
b. Eluting proteins from the resin using a pH gradient, to take account of any differential binding properties of recombinant and contaminant proteins.
c. Including low concentrations of imidazole (10–20 m*M*) in the lysis and washing buffers, which will disrupt the binding of proteins having dispersed histidine residues but should not affect the binding of the tagged proteins having a cluster of histidine residues.
d. Use of a gradient of imidazole in the elution buffer. It is advisable to carry out pilot purification experiments to determine the effectiveness of any or all of these approaches.

Depending upon the level of expression of the recombinant protein, each pilot purification should be carried out using lysate prepared from 100–200 mL of culture.

1. Use an appropriate amount of overnight culture to inoculate prewarmed LB (kan + amp) medium (500–1000 mL) and grow with vigourous shaking for 45 min (*see* **Note 13**).
2. Induce the culture with 1 m*M* IPTG for the determined optimal time.
3. Pellet the bacterial cells and resuspend in Buffer A (20 mL per 1 liter culture). Divide the suspension into five equal portions to be used for the various purification regimes (#1 = ß-ME in Buffer A; #2 = pH gradient during elution; #3 = imidazole in Buffer A; #4 = imidazole gradient during elution; #5 = standard protocol).
4. For portion #1, adjust the lysate to 10 m*M* ß-mercaptoethanol. For portion #3, adjust the lysate to 10 m*M* imidazole. Do not adjust portions #2, #4, or #5. Mix all portions gently on a rotator wheel for 60 min at room temperature.
5. Centrifuge the lysates at 10,000*g* for 20 min to pellet the debris.
6. Transfer the supernatants to fresh tubes and add an appropriate amount of 50% Ni-NTA resin slurry (equilibrated in Buffer A, or Buffer A + 10 m*M* ß-ME, or Buffer A + imidazole as appropriate). Mix gently on a rotator wheel for 60 min at room temperature.
7. Load each lysate-resin mixture into a separate empty column. Collect the flowthrough and retain for analysis (if necessary).
8. Wash each column with 2 × 4 mL Buffer C or C^U (adjust to 10 m*M* imidazole for column #3). Retain washes for analysis (if necessary).
9. Elute the protein with appropriate buffers. For column #1, use Buffer D (or D^U); column #2, use 1–2 mL of buffer at each 0.5 pH unit step, ranging from pH 7.5 to pH 4.5; column #3, use Buffer D (or D^U); column #4, use 1–2 mL of each step of Buffer A + imidazole (5–200 m*M*) gradient; column #5, use Buffer D (or D^U). Collect 250–500 µL fractions.
10. Analyze aliquots of the fractions by SDS-PAGE. Visualize proteins by Coomassie staining.

3.3. Large-Scale Protein Purification

3.3.1. Large-Scale Culture and Column Purification

1. Use a single colony from a plate streak to inoculate 200 mL LB (kan + amp) medium. Grow with shaking overnight at 37°C.
2. Use the overnight culture to inoculate 1 liter prewarmed LB (kan + amp) medium. Divide into 2x 2 L flasks to ensure adequate aeration. Grow with shaking at 37°C for 45 min. Check the OD_{600} to confirm that the cultures have reached sufficient density.
3. Add IPTG to 1 m*M* concentration (6 mL of a 100 m*M* stock solution added to each flask).
4. Continue growth at 37°C for the time determined for optimal induction.

5. Harvest the bacterial cells (GSA rotor, 6000 rpm, 10 min). Decant off the supernatant.
6. Resuspend the pellets in Buffer A (including ß-ME or imidazole as required), using a total of 24 mL. Ensure that cell clumps are properly dispersed. Pool suspended pellets in an SS34 tube.
7. Mix gently on a rotator wheel for 60 min at room temperature.
8. Centrifuge 10,000*g*, 20 min, to pellet debris. Transfer supernatant to a sterile 50 mL tube (e.g., Falcon or Corning).
9. Add an appropriate amount of Ni-NTA resin that has been equilibrated in Buffer A. Mix gently on a rotator wheel for 60 min at room temperature (or overnight at 4°C).
10. Transfer the lysate-resin suspension to a clean, empty column. It will be necessary to allow the column to run while the resin is settling. Collect the flowthrough.
11. After all of the suspension has been loaded onto the column, wash the column with 10–20 mL Buffer A. Collect the flowthrough and retain for analysis if necessary.
12. Wash the column with 10 mL Buffer C (or C^U). Collect the flowthrough.
13. Elute the protein with Buffer D (or D^U). Collect 0.5–1 mL fractions (depending upon the packed volume of the column). Use 25–50 µL of each fraction in a rapid protein determination assay (e.g., BioRad Protein Assay) to identify which fractions contain protein. Analyze aliquots of the load, washes and elution fractions by SDS-PAGE (*see* **Note 19**). For a typical column purification profile, *see* **Fig. 5**.

3.3.2. Dialysis and Recovery of Protein

1. Pool fractions containing protein and dialyze at 4°C against a suitable buffer, such as 25 m*M* Na phosphate pH 7.4, 10% glycerol, 50 m*M* KCl, 5 m*M* $MgCl_2$, 1 m*M* DTT (*see* **Note 20**). Use three changes of dialysis buffer, 100x pooled fraction volume each change.
2. Remove the dialyzed material from the dialysis tubing. If a precipitate has formed, transfer all of the material as a suspension into 1.5 mL microcentrifuge tube(s) and spin at top speed for 5 min (at 4°C) to pellet the precipitated material (*see* **Note 21**).
3. Carefully remove the supernatant to a fresh tube. Carry out a protein determination assay and set aside 20 µL for SDS-PAGE analysis. Divide the dialyzed material into small aliquots (200–300 µL) and snap-freeze in dry ice. Store at –70°C.
4. Add 500–1000 µL fresh dialysis buffer to the precipitated material and resuspend. Remove a 20 µL aliquot of the suspension for analysis, and then spin the tube again and remove a 20 µL aliquot of the supernatant for analysis. Store the suspension at –70°C.
5. Analyze 2 µL, 5 µL, and 10 µL aliquots of each of the samples set aside (from **steps 3** and **4**, above) by SDS-PAGE. Visualize the proteins by Coomassie staining. **Fig. 6** shows recovery of 6xHis-SRY HMG-box protein after large-scale column chromatography and dialysis.

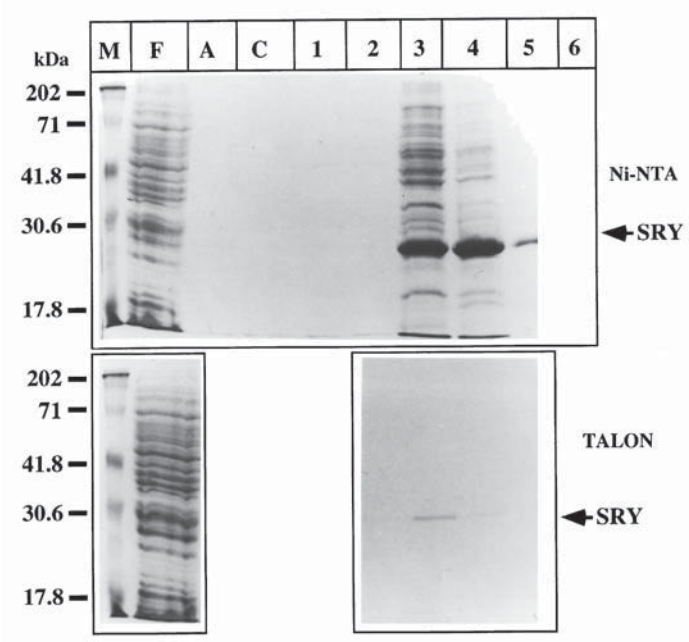


Fig. 5. Large-scale column purification under denaturing conditions. Bacterial cells from 1.5 L of culture (M15 transformed with pDS56-6xHis-SRY; induced for protein synthesis for 2 h) were resuspended in 30 mL Buffer A. The cleared lysate was divided into two portions: 20 mL was mixed with 6 mL 50% Ni-NTA slurry; the remaining 10 mL was mixed with 2 mL 50% TALON slurry. After gentle mixing overnight at 4°C, the slurry-protein mixtures were loaded into clean columns. Protein loading, washes, and elution were carried out as described in **Subheading 3.3.1.** (Ni-NTA) or according to the supplier's instructions (TALON). Protein was eluted in 1 mL fractions (buffer D) from Ni-NTA resin and in 350 µL fractions (buffer D^U) from TALON resin. From the Ni-NTA column, aliquots of the flowthrough (F), buffer A wash (A), buffer C wash (C) and elution fractions 1–6 were analysed on a 13% SDS-polyacrylamide gel. The flow-through and elution fractions 2–5 from the TALON column are shown for comparison. The gels were stained with Coomassie dye.

3.4. Other Methods for Problematic Situations (*see* **Note 22**)

3.4.1. Colony Hybridization to Identify Recombinant Clones

In some cases, the ligation of the insert into the expression vector may occur at greatly reduced efficiency. If this happens, and if the vector-alone ligation produces a high background after transformation of M15 cells, then the frequency of clones bearing recombinant plasmids can be quite low. Under such circumstances, it may be advantageous to screen the colonies from the transformation by hybridiaation with an insert-specific probe.

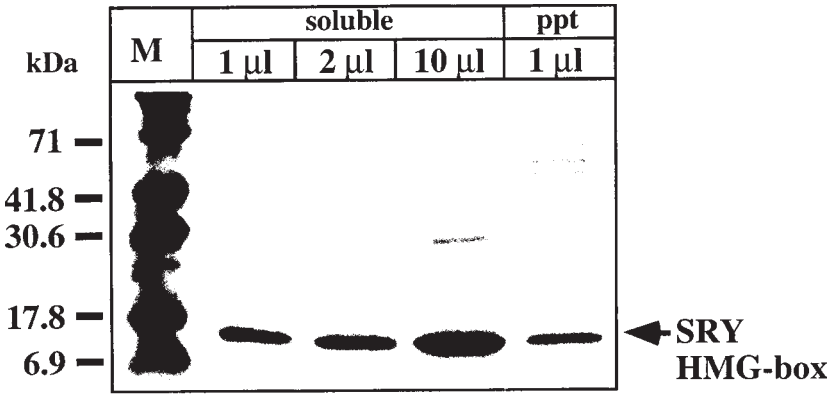


Fig. 6. Recovery of purified proteins following dialysis. 6xHis-(SRY)HMG-box protein from a 500 mL culture was purified under denaturing conditions using 1.5 mL (bead volume) of Ni-NTA resin. Column fractions containing eluted protein were pooled (1.6 mL total) and dialyzed against the buffer described in **Subheading 3.3.2.** After dialysis a precipitate had formed. The soluble protein was removed; 1 µL, 2 µL, and 10 µL aliquots were analyzed on a 15% SDS-polyacrylamide gel. The precipitate was resuspended in 100 µL dialysis buffer and 1 µL of the suspension was also analyzed. The gel was stained with Coomassie dye.

1. Place a marked nitrocellulose filter onto an LB (kan + amp) agar plate, and have a second, similarly labeled plate available to use as a master plate.
2. Using sterile toothpicks, pick colonies from the transformation plates. Spot the bacterial cells first onto the master plate and then onto an identical position on the nitrocellulose filter (a grid pattern is particularly useful). Include a colony from the vector-only transformation.
3. Place both plates at 37°C overnight to allow the colonies to grow.
4. Seal the master plate and store at 4°C. Remove the nitrocellulose filter from the other plate and process (denature and neutralize the colonies) as described in Sambrook et al. ***(21)***.
5. Bake the filter and then prehybridize in a suitable solution. Prepare an insert-specific probe (by PCR-labeling, nick translation, primer end-labeling, or any other convenient method) and allow the probe to hybridize to the filter under appropriate conditions (*see* **21–23**).
6. Wash to remove unhybridized probe and expose the filter to X-ray film. Align the autoradiogram and the filter to identify positive colonies.

3.4.2. Additional Purification Steps

Occasionally, the recombinant protein is so poorly expressed that the level of contamination of the purified material by bacterial proteins is unacceptably high. This was the case in purifying the AP-1 family member Fra-1 (E. McLellan and D. R. Cohen, unpublished data). Attempts to remove the bacte-

rial histidine-rich protein (the major contaminant as determined by protein sequencing) by use of imidazole, ß-mercaptoethanol, or pH gradients, were completely unsuccessful. Finally, the Fra-1 protein was gel purified, and the following procedure worked well to yield relatively pure, functional protein.

1. Run the material eluted from the Ni-NTA column on a preparative SDS-polyacrylamide gel.
2. Cut off a reference lane and stain in Coomassie blue to visualize the proteins.
3. Align the stained gel with the remainder of the gel and cut out the region corresponding to the protein of interest.
4. Electroelute the protein from the gel.
5. Remove the SDS from the eluted material using a detergent removing gel, such as Extracti-Gel™D from Pierce, following the manufacturer's instructions.
6. Dialyze the recovered protein against a suitable buffer.

4. Notes

1. Check that the digestion of the vector is complete by agarose gel electrophoresis. Compare 1 µL of the digestion reaction to approximately 100 ng of undigested plasmid DNA.
2. Back-extraction of the organic phase after the first PCI extraction improves DNA recovery.
3. Addition of 10 µg of yeast tRNA as a carrier during ethanol precipitation is recommended, and will improve the recovery of the vector without interfering with later ligation steps.
4. Check the recovery of the digested DNA by agarose gel electrophoresis; run a 1 µL aliquot and a 0.1 µL aliquot (i.e., 1 µL of a 1:10 dilution) of the resuspended DNA.
5. If the insert is to be prepared by resection from another plasmid, perform a restriction digestion reaction with appropriate enzymes. Digest sufficient plasmid to produce 1–2 µg of insert fragment. Purify the fragment by agarose gel electrophoresis and recover the DNA by electroelution or with the assistance of a commercially available gel extraction reagent.
6. Other thermostable DNA polymerases having proofreading activity, such as *Pfu* polymerase, *Tli* polymerase, or Vent/Deep Vent polymerases may also be used; enzyme combinations such as KlenTaq, eLONGase, or Expand are also suitable. These are available from a variety of commercial sources.
7. These conditions are an indication only; PCR amplification conditions may have to be determined on an individual basis. The lower annealing and extension temperatures used in the first cycle take account of the long noncomplementary "tail" of the primers.
8. It is advisable to confirm the success of the PCR step before proceeding with the next steps. Analyze a 5 µL aliquot of the amplification reaction by agarose gel electrophoresis.
9. If possible, avoid using an ethanol precipitation step in the recovery of the PCR product. Ethanol precipitation, if performed on amplified DNA that has not been

sufficiently "cleaned up" by PCI and CI extractions, seems to result in a nonreversible impairment of the DNA for subsequent functions, such as restriction digestion and ligation. There are a range of spin column products available for removal of unincorporated nucleotides. Some of these (e.g., Pharmacia MicroSpin S-200 HR Columns or cDNA Spun Columns) are also suitable for removing excess primers, but the product chosen will depend upon the size of the amplified product to be purified and the length of the oligonucleotide primers used for the amplification.

10. It may be necessary to carry out this reaction overnight to ensure sufficient digestion has occurred. There is no benefit in comparing undigested and digested material by gel analysis, as the change in fragment size following digestion will be insignificant.
11. Set up a panel of ligation reactions using 150–200 ng of linearized vector DNA and various amounts of insert DNA representing different molar ratios of vector to insert. Always include a vector-only control.
12. Sambrook et al. ***(21)*** gives methods for preparation of competent cells for use in transformation. Fresh or frozen competent cells are acceptable for this purpose.
13. After 45 min of incubation at 37°C, a 1:6 dilution of overnight culture should have reached a density equivalent to 0.7–1.0 OD_{600} units. It is advisable to check that this density has been reached before adding the IPTG to induce protein synthesis.
14. Protein minigels are useful for analysis of large numbers of crude lysates. Kaleidescope Prestained Standards from BioRad provide a convenient marker.
15. If the protein of interest is highly expressed, it will be readily apparent on the Coomassie stained gel which of the colonies contain recombinant plasmids (*see* **Fig. 2**, SRY, and Jun proteins). Occasionally, the protein is not very highly expressed, or the protein band is obscurred by high levels of endogenous bacterial proteins of similar size (**Fig. 2**, Fra-1 protein). In these cases, the SDS-PAGE analysis may not give a completely unequivocal result, and it is advisable to confirm that the colonies believed to be expressing the protein of interest do contain recombinant plasmids. Perform mini-plasmid DNA preps (e.g., a rapid boiling miniprep method; *see* **ref. *21***) using some of the overnight culture set aside, and then digest with enzymes that will resect the insert from the recombinant plasmid. Remember that the DNA sample contains both the expression plasmid and pDMI.1 plasmid from the M15 cells, and the latter will contribute extra bands visible in the digested sample. Digestion of DNA prepared from the vector-only colony will serve as a useful control.
16. Equilibrate the Ni-NTA resin in Buffer B. Transfer an appropriate volume of resin slurry from the stock to a 1.5 mL tube, pellet the resin at top speed for 20 s, remove the supernatant by careful aspiration and resuspend the resin with 2 volumes of Buffer B. Repeat this washing step twice and after the final aspiration step, add 1 volume of Buffer B to create a 50% slurry.
17. These samples are for analysis in the event that little protein is recovered from the resin. If sufficient binding of the recombinant protein has occurred, this analysis will not be necessary.

18. For poorly expressed proteins, the QIAexpressionist Handbook ***(24)*** suggests that larger volumes of culture should be grown and effectively "concentrated" by resuspending the cell pellet in smaller volumes of lysis Buffer A. The problem encountered by this approach is that the bacterial cells may not lyse completely in too small volumes of lysis buffer, and the lysate may be very viscous. In practice, larger volumes of lysis buffer are required when larger volumes of culture (and therefore larger numbers of bacterial cells) are harvested. A ratio of 4–5 mL Buffer A per gram of cell pellet gives very good results.
19. If the elution buffer is prepared using guanidine HCl, then the guanidine will precipitate in the presence of SDS. It is possible to load such samples onto SDS-polyacrylamide gels: add 5 µL of each fraction to 45 µL 1x dissociation buffer and after boiling for 5 min, keep the sample tubes in the hot water while loading 25–30 µL quickly onto the gel. Precipitate does form in the wells, but the samples will electrophorese properly. Alternatively, the samples may be dialyzed to remove the guanidine HCl before analysis (this is not very feasible for large numbers of small fractions), or the elution buffer may be prepared using urea rather than guanidine (the advantage of guanidine is that it is a stronger denaturant).
20. Appropriate dialysis buffer conditions may have to be determined for each individual protein. The conditions given here have proved suitable for a number of transcription factors, including c-Fos, wbFos, Fra-1, SRY, and the SRY HMG-box domain. Gradual renaturation of the proteins can be facilitated by including urea or guanidine HCl in the dialysis buffer, reducing the concentration stepwise with each change of buffer.
21. In practice, the precipitate that forms appears to contain a high proportion of *E. coli* proteins that have contaminated the purification of the desired recombinant protein (*see* **Fig. 6**). Thus, the precipitation of proteins during dialysis acts as a further purification of the 6xHis-tagged protein. Although some protein of interest will also be precipitated, much of this can be resolubilized if desired.
22. If controls are performed at all stages in the generation of the expression plasmid and during the expression and purification steps, then it will be relatively straightforward to identify potential problems that might contribute to lack of success. Starting with an expression construct that is known to give a correct, in-frame fusion of the 6xHis tag and the cDNA coding region is vitally important, and it is worth sequencing the junctions of the construct (if nothing else) before proceeding with other steps. The QIAexpressionist Handbook ***(24)*** from QIAGEN provides an excellent Troubleshooting Guide for problems that might be encountered during expression or purification of the 6xHis-tagged recombinant protein.

References

1. Porath J., Carlsson, J., Olsson, I., and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258,** 598-599.
2. Hoculi, E. Dobeli, H., and Schacher, A. (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatog.* **411,** 177–184.

3. Porath, J. (1992) Immobilised metal ion affinity chromatography. *Prot. Express. Purif.* **3,** 263–281.
4. Papavassiliou, A. G., Bohmann, K., and Bohmann, D. (1992) Determining the effect of inducible protein phosphorylation on the DNA-binding activity of transcription factors. *Anal. Biochem.* **203,** 302–309.
5. Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., Nordheim, A., and Stunnenberg, H.G. (1991) Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci. USA* **88,** 8972–8976.
6. Bugge, T. H., Pohl, J., Lonnoy, O., and Stunnenberg, H. G. (1992) RXRalpha, a promiscuous partner of retinoic acid and thyroid hormone receptors. *EMBO J.* **11,** 1409–1418.
7. Noris, E., Jupin, I., Accotto, G. P., and Gronenborn, B. (1996) DNA-binding activity of the C2 protein of tomato yellow leaf curl geminivirus. *Virology* **217,** 607–612.
8. Abate, C., Luk, D., and Curran, T. (1990) Transcriptional regulation by Fos and Jun in vitro: interaction among multiple activator and regulatory domains. *Mol. Cell. Biol.* **11,** 3624–3632.
9. Abate, C., Luk, D., Gentz, R., Rauscher III, F. J., and Curran, T. (1990) Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: both Fos and Jun contact DNA directly. *Proc. Natl. Acad. Sci. USA* **87,** 1032–1036.
10. Zaret, K. S. and Stevens, K. (1995) Expression of a highly unstable and insoluble transcription factor in *Escherichia coli*: purification and characterisation of the fork head homolog HNF3α. *Protein Expr. Purif.* **6,** 821–825.
11. Kashlev, M., Martin, E., Polyakov, A., Severinov, K., Nikiforov, V., and Goldfarb, A. (1993) Histidine-tagged RNA polymerase: dissection of the transcription cycle using immobilised enzyme. *Gene* **130,** 9–14.
12. Cohen, D. R., Sinclair, A. H., and McGovern, J. D. (1994) SRY protein enhances transcription of Fos-related antigen 1 promoter constructs. *Proc. Natl. Acad. Sci. USA* **91,** 4372–4376.
13. Kato, K., Makino, Y., Kishimoto, T., Yamauchi, J., Kato, S., Muramatsu, M., and Tamura, T. (1994) Multimerisation of the mouse TATA-binding protein (TBP) driven by its C-terminal conserved domain. *Nucleic Acids Res.* **22,** 1179–1185.
14. Hoffmann, A. and Roeder, R. G. (1991) Purification of his-tagged proteins in non-denaturing conditions suggests a convenient method for protein interaction studies. *Nucleic Acids Res.* **19,** 6337–6338.
15. Rauscher III, F. J., Morris, J. F., Tourney, O. E., Cook, D. M., and Curran, T. (1990) Binding of the Wilms Tumor locus zinc finger protein to the EGR-1 consensus sequence. *Science* **250,** 1259–1262.
16. Corvello, C. M., Metz, R., Bravo, R., and Armelin, M. C. (1995) Expression and characterisation of mouse cFos protein using the baculovirus expression system: ability to form functional AP-1 complexes with co-expressed cJun. *Cell. Mol. Biol. Res.* **41,** 527–535.

17. Lorch, Y. and Kornberg, R. D. (1994) Isolation of the yeast histone octamer. *Proc. Natl. Aca. Sci. USA* **91,** 11,032–11,034.
18. Rosenberg, A. H., Goldman, E., Dunn, J. J. Studier, F. W., and Zubay, G. (1993) Effects of consecutive AGG codons on translation in Escherichia coli demonstrated with a versatile codon test system. *J. Bacteriol.* **175,** 716–722.
19. Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.-T., and Dobberstein, B. (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. *Meth. Enzymol.* **155,** 416–433.
20. Certa, U., Bannwarth, W., Stuber, D., Gentz, R., Lanzer, M., LeGrice, S., Guillot, F., Wendler, I., Hunsmann, G., Bujard, H., and Mous, J. (1986) Subregions of a conserved part of the HIV gp41 transmembrane protein are differentially recognised by antibodies of infected individuals. *EMBO J.* **5,** 3051–3056.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY.
22. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D. Sedman J. G., Smith J. A., and Struhl, K., eds. (1995) *Current Protocols in Molecular Biology.* Wiley, NY.
23. Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J. (1990) *PCR Protocols—A Guide to Methods and Applications,* Academic, San Diego, CA.
24. *The QIAexpressionist: A handbook for high-level expression and purification of 6xHis-tagged proteins*, (1997) third ed., Qiagen.

18

Generation of Transcription Factors in Rabbit Reticulocyte Lysate Depleted of Endogenous DNA-Binding Protein

Ulrich Kruse, Thorsten T. Ebel, and Albrecht E. Sippel

1. Introduction

The rabbit reticulocyte lysate (RRL) is a convenient system for in vitro translation of transcription factors ***(1)***. By using in vitro-synthesized transcription factors, the time-consuming and often difficult process of protein purification from cellular extracts of in vivo expression systems is avoided. Although the endogenous level of RNA in the RRL is removed by treatment with micrococcus nuclease, there is a substantial level of proteins in the lysate, among them many DNA-binding proteins. These endogenous proteins can interfere with the function of newly synthesized transcription factors. Problems can be caused by signal overlaps in electrophoretic mobility shift assays (EMSA) and by dimerization of endogenous proteins and in vitro-translated products.

We have used depleted RRL in order to synthesize and characterize members of the Nuclear Factor I family of transcription factors (NFI/CTF or NFI/TGGCA proteins). Nuclear Factor I proteins bind as dimers to the palindromic consensus sequence 5'-YTGGCA$(N)_3$TGCCAR-3' functioning as transcription and replication factors ***(2)***. All NFI proteins have a highly conserved amino-terminal domain in common that is sufficient for DNA binding, dimerization, and stimulation of adenovirus DNA replication ***(3,4)***. The DNA-binding domain of NFI shows no obvious sequence similarity to any known class of DNA-binding domains, such as zinc finger, leucine zipper, or helix-loop-helix motifs. Members of the NFI family of proteins are ubiquitously expressed in all vertebrate cell types tested so far, including reticulocytes. When full-length NFI cDNAs are synthesized in RRL and DNA binding is examined by EMSA,

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

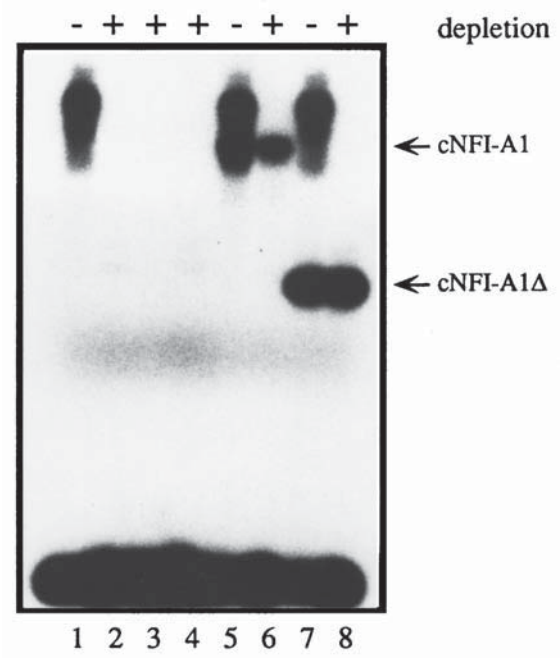


Fig. 1. Comparison of untreated and NFI-depleted rabbit reticulocyte lysate in EMSA. Parallel transcription/translation reactions expressing either no (lanes 1–4), a full length (lanes 5 and 6), or a truncated (lanes 7 and 8) chicken NFI-A1 cDNA ***(3)*** were performed as described in Figure 2. Aliquots of each reaction and a ^{32}P-labeled DNA fragment containing the NFI consensus binding site were used in EMSA as described (3). 0.5 mL corresponding to 1/50 of the translation reaction were used in the binding reactions (lanes 1 and 5-8), except for lane 2 (1 mL), lane 3 (2.5 mL) and lane 4 (5 mL) to confirm complete removal of endogenous rabbit NFI proteins. (This figure was reproduced from ***(5)*** by permission of Oxford University Press).

it is difficult to detect specific shift signals of full-length NFI proteins (*see* **Fig. 1**, lane 5) because of a high background signal arising from endogenous NFI complexes (*see* **Fig. 1**, compare lanes 1 and 5) .

In order to remove endogenous NFI proteins from RRL, we have developed a simple method using streptavidin-coated magnetic beads and a biotinylated DNA fragment comprising the consensus NFI binding site ***(5)***. This protocol resulted in complete removal of NFI DNA binding activity from the lysate (*see* **Fig. 1**, lanes 2 to 4) and allowed unambiguous detection of the newly synthesized full-length NFI-A1Δ isoform (*see* **Fig. 1**, lane 6). In addition, the identical strength of the shift signals caused by the newly synthesized NFI-A1D protein in untreated (*see* **Fig. 1**, lane 7) or depleted (*see* **Fig. 1**, lane 8) RRL

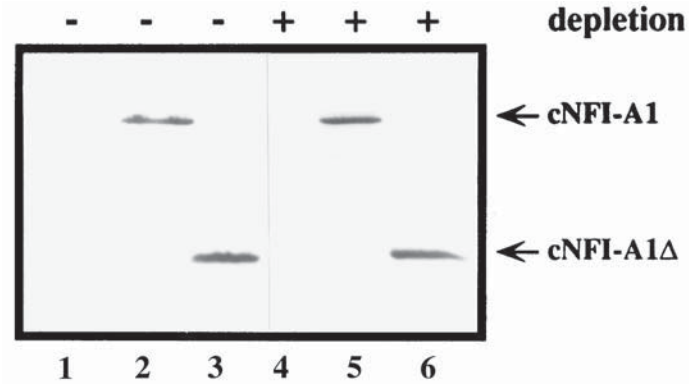


Fig. 2. Comparison of the translation efficiency of untreated and NFI-depleted TNT™-RRL. TNT™-RRL from one batch was split and one-half was depleted from endogenous NFI. Parallel transcription/translation reactions were performed in which each lysate was programmed either with no (lanes 1 and 4), 1 µg of a full-length chicken NFI-A1 cDNA (lanes 2 and 5), or 1 µg of a truncated (lanes 3 and 6) NFI-A1Δ cDNA ***(3)***. All reactions were performed in a total volume of 25 µL following manufacturer's instructions and using ^{35}S-methionine (Amersham, Piscataway, NJ; 1000 Ci/mmol) for protein labeling. 2.5 µL of each reaction were examined on a denaturing 12% polyacrylamide gel. (This figure was reproduced from ***[5]*** by permission of Oxford University Press).

shows that the depletion procedure did not affect the synthesis activity of the RRL. To confirm this observation, we compared ^{35}S-labeled translation products from untreated and NFI-depleted RRL on a denaturing polyacrylamide gel. There was no detectable influence on the amount or quality of the expressed protein (*see* **Fig. 2**, lanes 2 and 3 compared with lanes 5 and 6, respectively).

The depletion method can be used to prevent possible interference of DNA binding proteins present in the rabbit reticulocyte and homologous in vitro translated transcription factors in functional assays. The procedure is rapid, highly efficient, and causes no loss of translation efficiency. This method should be widely applicable for depletion of specific endogenous DNA-binding proteins from RRLs provided their target DNA-binding sequence is known.

Here we show one application (*see* **Fig. 3**), a quantitative EMSA with NFI transcription factors synthesized in depleted RRL. After removal of the endogenous sequence-specific DNA-binding activity, the unobstructed comparison of the relative DNA-binding strengths of full length and truncated NFI variants is possible. For NFI-A1, B2, and C2, we detect a significant increase in DNA-binding affinity upon removal of the carboxyl-terminal part of the protein.

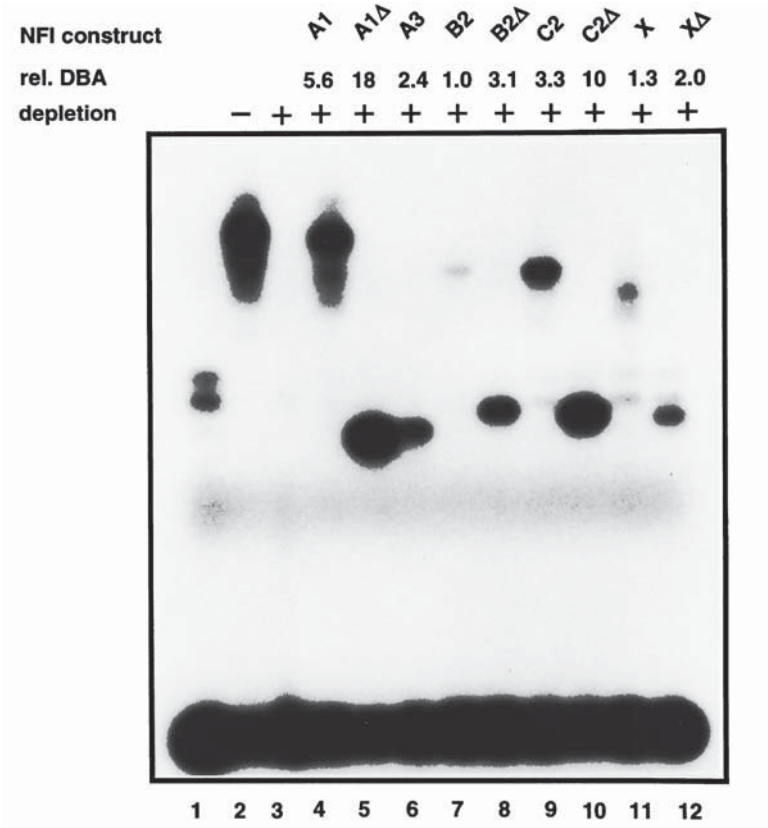


Fig. 3. Determination of relative DNA binding affinities of NFI variants. Full length NFI proteins (A1, B2, C2, and X) (3) and truncated versions comprising the DNA-binding and dimerization domain (A1Δ, A3, B2Δ, C2Δ, and XΔ) were synthesized in the presence of ^{35}S-methionine. Equimolar amounts of proteins were analyzed by EMSA using a ^{32}P-labeled DNA fragment as described *(3)*. DNA-binding reactions without lysate (lane 1), with undepleted lysate (lane 2), and depleted unprogrammed lysate (lane 3) served as controls. Signals of protein-DNA complexes were quantified with a phospoimager and relative DNA-binding activity (rel. DBA) calculated normalized to the weakest signal (lane 7, B2) (*see* **Note 3**).

2. Materials

For all solutions, RNase-free deionized water (DEPC water) is used.

1. Biotinylated oligonucleotides: biotinylated oligonucleotides are commercially available from a number of companies. Direct incorporation of biotin at the 5'

end of the oligonucleotide during DNA synthesis using a biotin phosphoramidite with a spacer arm (at least six C-atoms) is recommended (*see* **Note 1**). Complementary oligonucleotides comprising the recognition site for the DNA-binding protein of interest are synthesized and subsequently annealed. The length of oligonucleotide should be 20–30 bp (*see* **Note 1**). Oligonucleotides are purified by HPLC after synthesis.

2. Annealing buffer for oligonucleotides: 10 m*M* Tris-HCl, 100 m*M* NaCl, pH 7.4.
3. Phosphate buffer saline/bovine serum albumin (PBS/BSA) wash solution: PBS, 0.1% BSA.
4. Magnetic particle concentrator (MPC; Dynal) for Eppendorf microcentrifuge tubes and streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin; Dynal, Inc., Lake Success,NY 11042).
5. TNT™ coupled reticulocyte system (Promega, Madison, WI). Proteins can be produced directly from protein-encoding DNA sequences cloned downstream of T3, T7, or SP6 RNA polymerase promoters.

3. Methods

3.1. Preparation of a Biotinylated, Double-Stranded DNA Probe (*see* Note 1)

1. Dissolve each biotinylated oligonucleotide in 10 m*M* Tris-HCl, 100 m*M* NaCl at concentration of 50 pmol/µL.
2. Mix equal volumes of complementary oligonucleotides, resulting in a final concentration of 25 pmol/µL.
3. Anneal DNA strands by heating to 90°C for 1 min and allow to cool down to room temperature.
4. Keep annealed oligonucleotides on ice or store at –20°C.

3.2. Pretreatment of Streptavidin-Coated Magnetic Beads

1. Resuspend beads by gently shaking the vial to obtain a homogeneous suspension.
2. Add the appropriate amount of beads to a microcentrifuge tube and mix gently; 2.5×10^7 beads are sufficient to deplete 200 µL of lysate.
3. Place the tube in the magnetic particle concentrator (MPC) for 30 s.
4. Remove the supernatant by aspiration with a pipet while the tube remains in the MPC.
5. Remove the tube from the MPC. Add 1.0 mL PBS/BSA wash solution and resuspend gently.
6. Repeat **steps 3** and **4** two more times (a total of three wash cycles).
7. Remove the tube from the MPC and resuspend beads in 15 µL of deionized water.

3.3. Depletion of RRL (*see* Note 2)

1. Thaw 200 µL TNT-RRL quickly and place on ice.
2. Add 1 µL (25 pmol) of double-stranded DNA fragment containing the appropriate DNA-binding site.

3. Mix the solution gently by pipeting up and down. Incubate for 10 min on ice.
4. Add 2.5×10^7 pretreated streptavidin coated magnetic beads (*see* **Note 2**). Mix by gently pipetting the solution up and down. Incubate for 5 min on ice. Every 30 s, resuspend the beads by shaking the tube gently to obtain a homogenous dispersion of beads in solution.
5. Place reaction tube in the MPC for 2 min.
6. Remove supernatant carefully and transfer to a new tube and keep on ice. The supernatant represents the depleted RRL and can be used directly for in vitro transcription/translation reactions or it can be frozen and stored in aliquots at –80°C. Do not freeze/thaw lysate more than two times.

3.4. In Vitro Synthesis of Transcription Factors Using a Coupled Transcription/Translation Reaction

Standard RRL translation systems commonly use RNA synthesized in vitro from SP6, T3, or T7 RNA polymerase promoters ***(6)***. This process requires separate transcription and translation reactions with several steps between the reactions (e.g., removal of template DNA by DNase digestion and purification of mRNA). The TNT™ coupled reticulocyte lysate system (Promega) bypasses these steps by incorporating the transcription reaction directly into the translation mix. Translation reactions are performed according to manufaturer's instructions.

3.5. Analysis of Proteins (*see* Note 3)

In vitro synthesized proteins can be analyzed by EMSA, methylation protection, or footprinting techniques (see other chapters in this volume). However, EMSA is the most sensitive and convenient assay. The choice of method depends on the amount of protein synthesized (*see* **Note 3**) and the affinity of the protein for its DNA-binding site. Proteins can also be analyzed by EMSA in regard to homo- or heterodimerization by mixing or cotranslation of full-length and truncated forms of the protein (***3,7***). **Fig. 3** shows a quantitative functional DNA-binding assay of in vitro synthesized transcription factor variants.

4. Notes

1. Larger DNA fragments can be biotinylated by using PCR with one biotinylated primer or nick translation using DNA polymerase I (Klenow fragment) with one biotin dNTP. All excess biotin must be removed after biotinylation, since free biotin will occupy binding sites and reduce the binding capacity for biotinylated DNA.
2. It is recommended to perform a titration to optimize the quantity of beads used because DNA fragment size and biotinylation procedures can affect binding capacity of the beads.
3. The amount of synthesized protein can be estimated by in vitro translation in the presence of ^{35}S-methionine, followed by SDS-polyacrylamide gelelectrophoresis

and autoradiography. If the protein is used for EMSA with a ^{32}P-labeled DNA probe, radiation originating from the ^{35}S-labeled protein can create a diffuse background signal. This can be avoided by placing two sheets of paper between the EMSA gel and X-ray film, thereby shielding the signal from the ^{35}S-labeled protein.

References

1. Pelham, H. R. B. and Jackson, R. J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem*. **67,** 247–256.
2. Rupp, R. A. W., Kruse, U., Multhaup, G., Göbel, U., Beyreuther, K., and Sippel, A. E. (1990) Chicken NFI/TGGCA proteins are encoded by at least three independent genes: NFI-A, NFI-B and NFI-C with homologues in mammalian genomes. *Nucleic Acids Res.* **18,** 507–516.
3. Kruse, U. and Sippel, A. E. (1994) Transcription factor nuclear factor I proteins form stable homo- and heterodimers. *FEBS Letts.* **348,** 46–50.
4. Dekker, J., van Oosterhout, J. A. W. M., and van der Vliet, P. C. (1996) Two regions within the DNA binding domain of nuclear factor I interact with DNA and stimulate Adenovirus DNA replication independently. *Mol. Cell. Biol.* **16,** 4073–4080.
5. Ebel, T. T. and Sippel, A. E. (1995) A rapid method to deplete endogenous DNA-binding proteins from reticulocyte lysate translation systems. *Nucleic Acids Res*. **23,** 2076–2077.
6. Olliver, L. and Boyd, C. D. (1996) In vitro translation of messenger RNA in rabbit reticulocyte lysate cell-free system, in *Basic DNA and RNA Protocols* (Harwood, A.J., ed.), Humana, Totowa, NJ, pp. 477–484.
7. Hope, I. A. and Struhl, K. (1987) GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J*. **6,** 2781–2784.

19

Electrophoretic Mobility Shift Assays

Peter L. Molloy

1. Introduction

Perhaps the most common technique used in the study of DNA-binding proteins is the electrophoretic mobility shift assay (EMSA) or gel shift assay. It can be used with crude protein mixtures or purified proteins in studies of, for example, the DNA sequence requirements of binding, kinetics of binding, identification and characterization of binding proteins, and cofactor requirements. Using radioactively labeled DNA probes, detection of proteins present in subfemtomolar amounts is readily possible. The principle of the assay is very simple—DNA fragments and proteins are mixed in a suitable buffer and binding is allowed to occur. The mixture is then separated by nondenaturing gel electrophoresis; stable complexes of DNA and protein are generally significantly retarded in mobility in comparison with the free DNA and the separated complexes are normally viewed by phosphorimaging or autoradiography. Despite the basic simplicity of the method there are many factors in addition to the need to optimize basic buffer requirements important to its successful application. In all reactions, especially those involving crude cellular or nuclear extracts, there are a number of competing reactions involving both the specific DNA probe and its protein target. The DNA probe may be bound nonspecifically by other proteins, removing it from the pool for specific binding and leading to nonspecific bands, smearing of the probe or trapping in the sample well. Similarly, the specific DNA binding protein will bind with variable affinity to competitor nonspecific DNA added to the reaction and may be specifically bound or nonspecifically aggregated with other proteins in the reaction. The aim in any application should be to maximize the amount of probe forming specific complexes while minimizing the sequestering of probe or target protein by other components in the reaction.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

In this chapter a basic protocol for EMSAs using nuclear extracts (*see* **Note 1**) is provided along with discussion notes on optimization of binding conditions (*see* **Notes 11–22**) and testing for specificity of binding (*see* **Notes 23–27**). Also included are examples of ways in which the basic assay can be extended to obtain more information on specific protein-DNA complexes (*see* **Notes 32–37**).

2. Materials

All solutions should be prepared using high-quality distilled or deionized water (such as MilliQ, Millipore, Bedford, MA).

1. 2% Nonidet P40.
2. Oligonucleotides: oligonucleotides purified using OPC columns (Applied Biosystems), or similar, are normally sufficiently pure for use in probe preparation.
3. 100 m*M* $MgCl_2$: prepare from fresh solid using sterile water and filter sterilize.
4. TNAE buffer: 6.7 m*M* Tris-HCl, pH 7.9, 3.3 m*M* sodium acetate, 1 m*M* Na_2EDTA. Adjust the pH of a 10X stock buffer to 7.9.
5. TBE buffer: 89 m*M* boric acid, 89 m*M* Tris base, 2 m*M* Na_2EDTA. Prepare a 5X stock solution.
6. Tris-Glycine buffer: 50 m*M* Tris base, 380 m*M* Glycine, 1.67 m*M* Na_2EDTA. Prepare a 5X stock solution.
7. EMSA gel loading buffer: Buffer A without DTT and containing 0.02% bromophenol blue.
8. 40% acrylamide/1% bisacrylamide.
9. 10% ammonium persulphate.
10. TEMED.
11. 10% acetic acid.
12. Binding Buffer A: 20 m*M* HEPES , pH 7.9, 100 m*M* KCl, 2 m*M* dithiothreitol (DTT), 1 m*M* EDTA, 20% glycerol. The buffer is prepared by first dissolving the HEPES base (Calbiochem, La Jolla, CA), KCl and EDTA and adjusting the pH to 7.9 using 2 *M* KOH. The glycerol is then added and the volume adjusted prior to autoclaving. This buffer can be stored at room temperature. After adding DTT, the buffer should be stored in aliquots at –20°C.
13. poly (dI-dC)·poly(dI-dC), poly (dG-dC)·poly(dG-dC), poly(dA-dT)·poly(dA-dT) (Pharmacia Biotech, Uppsala, Sweden): dissolve at 1 mg/mL in 10 m*M* Tris.HCl, 0.1 m*M* EDTA, pH 8.0. Store at –20°C.
14. 10 mg/mL nuclease free BSA (Promega, Madison, WI): store at –20°C.
15. Mammalian cell nuclear extract, dialyzed against Buffer A (*see* **Note 1**).
16. α-[^{32}P]-dATP or dCTP (>3000 Ci/mmol, 10 mCi/mL).
17. γ-[^{32}P]-ATP (>3000 Ci/mmol, 10 mCi/mL).
18. Klenow fragment: 1 U/μL.
19. 10 m*M* dNTPs.
20. T4 polynucleotide kinase: 10 units/μL.

21. 10X T4 polynucleotide kinase buffer: 700 m*M* Tris-HCl pH 7.6, 100 m*M* $MgCl_2$, 50 m*M* DTT (or as supplied with enzyme).
22. Restriction enzymes and 10X buffers.
23. Vertical gel apparatus: 20 × 20 cm plates, with 0.75–1 mm spacers and combs.
24. Phosphorimaging system or X-ray film and cassettes.

3. Methods

3.1. Preparation of Restriction Fragment Probes (*see* Notes 2–5)

1. Digest 2 µg of plasmid DNA in 20 µL reaction with appropriate enzymes for releasing probe fragment (*see* **Notes 2–4**).
2. To the digest add 2 µL of suitable ^{32}P-labeled deoxynucleotide for end-labelling along with the other three unlabeled dNTPs to a concentration of 100 µM. Add 1 U of Klenow fragment of DNA polymerase I and incubate for 15 min at room temperature; add fourth unlabeled dNTP to 100 µM and incubate for a further 5 min. At this and subsequent steps use Perspex shielding to limit exposure to ^{32}P radiation.
3. Load digest in a 2.5-cm-wide well on a 1-mm-thick 6% acrylamide/0.3% bisacrylamide gel in TBE buffer and electrophorese at 10 V/cm until the bromophenol blue dye is near the bottom of the gel (*see* **Note 5**).
4. Dismantle the gel apparatus leaving the gel on one of the glass plates. Cover with plastic wrap and expose to X-ray film for 2–5 min, marking the film for alignment with the gel.
5. Identify the position of the labeled band and excise the gel slice with a scalpel blade.
6. Elute the fragment in 100 µL of water, overnight on a rocking platform. For short fragments (<50 bp) this should be done in the cold room, otherwise at room temperature.
7. Store the radiolabeled fragment in the presence of 0.5 m*M* DTT or 1 m*M* ß-mercaptoethanol (to limit radiolytic breakdown).

3.2. Preparation of Oligonucleotide Probes (*see* Notes 5–8)

1. In a 20 µL final volume, mix 2 pmol of oligonucleotide, 2 µL of 10X kinase buffer, 3 µL of γ-[^{32}P]-ATP and 1 µL of polynucleotide kinase and incubate at 37°C for 30 min (*see* **Notes 6–8**).
2. Mix together two kinased complementary oligonucleotides (2 pmol of each), add 5 µL of 100 m*M* $MgCl_2$ and adjust to 50 µL. Heat to 60°C and allow to anneal by cooling slowly to room temperature.
3. Add gel loading dye and isolate double-stranded probe from gel, as for restriction fragment probes, but using a 10% acrylamide/0.5% bisacrylamide gel run in TBE buffer (*see* **Note 5**).

3.3. Preparation of Unlabeled Competitors DNAs (*see* Notes 9–10)

1. For restriction fragments follow protocol 3.1 above, but increase the amount of DNA to 5 µg, the incubation volume to 50 µL, and use all four unlabeled

Table 1
Suggested Reactions for Optimizing Binding

Reaction #	1	2	3	4	5	6	7	8	9	10
Nuclear extract µL	0	1	2	4	8	12	4	4	4	4
Buffer A µL	12	11	10	8	4	0	8	8	8	8
BSA µL	1	1	1	1	1	1	1	1	1	1
poly dIC µL	1	1	1	1	1	1	0.5	1.5	2	1
MgCl2 µL										1
DNA probe µL	2	2	2	2	2	2	2	2	2	2
H_2O µL	3	3	3	3	3	3	3.5	2.5	2	2

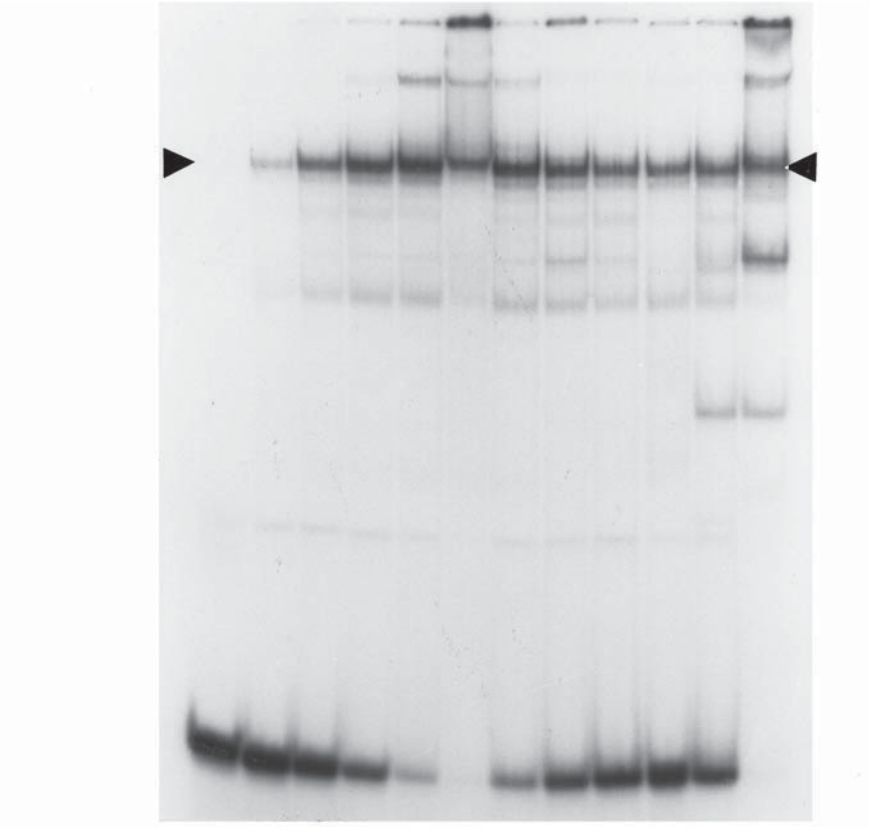
deoxynucleotides at a final concentration of 100 µ*M*. For isolation of the fragment, use a thicker gel (2 mm) or a wider well (*see* **Note 9–10**).

2. For oligonucleotides, anneal 10 pmol of complementary oligonucleotides in 50 µL and gel purify as in the protocol in **Subheading 3.2.** (*see* **Notes 9–10**).

3.4. DNA Binding Reactions (*see* **Notes 11–27** and **32–37**)

Optimal binding conditions need to be determined for each protein–DNA interaction being studied. A range of conditions to optimise levels of protein and nonspecific competitor, as shown in **Table 1**, should be tested initially. Many parameters of DNA-binding reactions can be varied to optimize the amount of specific complexes formed, to minimize nonspecific complexes, and to provide sharper bands. These are discussed in **Notes 11–22**. A good starting point suitable for a wide range of proteins is as follows.

1. To a 1.5 mL microfuge tube, add the following components in the order shown for a 20 µL reaction mix: 3.5 µL H_2O, 12 µL Protein extract (4 µL) plus buffer A (8 µL), 1.0 µL 2% NP-40 , 1.0 µL 10 mg/mL nuclease-free BSA, 0.5 µL 2 mg/mL poly (dI-dC)·poly(dI-dC), 2.0 µL(10 fmol) DNA fragment.
2. Incubate at 30°C for 30 min.
3. Add 2 µL of EMSA loading buffer prior to loading on the gel.
4. **Figure 1** shows an optimization experiment using the binding site for an ets-related protein (GGAA core sequence), found in the N-ras gene promoter. The amount of specific complex formed is seen to increase with increasing levels of



	1	2	3	4	5	6	7	8	9	10	11	12
Nuc. extract (µl)	0	0.2	0.5	1.0	2.0	4.0	1.0	1.0	1.0	1.0	1.0	1.0
poly dI-dC (µg)	1.0	1.0	1.0	1.0	1.0	1.0	0.5	1.5	2.0	1.0		
poly dX-dY, 1µg											dA-dT	dG.dC
$MgCl_2$, 5 mM	-	-	-	-	-	-	-	-	-	+	-	-

Fig. 1. Optimization of EMSA conditions. Binding reactions were set up using a 50 bp fragment containing a binding site for an ets-related protein. Reaction components were as in **Subheading 3.3., Step 1,** except for the variation in specific components as indicated in the table beneath the gel lanes. Free and bound complexes were separated as described in **Subheading 3.4.** The major specific complex is indicated by the arrow.

nuclear extract, until at higher levels a significant fraction of the labeled probe is trapped in the well. A level of 0.5–1 µL of nuclear extract with 1 µg of poly dI-dC competitor provides a clear signal with minimal background. The value of testing alternate nonspecific competitor DNAs is shown by the appearance of additional complexes in the presence of poly dA-dT and poly dG-dC. The specificity of all complexes would need to be tested as discussed in **Notes 23–27**.

3.5. Electrophoresis of Complexes (*see* Notes 28–31)

1. Prepare 50 mL of gel mix for a 0.75- or 1-mm thick, 20 × 20 cm gel with a 12–15 well comb (wells 6–8 mm wide): 1.0 mL 50X TNAE buffer, 6.25 mL 40% acrylamide/1% bisacrylamide mix (*see* **Note 31**), 42.3 mL water, 0.4 mL 10% ammonium persulfate and 40 µL TEMED.
2. Set up the gel apparatus in a cold room or refrigerator with a pump to recirculate the buffer between the upper and lower chambers and electrophorese at 100–150 V for 1 h before loading samples.
3. Load samples into the wells and continue electrophoresis at 300 V until the bromophenol blue has migrated about three-fourths of the distance down the gel. *Be careful as small DNA probes can run considerably ahead of the tracking dye.*
4. Separate the glass plates and fix the gel in 10% acetic acid.
5. Transfer the gel to thick (e.g., 3MM) filter paper and dry prior to phosphorimaging or autoradiography.

4. Notes

1. *Source of DNA Binding Protein.* Proteins used in EMSAs may come from a variety of sources: crude cellular or nuclear extracts, further purified fractions, expressed and purified recombinant proteins, or proteins produced by in vitro transcription/translation of cloned genes. In the standard binding conditions described below, nuclear extract dialyzed against binding buffer A is used as the source of proteins. For specific applications allowance needs to be made for contributions made by the buffer in which the protein is contained (e.g., if obtained by in vitro transcription/translation).

Preparation of Probes and Competitors

2. Digestion and end-labeling of 1–2 µg of plasmid DNA should yield enough of any restriction fragment to perform 50–100 individual binding reactions (2 µg of a 3 kb plasmid equates to 1 pmol, with about 10 fmol being used per reaction).
3. Ideally restriction fragments should be between 40 and 100 bp in length. This provides a balance between affinity, which is often enhanced by nonspecific interaction of proteins with DNA surrounding their specific binding site, and increased nonspecific binding of other proteins with increased fragment length. If the location of binding sites is known, or suspected, it is preferable to try to have these located centrally within the fragment.
4. Restriction sites are best chosen to leave 5' overhangs, which can be filled in and radiolabeled using the Klenow fragment of DNA polymerase I and an appropriate α-[^{32}P]-dNTP in the presence of the other unlabeled deoxynucleotides. It is important to completely end-fill (or remove 3' overhangs) from restriction fragments as single-stranded ends can provide avid binding sites for some proteins.
5. Probes should always be gel purified. For probes 50 bp or less, care should be taken to avoid DNA melting as single-strand DNAs can produce artifactual results. We routinely run gels and elute small DNA fragments in the cold room.

The probe can be eluted from the gel slice in a minimal volume of water or TE and either used directly or ethanol precipitated and resuspended if it is necessary to concentrate the fragment.

6. Oligonucleotide probes used in EMSAs are generally 20–40 bp long and should be designed to allow convenient labeling. It is important with oligonucleotide probes to separate the double-stranded product from the single-stranded oligonucleotides by gel electrophoresis.
7. If the complementary oligonucleotides anneal without overhanging ends, labeling is best accomplished by kinasing one or both single-stranded oligonucleotides (using T4 polynucleotide kinase and γ-[^{32}P]-ATP) prior to annealing. Oligonucleotides that anneal to leave 5' overhangs can be readily labeled by end-filling, using the Klenow fragment of DNA polymerase I and an appropriate α-[^{32}P]-dNTP in the presence of the other unlabeled deoxynucleotides.
8. If mutations or variants of a particular region are to be studied, it is often convenient to prepare a primer of 12–15 bases corresponding to the sequence adjacent to the region to be varied. This primer can then be used to elongate on a set of template oligonucleotides (30–40 bases long) containing centrally the region of interest.
9. Competitor DNA fragments are prepared in the same manner as the radiolabeled probes except that only unlabeled nucleotides are used for end-filling reactions. To obtain sufficient material to enable competitor/probe ratios of up to 100, it is advisable to start with 5 µg of plasmid DNA or 5–10 pmol of oligonucleotides.
10. The position of bands in a preparative gel can be identified by ethidium bromide staining (soaking gel in 0.5 µg/mL ethidium bromide solution for 5–10 min) or by running the equivalent radiolabeled probe adjacent to the preparative track. To achieve sufficient concentration of competitor, the eluted DNA needs to be concentrated by ethanol precipitation.

Optimization of Reaction Parameters

11. *Order of Addition of Components:* There is no strict requirement, but the DNA probe and the protein should not come in contact until the final mixing step. The DNA probe and non-specific competitor DNA may be added together or the non-specific competitor may be added prior to the probe.
12. *Temperature:* Binding reactions are most commonly done at 30°C. For some proteins better binding may be seen at lower temperatures (4–25°C); Sp1, for example, binds better at 20°C than at 30°C.
13. *Time:* Incubations are generally done for 10–30 min, as for most proteins this provides sufficient time for equilibrium to be reached.
14. *pH:* While a pH of 7.9 is routinely used with extracts of mammalian cells, there is no reason that this will be optimal for all proteins, and it is worthwhile to evaluate a range from pH 6.5 to 8.5.
15. *Ionic Strength:* the concentration of monovalent cations is an important parameter in determining the relative amounts of binding of specific and nonspecific complexes. Many nonspecific protein–DNA interactions are ionic in nature and

will be progressively suppressed at higher salt concentrations. Specific hydrogen bonding and hydrophobic interactions between the bases and amino acids will be less affected. A significant part of the binding energy of most specific protein-DNA complexes is still provided, however, by interactions of positively charged amino acids with the phosphate backbone of the DNA, and higher salt concentrations will eventually disrupt complexes. When working with crude protein extracts, salt concentration in the range of 50–100 m*M* are generally optimal, but individual complexes can sometimes be differentiated by their stability at concentrations up to 200 m*M*. For some proteins (especially when purified) it can be advantageous to use lower or no added monovalent ion.

16. *Divalent Ions:* The most common divalent ion is Mg^{2+}. The concentration optimal for binding generally is in the range 0–10 m*M*, with 5 m*M* being a good initial concentration to test. The binding of some proteins is very sensitive to the presence of Mg^{2+}. For example, the presence of Mg^{2+} alters the equilibrium binding of the helix-loop-helix protein USF, and also dramatically increases both the association and dissociation rates and alters its binding specificity ***(1,2)***. It is therefore advisable to always evaluate binding in both the presence and absence of Mg^{2+}. With crude protein extracts the presence of Mg^{2+} can be problematic if nucleases are present, and it may be necessary to minimize times and/or lower the temperature of incubations in order to limit degradation of the probe.
17. For proteins such as Zn finger proteins and the metal response element binding factor, addition of the appropriate metal ion (Zn^{2+} or Cd^{2+}, respectively) can improve or be necessary for binding. In some cases it may be necessary to include the divalent ion or exclude metal chelators from the electrophoresis buffer ***(3)***.
18. *Nonionic Detergents:* The addition of nonionic detergent, such as NP-40, Triton X100, or Tweens help to minimize protein aggregation and generally results in less smearing of bands and less trapping of the DNA probe in the wells. Concentrations of 0.1 or 0.2% NP-40 are normally used, but higher concentrations (up to 1 or 2%) can be used and high concentrations of a variety of nonionic detergents may enhance binding.
19. *Protein:* For crude whole cell or nuclear protein extracts it is advisable to initially assay a wide range of protein concentrations, e.g., from 2–20 µg in a 20 µL reaction. A typical nuclear extract contains 2–4 µg of protein per µL. For more purified proteins lesser amounts can be used but these will need to be titrated to determine optimal levels.
20. *Nonspecific Competitor DNA:* Both the type and the amount of competitor DNA needs to be determined to provide the best sensitivity of detection each specific complex. Competitors that have similarity to the binding site of the protein(s) being studied have the potential to interfere significantly with specific binding. For example, poly dA-dT is a poor choice of competitor for studying binding of the TATA binding protein, but very good for Sp1 (core binding site CCCGCC).
21. *Reducing Agents:* Binding of some proteins is specifically sensitive to redox potential, whereas for others, such as USF, binding is improved significantly if the DTT concentration is raised to 5 m*M*.

22. *Protease Inhibitors:* Cellular or nuclear extracts should normally be prepared in the presence of protease inhibitors (PMSF or Pefabloc® [Boehringer, Darmstadt, Germany], soybean trypsin inhibitor, leupeptin, benzamidine, etc.). Depending on the source of the protein it may prove necessary to have protease inhibitors present during binding reactions. EGTA (0.1 m*M*) can be used in the presence of Mg^{2+} to inhibit calcium-dependent proteases.

Controls for Specificity

23. *Competitors:* The most common test for specificity of complex formation is the ability of excess, unlabeled competitor DNAs to reduce the amount of complex formed with the specific labeled probe. Conditions should be chosen such that the DNA-binding protein is not saturating on the probe, in which case substantially more competitor may be required before significant decrease in complex formation with the labeled probe will be seen. Conditions under which 20–30% of the probe is bound normally provide a good starting point. Using 5 fmol of labeled probe molar excesses of 10-, 50-, and 200-fold can be used initially. Competitor DNAs may be preincubated with the protein or added at the same time as the labeled probe, but for proper kinetic comparison, the probe and competitor should be added together. It is important to remember, especially with impure protein preparations, that binding kinetics are not simple and that the final level of binding seen depends not just on the specific binding parameters of the protein with its specific DNA target, but also the binding of the protein to both specific and nonspecific competitor DNAs and to other proteins in the reaction, as well of the binding of the DNA probe to other proteins.
24. It is convenient to use at least three competitor DNAs: one corresponding directly to the binding site being studied, an equivalent DNA containing specific point mutations within the putative binding site, one unrelated DNA, and, if the site is predicted to bind a known factor, a fragment corresponding to an established binding site for the factor.
25. Either restriction fragments or annealed double-stranded oligonucleotides may be used as competitors, though it is generally easier to prepare larger quantities of oligonucleotides. If comparisons are to be made between competitors, it is important to use DNAs of similar size, as affinity often increases significantly with increasing fragment size.
26. *Mutant Probes:* probes with specific point mutations incorporated into the putative binding site can be used both to validate that a sequence region is critical for binding (e.g., by replacement of sequence with a linker) and to define critical bases for binding by using probes with selected single base changes.
27. *Antibodies:* The above methods identify the specificity characteristics of the target DNA-binding sequence. Antibodies can be used to identify the protein binding to the target sequences. If antibodies bind to the DNA-binding protein without interfering with the protein–DNA binding, complexes of larger molecular weight are formed and can be visualized following gel electrophoresis as more slowly migrating "supershifted" bands. A number of commercially available monoclonal

antibodies to common transcription factors can be used in supershift assays. Alternatively, if antibody binding interferes with formation of protein–DNA complexes, loss of binding can be seen; this is common with polyclonal antibodies. In this application, antibodies should be incubated with proteins prior to addition of DNA. Control reactions should include nonspecific serum and also a binding reaction involving an unrelated DNA–binding protein.

Electrophoretic Conditions

28. Polyacrylamide EMSA gels can be prepared using any of the buffers described in **Subheading 2.1.** (TBE should be used at 0.5 or 0.25X). It is a good idea to try them all, as resolution of complexes can differ between the systems. The low-salt TNAE buffer is optimal for minimizing dissociation of complexes, but because of its low buffering capacity, the running buffer needs to be recirculated. Not all complexes will be stable in the higher-salt buffers, but they are more convenient in that they do not require recirculation, and background nonspecific binding to the probe can be reduced.
29. To maximize complex stability, it is preferable to run gels in the cold (4°C), but for many complexes, electrophoresis at room temperature is satisfactory.
30. Agarose gels can also be used in EMSA assays, particularly for separation of large complexes such as nucleosomes or when using large DNA fragments as probes.
31. Both the concentration of acrylamide and the ratio of acrylamide/bisacrylamide can be varied to improve band resolution and separation (e.g., 4%, 80:1 for large, slowly migrating complexes).

Extended Applications

32. *Binding Constants/Kinetics:* Using purified proteins and DNA fragments it is possible to accurately determine association and dissociation rates and equilibrium binding constants using appropriately designed EMSA experiments. Papers by Meisterernst et al. ***(4)***, Carthew et al. ***(1)***, and Cann ***(5)*** provide examples and background discussion. Competing interactions (of specific probe with other proteins and protein with nonspecific competitor) limit the data that can be obtained using crude extracts. Dissociation rates can be determined, however, by following the decrease in the amount of complex following the addition of a large amount of unlabeled specific competitor DNA to a reaction containing preformed DNA protein complexes. Any formation of new complexes will be distributed between the labeled probe and the large excess of unlabeled probe, so complexes visualized in gels are those remaining at different time points after addition of the competitor.
33. *Relative Binding Affinity:* A convenient way to determine the relative affinity of a protein for two different binding sites is to incubate two different length DNA fragments containing binding sites in a single reaction ***(2)***. By choosing fragments of sufficiently different size, it is normally possible visualize complexes on each fragment as separate bands. By using one site in a reference fragment, a number of sites (including the same site as the reference site) can be readily compared in a single gel. An example of binding of a number of different sites to the

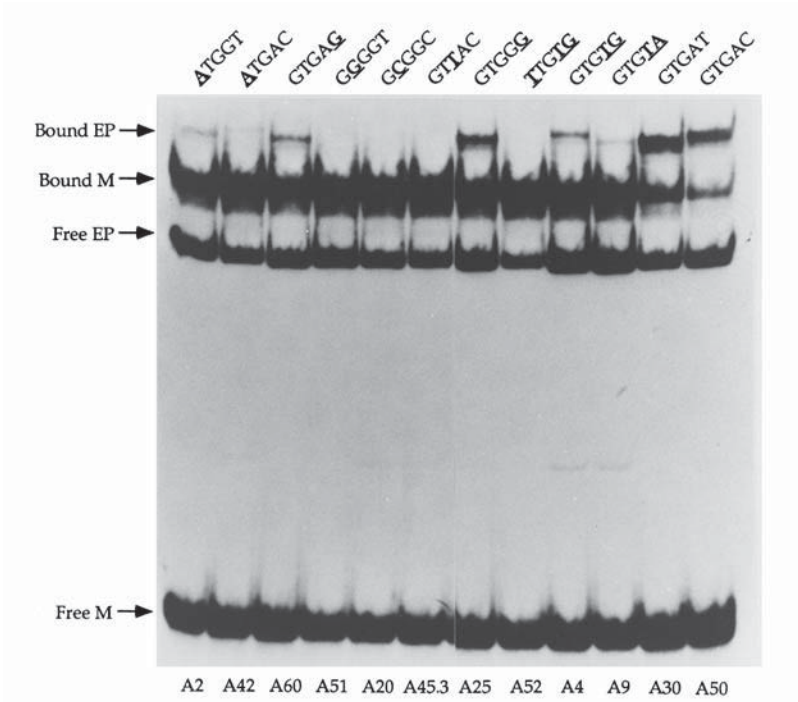


Fig. 2. Double fragment Gel Shift Assay. 9 µg of a partially purified fraction containing the transcription factor USF (from a HeLa cell nuclear extract) was used in binding reactions containing 5–10 fmol of two labeled probes and 200 ng of poly(dI-dC).poly(dI-dC). Complexes were separated on a 5% (20:1) acrylamide gel in TNAE buffer. Free M and Bd. M are the free and bound control fragment, a 50mer containing a high-affinity USF binding site. Free EP and Bd. EP are the free and bound test fragments that contain USF binding sites of different affinity on a 320 bp EcoRI - PvuII restriction fragment. Clone numbers are shown below and sequences of half sites above the lanes. Nonconsensus bases are shown in bold and underlined (Figure reprinted from ***[2]***).

USF protein is shown in **Fig. 2**. Here, the reference binding site fragment, M, is 50 bp and that containing the different binding sites, EP, is a 320 bp EcoRI to PvuII restriction fragment. Because all parameters within the reaction are equivalent reliable relative binding affinities can be determined even in crude protein extracts.

34. *Protease "Clipping":* This technique can be applied to provide evidence that bands of equivalent mobility formed on different DNA-binding sites are derived from binding to the same protein. Many proteins have regions that are more susceptible to proteases, e.g., in linker regions between protein domains, or have susceptibility to particular sequence-specific proteases. The core DNA-binding region of the protein, when bound to DNA, is often particularly protease resistant. By performing a time course digestion of protein–DNA complexes before loading on to an EMSA gel, the patterns of digestion products can be compared

for similarity. It is preferable to use a protease whose action can be stopped at the end of the incubation prior to gel running; incubation with trypsin followed by stopping the reaction with soybean trypsin inhibitor or Pefabloc (Boehringer) is an example. The calcium-dependent protease, calpain, is particularly useful as its proteolytic activity is quite limited and digestion is readily stopped by addition of the chelator EGTA ***(6)***.

35. *Binding Stochiometry:* Mobility shift analysis can be used to study the binding stochiometry of proteins. This is most conveniently done by preparing a truncated form of the protein (either by controlled limited proteolysis or by genetic manipulation). Incubations containing the normal protein alone, the truncated protein alone, or mixtures of the two are incubated with DNA probes. For a protein that binds as a dimer, a complex of intermediate mobility would be expected only for the mixture of the two different length proteins.
36. *Mol Wt Determination:* While there is generally an approximate relationship between fragment mobility and protein size, the effects of different protein charge and shape make it difficult directly to make a reliable estimate. By performing EMSAs in gels of different percent polyacrylamide, Orchard and May ***(7)*** have shown reliable molecular weight estimates can be made.
37. *Site Selection:* Using pools of oligonucleotides containing a region of random sequence EMSA can be efficiently used to select from the population those molecules bound to a particular protein. By using successive rounds of PCR and EMSA, enriched populations of DNA binding sites can be isolated and studied ***(2,8)***.

References

1. Chodosh, L. A., Carthew, R. W., and Sharp. P. A. (1986) A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* **6**, 4723–4733.
2. Bendall, A. J. and Molloy, P. L. (1994) Base preferences for binding by the bHLH-Zip protein USF: effects of $MgCl_2$ on specificity and comparison with binding of Myc family members. *Nucleic Acids Res.* **22**, 2801–2810.
3. Anderson, R. D., Taplitz, S. J., Oberbauer, A. M., Calame, K. L., and Herschman, R. (1990) Metal-dependent binding of a nuclear factor to the rat metallothionein-I promoter. *Nucleic Acids Res.* **18**, 6049–6055.
4. Meisterernst, M., Gander, I., Rogge, L., and Winnacker, E-L. (1988) A quantitative analysis of nuclear factor I/DNA interactions. *Nucleic Acids Res.* **16**, 4419–4435.
5. Cann, J. R. (1989) Phenomenological theory of gel electrophoresis of protein-nucleic acid complexes. *J. Biol. Chem.* **264**, 17,032–17,040.
6. Watt, F. and Molloy, P. L. (1993) Specific cleavage of transcription factors by the thiol protease, m-calpain. *Nucleic Acids Res.* **21**, 5092–5100.
7. Orchard, K. and May, G. E. (1993) An EMSA-based method for determining the molecular weight of a protein-DNA complex. *Nucleic Acids Res.* **21**, 3335–3336.
8. Oliphant, A. R., Brandl, C. J., and Struhl, K. (1989) Defining the sequence-specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein. *Mol. Cell. Biol.* **9**, 2944–2949.

20

In Vitro Promoter Analysis Using Nuclear Extracts and G-Free Cassette Vectors

Martin J. Tymms

1. Introduction

Nuclear extracts prepared by appropriate methods contain all of the components for transcription ***(1)***. Although nuclear extracts are not defined like transcription systems utilizing purified transcription components ***(2–4)*** the extracts can be used to study aspects of promoter regulation. If suitable nuclear extracts can be prepared that contain the necessary transcription factors in an active state, and suitable DNA templates containing regulator sequences are available, processes such as differential gene expression ***(5,6)*** and hormonal simulation of transcription can be reconstituted ***(7,8)***.

In many cases, however, faithful in vitro transcription cannot be reconstituted. A serious problem with in vitro transcriptions that is not easily resolved is the absence of the structure which is found in vivo. The DNA in a eukaryotic nucleus is packaged into a nucleosome array, and the precise positioning of nucleosomes along the promoter region is critical for appropriate regulation of gene expression ***(9–11)***. Some inducible genes are accessible to transcription factors without chromatin remodeling, but others require remodeling before DNA is accessible ***(12)***.

In vitro transcription reactions using nuclear extracts are particularly useful for studying the role of specific transcription factors in promoter regulation. Antibodies to transcription factors can be used to dissect the role of specific transcription factors in in vitro transcription driven by nuclear extracts. For example, an antibody that binds to a transcription factor and prevents binding to DNA can be used to assess the role of that factor in the overall transcriptional activity. Alternatively, specific antibodies can be used to remove transcription factors from nuclear extracts prior to the transcription reaction ***(13)***.

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms © 2000 Humana Press Inc., Totowa, NJ

If antibodies are not available, alternative strategies can be used. Because transcription factors bind in a sequence-specific manner, the addition of DNA sequences that will bind a factor results in competition between the transcription template and the added DNA, which usually consists of a molar excess of short double-stranded oligonucleotides 18–30 base pairs in length. This type of analysis can establish the importance of the promoter element by effectively stopping the protein(s) binding the element from interacting with the transcriptional template (*see* Chapter 21, this volume).

In vitro transcription systems allow the addition of a wide range of exogenous proteins to study transcriptional regulation. Recombinant transcription factors or purified factors can be added to in vitro transcription reactions. This approach can be useful when the nuclear extracts have been depleted of endogenous transcription factors. For example, a nuclear extract depleted of AP1 transcription factors can be used to study the transcriptional activation of different combinations of recombinant AP1 transcription factors ***(14)***. This approach, furthermore, can be used to study the activity of modified proteins and explore structure function relationships ***(15,16)***.

In addition to manipulations in the protein components involved in regulating transcription, the DNA template can also be modified ***(17)***. This can include testing different promoter deletions or making smaller deletions or mutations. This analysis can be used to establish the importance of regulatory elements.

A number of different protocols have been developed for quantitation of in vitro transcription reactions with nuclear extracts. The transcription reactions can be carried out with radiolabeled nucleotides and the products of transcription directly detected, or the reaction can be carried out with "cold" nucleotides and the subsequent detection of specific transcripts performed in using specific protocols employing radioisotope. The RNase protection ***(18)*** and S1 nuclease protection protocols ***(19)*** use radiolabeled antisense RNA and DNA probes, respectively, to detect specific RNA transcripts generated from the transcription template. Another protocol commonly employed is primer extension analysis in which a transcript-specific DNA primer and reverse transcriptase are used to make a radiolabeled cDNA copy of transcripts up to the site of initiation ***(20)***.

In vitro transcription can also be quantitated via the direct incorporation of radiolabeled ribonucleotide ***(21)***. One problem with this approach is that DNA contamination in nuclear extract can cause spurious transcription resulting in significant background incorporation, which can mask specific transcription products. One of the best techniques for detecting specific transcription in crude transcription systems is the "G-free" cassette method ***(7,22)***. The G-free technique allows the detection of specific transcripts directly in the transcription reaction without a separate detection step. The key to this ingenious method is

the use of special plasmid vectors with "reporter" DNA sequences that do not contain G residues in the mRNA strand. Promoter elements are cloned into these vectors with no G residues between the start of transcription and the G-free cassette (a synthetic DNA sequence that contains no G residues in the noncoding strand). G-free vectors available are shown in **Fig. 1**. The technique relies on three modifications to normal transcription reactions to significantly reduce non-specific transcripts:

1. No GTP is added to the reaction mix: correctly initiated transcripts from the added plasmid template will not contain any G residues.
2. GTP analog 3'-*O*-methyl GTP is added: this analog when incorporated in the place of GTP causes termination of transcription. Because the correctly initiated transcripts do not have G residues, the analog will only terminate spurious transcripts.
3. As the final coup ribonuclease T_1 is added to the reaction: ribonuclease T_1 cleaves RNA adjacent to G nucleotides and will degrade RNA which contains G residues. As the transcripts from the G-less vector template do not contain G residues, they are not degraded by ribonuclease T1. The combination of all three of these measures results in clear transcription signals in crude nuclear extracts containing high levels of nuclear DNA contamination. This chapter details a basic protocol for the use of the G-free cassette system with crude nuclear extract.

2. Materials

High-quality deionized water such as that produced with a Millipore Milli-Q purification system or other system giving high-quality water should be used for making up all solutions.

1. DEPC water: add 0.1%(v/v) diethylpyrocarbonate to Milli-Q water and mix for 1 h and autoclave.
2. PBS: 137 m*M* NaCl, 2.7 m*M* KCl, 9.5 m*M* sodium phosphate, pH 7.3.
3. Buffer A: 20 m*M* Tris-HCl, pH 7.9, 2 m*M* DTT, 1 m*M* EDTA, and 10%(v/v) glycerol.
4. Buffer C: 20 m*M* Tris-HCl, pH 7.9, 600 m*M* NaCl, 2 m*M* DTT, 1.5 m*M* $MgCl_2$, 0.2 m*M* EDTA, and 20%(v/v) glycerol.
5. Buffer D: 20 m*M* Tris-HCl, pH 7.9, 20 m*M* NaCl, 2 m*M* DTT, 5 m*M* $MgCl_2$, 0.2 m*M* EDTA, and 20%(v/v) glycerol.
6. TE: 10 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA. Autoclave.
7. 1 *M* NaCl: autoclave.
8. 1 *M* HEPES pH 7.9.
9. 100 m*M* $MgCl_2$.
10. 100 m*M* ATP: high purity, (Pharmacia, Uppsala, Sweden).
11. 50 m*M* 3'-*O*-methyl-GTP: 3'-O-methylguanosine 5''-triphosphate(sodium salt) (Pharmacia) made up DEPC-treated water and frozen in aliquots at –20°C.
12. 100 m*M* CTP: high purity (Pharmacia).
13. 2.5 m*M* UTP: dilute 100 m*M* stock (Pharmacia) with DEPC-treated water.

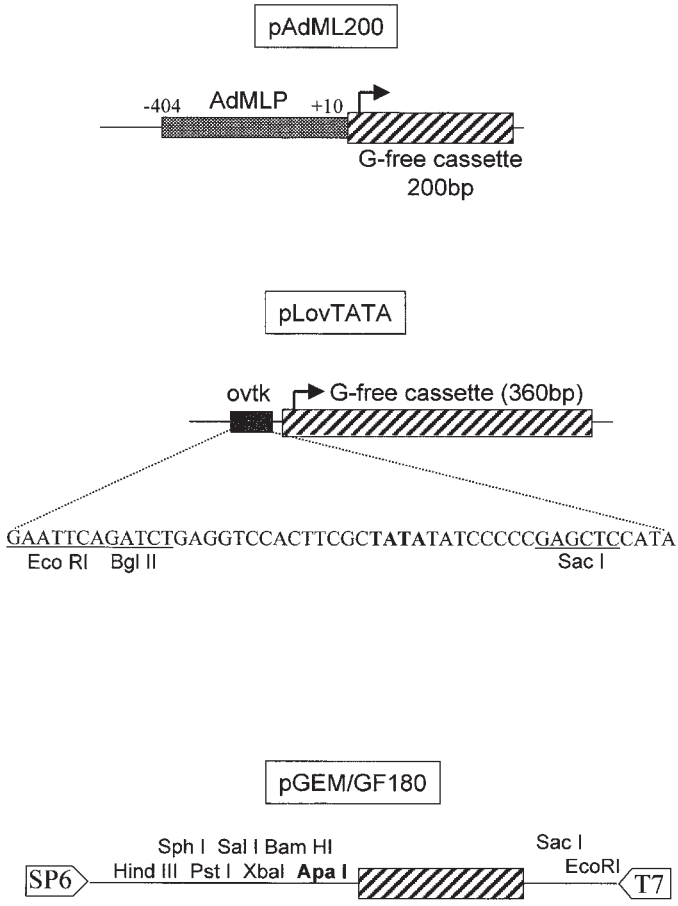


Fig. 1. G-free reporter plasmids. **pAdML200**: this control plasmid gives a 190 bp transcript from correct initiation and a 200 bp transcript from incorrect initiation(s) upstream. This plasmid is a derivative of pMLC$_2$AT19 ***(22)*** containing the adenovirus major-late promoter(AdML) upstream of a 200-bp G-free cassette ***(23)***. **pLovTATA**: this vector contains a synthetic minimal promoter derived from the HSV thymidine kinase promoter and the chicken ovalbumin promoter (ovtk) ***(23)*** cloned into the vector pC$_2$AT19 which based on pUC13 ***(1)***. Elements such as the glucocorticord and progesterone response elements have been cloned into the Bgl II site to give a construct showing progesterone responsiveness in vitro ***(7,23)***. **pGEM/GF180** (Tymms, M., unpublished)**:** This plasmid is derived from pGEM3Zf(+) (Promega Corp, Madison,WI) and contains a 180 bp G-free cassette, but no promoter sequences. Promoter fragments can be cloned into the polylinker provided that transcriptional initiation does not occur before the ApaI site. The vector can be digested with ApaI and made blunt with T4 DNA polymerase to accept blunt fragments without G residues upstream of the G-free cassette. The vector is available from the author.

14. Transcription stop solution: 25 m*M* Tris-HCl, pH 7.5, 0.5% SDS, 10 m*M* EDTA, 400 µg/mL proteinase K (Boehringer Mannheim, Darmstadt, Germany), 200 µg/mL yeast RNA. Make up fresh from stock solutions.
15. Urea solution: 10 m*M* Tris-HCl, pH 7.5, 1 m*M* EDTA, 8 *M* urea. Make up fresh and filter with a 0.2 µm filter.
16. 3 *M* sodium acetate, pH 5.3.
17. Chloroform/isoamyl alcohol: Analytical grade chloroform containing 4%(v/v) isoamyl alcohol.
18. 10X TBE: 0.89 *M* Tris-base, 0.89 *M* boric acid and 20 m*M* EDTA.
19. 100% ethanol: analytical grade ethanol.
20. 80% ethanol: 80%(v/v) analytical grade ethanol in DEPC water.
21. Acrylamide stock: 29%(w/v) acrylamide, 1%(w/v) bisacrylamide in water.
22. 6% gel mix: dissolve 480 g of high-purity urea in 200 mL acrylamide stock, 100 mL 10X TBE and adjust the volume to 1 L with water. Gel mix can be stored for 1–2 mo at 4°C.
23. DTT: 100 m*M* dithiothreitol, 1 m*M* EDTA.
24. Gel loading solution: 80%(v/v) formamide, 2 m*M* EDTA, 0.1%(w/v) xylene cyanol, 0.1% bromophenol blue.
25. Ribonuclease T_1: ribonuclease T_1 (Boehringer 109 207) at 100,000 U/mL
26. Yeast RNA: 5 mg/mL in water (*see* **Note 1**).
27. Phenol: melt phenol at 60°C and add an equal volume of 50 m*M* Tris-HCl, pH 9.0. Vigorously mix and allow the phases to separate and collect the lower phase. To the buffer-saturated phenol, add 0.1%(w/v) hydroxy quinoline and stir to dissolve. Aliquot into 10 mL portions at store at –20°C.
28. Radiolabeled-UTP: [α-^{32}P] or preferably [α-^{33}P]-labeled UTP can be used.
29. Dialysis tubing.
30. Dounce homogenizer: size 7 with a B-pestle (Wheaton,NJ).
31. Sequencing gel apparatus and power supply.

3. Methods

3.1. Preparation of Nuclear Extract (*see* Note 2)

1. Harvest cells from 10–15 near-confluent flasks (each 150 cm^2). Approx 4×10^8 cells.
2. Wash with PBS and collect the cells in a 50 mL polypropylene tube on ice and pellet cells by low-speed centrifugation in a bench-top centrifuge. Measure the packed cell volume (PCV).
3. Add 5X PCV of Buffer A and resuspend the cells in buffer by gently inverting the tube several times and incubate on ice for 10 min.
4. Transfer the cell suspension to a polypropylene centrifuge tube and centrifuge at 3000*g* for 10 min at 4°C. Discard the supernatant.
5. Add twice the packed cell volume of Buffer A. Mix gently and transfer to an ice-cold Dounce homogenizer. Perform homogenization at 4°C (preferably in a cold room) with at least 20 gentle strokes of the pestle being careful not to create frothing.

6. Transfer the homogenate to a precooled 50 mL polypropylene centrifuge tube and centrifuge at 3000*g* for 10 min at 4°C.
7. Remove the supernatant and add 2/3 PCV of Buffer C to the pellet. Transfer to the Dounce homogenizer and homogenize with 20 strokes of the pestle (*see* **Note 3**).
8. Transfer the homogenate to a 50 mL polypropylene centrifuge tube, incubate on ice for 30 min with swirling every 5 min.
9. Centrifuge for 20 min at 25,000*g* at 4°C.
10. Collect the supernatant, which is the nuclear extract, in an ice-cold, sterile polypropylene tube. Discard the pellet.
11. Dialyze the nuclear extract in a prewashed dialysis bag against 1 L of ice-cold Buffer D overnight with one change of buffer after 4 h.
12. Collect the dialyzed sample from the dialysis bag. Centrifuge at 25,000*g* for 15 min at 4°C to remove denatured proteins.
13. Flash-freeze 100 µL aliquots of the extract in 1.5 mL microtubes with liquid nitrogen or dry ice and store at –70°C.
14. Measure protein concentration by the method of Bradford ***(24)*** or another appropriate method. The protein concentration should be in the range of 3–5 mg per mL.

3.2. *In Vitro Transcription* (*see* **Note 4**)

1. To each tube add: 2 µL DNA mixture containing 100–200 ng of test template and 25–50 ng control template (*see* **Note 5**).
2. Add 2–10 µL nuclear extract (approx 3–5 mg/mL protein) and 5–13 µL Buffer D to a final volume of 15 µL.
3. Incubate 15 min at 20°C.
4. Make a master transcription mix by assembly of the following components which are sufficient for 10 reactions: 3 µL 1 *M* NaCl; 2.6 µL, 1 *M* HEPES pH 7.9, 7.5 µL 100 m*M* $MgCl_2$; 1.25 µL of 100 m*M* ATP; 1.25 µL of 100 m*M* CTP ; 1 µL of 2.5 m*M* UTP (*see* **Note 6**); 5 µL of 50 m*M* 3'-0-methyl GTP (*see* **Note 7**); 1 µL of 100 U/µL ribonuclease T_1 (*see* **Note 8**); 5 µL 100 m*M* DTT (*see* **Note 9**); 3 µL of ^{33}P-UTP (*see* **Note 10**) and water to a total of 80 µL.
5. Add 8 µL of Master Mix to each reaction, mixing by pipeting and incubate 45–60 min at 30°C (*see* **Note 11**).
6. Add 100 µL transcription stop solution, vortex, and incubate at 20°C for 10 min.
7. Add 200 µL urea solution, vortex, and incubate at 20°C for 10 min.
8. Add 160 µL phenol and 160 µL choroform/isoamylalcohol to each tube, tightly cap, and vortex for 20 s.
9. Centrifuge at 10,000*g* for 10 min and remove the aqueous supernatant to a new tube. **Note:** All waste produced from this and subsequent extractions are radioactive.
10. Repeat the phenol–chloroform extraction (*see* **steps 8–10**).
11. To each tube add 30 µL of 3 *M* sodium acetate pH 5.3, 10 µg yeast RNA and 900 µL 100% ethanol, vortex, and store at –70°C for 1 h.
12. Centrifuge tubes at 10,000g for 10 min at 4°C and carefully remove most of the supernatant without disturbing the RNA pellet.

13. Add 400 µL of ice-cold 80% ethanol, centrifuge at 10,000*g* for 1 min and remove all ethanol from the pellet and allow air dry for 15–20 min.
14. Resuspend the RNA pellet in 5 µL gel loading solution.
15. Heat samples at 90°C before loading a 6% polyacrylamide sequencing gel (*see* **Note 12**).
16. Run a radiolabeled size standard on the gel to verify the size of transcription product (*see* **Note13**).
17. Treat the gel with gel fixative and dry onto 3 MM paper.
18. Expose the dried gel to X-ray film or preferably use a phosphoimager such as a Fuji Bioimager. **Figure 1** looks at optimization of signal using the G-less cassette system.

4. Notes

1. Any RNA can be used as a carrier, provided it is free of contaminating ribonuclease. Commercially available yeast RNA from Boehringer can be used after cleaning up by phenol–chloroform extraction.
2. There are a large number of methods described for preparing nuclear extracts based on the method originally decribed by Dignam ***(1)***. The method detailed here is based on Bagchi ***(7)***.
3. If a conductance meter is available, the effective concentration of sodium chloride can be measured. It should be between 350–400 m*M*.
4. The basic transcription protocol described should be used as a guide to optimize the reaction for the template and nuclear extract used. The conditions will need to be optimized for each nuclear extract prepared. The use of a control template is optional, but allows correction for sample loading variations.
5. The amount of test and control template in reactions will depend upon the activity of the templates in the nuclear extract used. The control template concentrations should be adjusted to give a signal of similar or lower intensity than the test tempate.
6. The quantity of UTP in the reaction can be reduced to increase the signal intensity, but the level of premature transcriptional terminations may also increase.
7. The level of 3'-*O*-methyl GTP required in the reaction to suppress spurious transcriptional initiation will depend upon the amount of nuclear extract used and the purity of the extract. *See* **Fig. 2A** for an illustration of the effect of titration of 3'-*O*-methyl GTP.
8. The amount of ribonuclease T_1 added to the reaction to give the cleanest signal need to be empirically determined. If the nuclear extract is significantly contaminated with chromosomal DNA and contains high exogenous levels of GTP, the addition of higher levels of ribonuclease T_1 may be more effective than adding additional 3'-*O*-methyl GTP, which is expensive.
9. If nuclear extracts are supplemented with purified recombinant transcription factors such as Fos and Jun, a DTT level of 4–5 m*M* has been found to improve transactivation.
10. ^{32}P-UTP can also be used in these reactions, but ^{33}P-UTP is recommended because of its lower hazard level in the complex manipulations involved in preparing samples for electrophoresis.

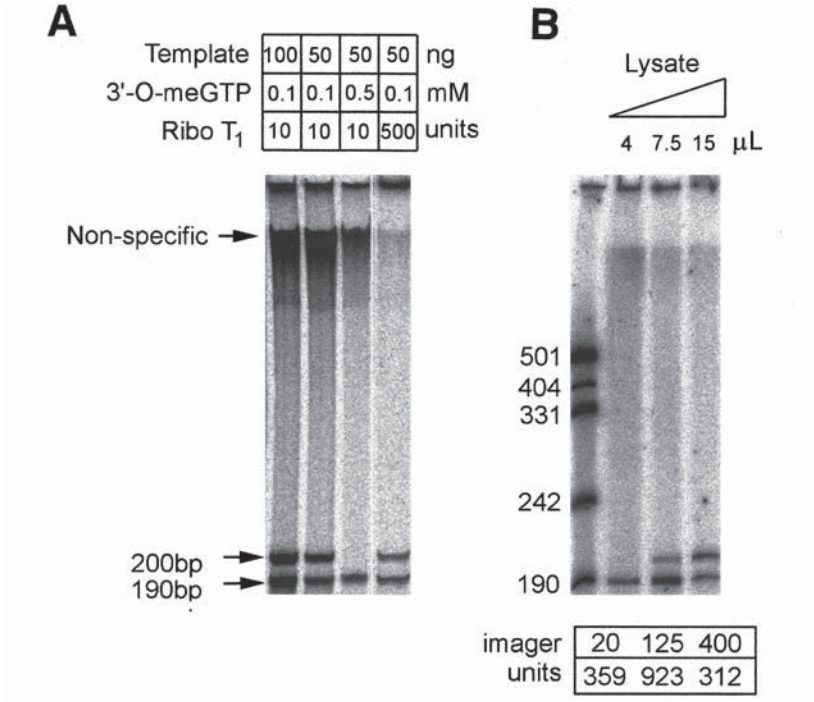


Fig. 2. Optimization of transcription using pAdML200. **(A)** The effect of ribonuclease T_1 and 3'-*O*-MeGTP. Transcription reactions using 7.5 µL of HeLa and the indicated levels of template, ribonuclease T_1, and 3'-*O*-MeGTP. Note that increases in the levels of both 3'-*O*-MeGTP and ribonuclease T_1 reduce the level of nonspecific signal, which arises from genomic DNA contamination of the nuclear extract. Correctly initiated transcripts are 190 bp, whereas incorrect upstream initiations are digested by ribonuclease T_1 to give a 200 bp transcript ***(23)***. **(B)** The effect of increasing levels of nuclear extract. Transcription reactions using 50 ng pAdML200, 1 m*M* 3'-*O*-MeGTP and 50 U ribonuclease T_1 and the indicated volumes of nuclear extract from Jurkat T cells. The quantitation of signals (in arbitary units) for both the correctly and incorrectly initiated transcripts are indicated below each lane. Note that although all transcripts are correctly initiated with the lowest level of extract, maximal transcription is achieved with a higher concentration of extract.

11. Both the incubation temperature and incubation time may be optimized.
12. Any standard sequencing apparatus is satisfactory. It is generally unnecessary to run the gels longer than necessary to move the bromphenol tracking dye 20–25 cm.
13. Dephosphorylated DNA markers which are ready to end-label with [γ-^{32}P]-ATP are available from a number of commercial suppliers.

References

1. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **1**,1475–1489
2. Tyree, C. M., George, C. P., Lira-DeVito, L. M., Wampler, S. L., Dahmus, M. E., Zawel, L., and Kadonaga, J. T. (1993) Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.* **7**, 1254–1265.
3. Goodrich, J. A. and Tjian, R. (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **77**, 145–156.
4. Hansen, S. K., Takada, S., Jacobson, R. H., Lis, J. T., and Tjian, R. (1997) Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* **91**, 71–83.
5. Flanagan, W. M. and Crabtree, G.R. (1992) In vitro transcription faithfully reflecting T- cell activation requirements. *J. Biol. Chem.* **267**, 399–406.
6. Gorski, K, Carneiro, M., and Schibler U. (1986) Tissue-specific in vitro transcription from mouse albumin promoter. *Cell* **47**, 767–776
7. Bagchi, M. K., Tsai, S. Y., and Tsai, M.-J. (1995) In vitro reconstitution of progesterone-dependent RNA transcription in nuclear extracts of human breast carcinoma cells, in *In Vitro Transcription and Translation Protocols* (Tymms, M. J., ed.), Humana, Totowa, NJ, pp. 107–120.
8. Fondell, J. D., Brunel, F., Hisatake, K., and Roeder, R. G. (1996) Unliganded thyroid hormone receptor alpha can target TATA- binding protein for transcriptional repression. *Mol. Cell Biol.* **16**, 281-287.
9. Wallrath, L. L., Lu, Q., Granok, H., and Elgin, S. C. R. (1994) Architectural variations of inducible eukaryotic promoters:preset and remodeling chromatin structures. *Bioessays* **16**, 165–170.
10. Razin, A. (1998) CpG methylation, chromatin structure and gene silencing-a three-way connection. *EMBO J.* **17**, 4905–4908.
11. Kadonaga, J. T. (1998) Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92**, 307–313.
12. Archer, T. K., Cordingley, M. G., Wolford, R. G., and Hager, G. L. (1991) Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. *Mol. Cell Biol.* **11**, 688–698.
13. Chang, L. A., Smith, T., Pognonec, P., Roeder, R. G., and Murialdo, H. (1992) Identification of USF as the ubiquitous murine factor that binds to and stimulates transcription from the immunoglobulin lambda 2- chain promoter. *Nucleic. Acids. Res.* **20**, 287–293.
14. Abate, C., Luk, D., and Curran, T. (1991) Transcriptional regulation by Fos and Jun In Vitro: interaction among multiple activator and regulatory domains. *Mol. Cell. Biol.* **11**, 3624–3632.
15. Goldberg, Y., Treier, M., Ghysdael, J., and Bohmann, D. (1994) Repression of AP-1-stimulated transcription by c-Ets-1. *J. Biol. Chem.* **269**, 16,566–16,576.

16. Song, C. Z., Tierney, C. J., Loewenstein, P. M., Pusztai, R., Symington, J. S., Tang, Q. Q., Toth, K., Nishikawa, A., Bayley, S. T., and Green, M. (1995) Transcriptional repression by human adenovirus E1A N terminus/conserved domain 1 polypeptides in vivo and in vitro in the absence of protein synthesis. *J. Biol. Chem.* **270**, 23,263–23,267.
17. Yu, M., Yang, X. Y., Schmidt, T., Chinenov, Y., Wang, R., and Martin, M. E. (1997) GA-binding protein-dependent transcription initiator elements. Effect of helical spacing between polyomavirus enhancer a factor 3(PEA3)/Ets-binding sites on initiator activity. *J. Biol. Chem.* **272**, 29,060–29,067.
18. Tymms, M. J. (1995) Quantitative measurement of mRNA using the RNase protection assay, in *In Vitro Transcription and Translation Protocols* (Tymms, M. J., ed.) Humana, Totowa, NJ, pp. 31–46.
19. Viville, S. and Mantovani, R. (1994) S1 mapping using single-stranded DNA probes, in *Protocols for Gene Analysis* (Harwood, A. J., ed.) Humana, Totowa, NJ, pp. 299–305.
20. Simpson, C. G. and Brown, J. W. (1995) Primer extension assay, in *Plant Gene Transfer and Expression Protocols* (Jones, H., ed.) Humana, Totowa, NJ, pp. 249–256.
21. Pugh, B. F. (1995) Preparation of HeLa nuclear extracts, in *In Vitro Transcription and Translation Protocols* (Tymms, M. J., ed.) Humana, Totowa, NJ, pp. 349–357.
22. Sawadogo, M. and Roeder, R. G. (1985) Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci. USA* **82**, 4394–4398.
23. Klein-Hitpass, L., Tsai, S. Y., Weigel, N. L., Allan, G. F., Riley, D., Rodriguez, R., Schrader, W. T., Tsai, M. J., and O'Malley, B. W. (1990) The progesterone receptor stimulates cell-free transcription by enhancing the formation of a stable preinitiation complex. *Cell* **60**, 247–257.
24. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

21

In Vitro Transcription Using Competitor Oligonucleotides to Deplete Specific Transcription Factors

Fujiko Watt

1. Introduction

In vitro transcription provides information not necessarily available from transfection assays about the regulatory activity of a promoter transcribed by RNA polymerase II. In the particular application discussed here, when an oligonucleotide duplex containing a recognition site for a specific transcription factor is included in the transcription reaction using nuclear extract, the factor can be selectively depleted from a pool of binding factors that are required to regulate transcription from the class II promoter.

When promoter analysis is conducted by transient expression assays, a number of mutated promoter constructs must be prepared for comparison. Sometimes it is technically difficult and too time-consuming to prepare a particular mutated construct. In addition, the use of competitor oligonucleotides in transient assays requires a large amount of a competitor, enough to compete with the transfected reporter-gene. However, it remains difficult to control the efficiency and uptake of oligonucleotides into the nucleus. These difficulties prompted me to develop a technique that uses specific oligonucleotides in transcription assays in vitro.

With in vitro transcription, parameters such as incubation temperature and the concentrations of DNA, protein and salt can be independently altered. Thus, in this defined environment, it is possible to quantitatively and qualitatively analyze how interactions between nuclear proteins and the DNA template regulate the promoter activity. In outline, the method described in this chapter uses double-stranded competitor oligonucleotides in transcription assays in vitro, using nuclear extract that provides the components required to form initiation

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

complex, i.e., RNA polymerase II and general transcription factors. Competitor oligonucleotides are designed either to contain the cognate sequence of a known transcription factor or the sequence encompassing the regulatory region. To deplete a specific DNA-binding protein before it binds to its cognate site on the promoter and recruits the components to form initiation complex, competitor oligonucleotide is added to the nuclear extract before adding template DNA. The effect of depletion of a specific factor is then monitored by examining the RNA products, which are converted to cDNA by reverse transcription using a specific primer **(Fig. 1)**. The method facilitates screening for the hierarchy of the regulatory elements, which are essential for, or which contribute to, the transcription from the promoter of interest. In addition, where multiple initiation sites are present in a promoter, the relationship between a set of regulatory elements and an individual initiation site nearby can be determined, i.e., the amount of cDNA as a measure of the level of specific transcription, and the size of it to identify a specific transcription start site **(Fig. 2)**. The competitor oligonucelotides, if so designed, can be used for analyses that complement the data obtained by in vitro transcription. For instance, they can be used to characterize binding factor(s) by EMSA.

The method described here uses supercoiled templates and detection of RNA products by reverse transcription. It could also be done using a linearized template and by a run-off transcription with the label incorporation during the reaction.

2. Materials

1. Template DNA: Supercoiled plasmid DNA serves as a more efficient template and should also be very clean, i.e., prepared by double CsCl gradient purification (*see* **Note 1**) or one CsCl purification followed by proteinase K digestion and phenol-chloroform extraction ***(1)***. When multiple templates are used, the DNA concentration of each template should be carefully adjusted so that equimolar amounts are compared for promoter activity. This is particularly important when mutated templates are included. The most accurate way of achieving this is to use dot-blot hybridization ***(1)***, together with OD measurement at 260 nm and visualization of linearized plasmid DNAs on an agarose gel. The concentration of the template is typically adjusted to 100–200 ng/µL (~20–100 fmol/µL) in TE.
2. Competitor oligonucleotides: These contain transcription factor(s) binding sites and are annealed to form double-stranded DNAs and end-filled to produce flush ends to prevent nonspecific binding of transcription factors to the sticky ends. Competitor concentrations are adjusted to 5 pmol/µL.
3. Primer extension oligonucleotide: The oligonucleotide for use in primer extension should be of discrete size, prepared by purification on denaturing polyacrylamide gel or by equivalent procedure.
4. Buffer A-0.1: 20 m*M* Hepes-KOH, pH 7.9, 1 m*M* EDTA, 20% glycerol, 0.1 *M* KCl, 4 mM DTT containing 0.5 m*M* PMSF or 0.4 m*M* of Pefabloc (Boehringer).

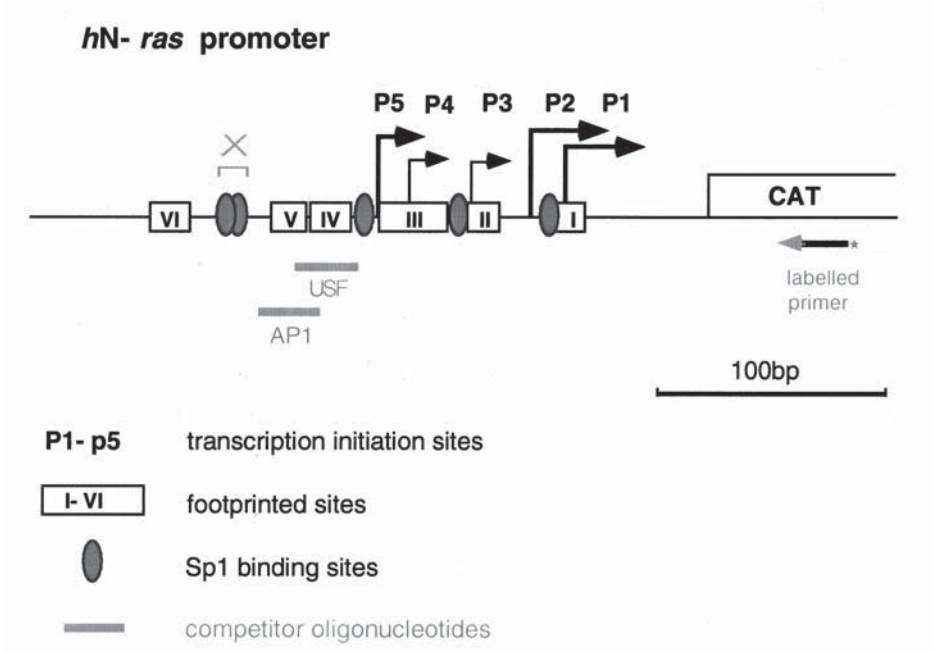


Fig. 1. Schematic diagram of a DNA template and competitor oligonucleotides. In this instance, the template DNA was made by inserting a 0.4-kb human N-ras promoter in front of the CAT-reporter gene (pN-ras-CAT). The human N-ras gene promoter is a TATA-less promoter, in which transcription starts from multiple sites ***(6)***. The five main transcription initiation sites detectable by in vitro transcription analysis are depicted as P1–P5 ***(7)***. Two oligonucleotides containing binding site for AP1, and USF, respectively, were used to compete out the binding of transcription factors, AP1 and USF, to their regulatory elements (footprinted sites, fp-V and -IV) on the promoter ***(8,9)***. The transcripts were analyzed by extension of a 25-nucleotide primer complementary to the sequence within the CAT-gene.

5. Buffer A-0 : Buffer A-0 is the same as Buffer A-0.1 except that it lacks KCl.
6. Ammonium sulfate.
7. Creatine mix: 0.5 µL of 0.2 *M* creatine phosphate, 0.5 µL 20 mg/µL (or 1 U) creatine kinase, 0.5 µL of 200 m*M* $MgCl_2$, and 0.5 µL (1 U/µL) RNAsin (Promega, Madison, WI) per 20 µL reaction. This should be made fresh for each experiment (*see* **Note 3**).
8. 8 m*M* rNTP mix: 8 m*M* each ribonucleotide.
9. Stop mix: 250 µg/mL Proteinase K, 0.25% SDS, 25 m*M* Tris-HCl, pH 7.5, 10 m*M* EDTA, and 50 µg/mL tRNA (see **Note 3**).
10. Annealing buffer: 10 m*M* Tris-HCl, pH 7.9, 1 m*M* EDTA, 1.25 *M* KCl.
11. 2X primer extension buffer: 40 m*M* Tris-HCl, pH 8.7, 20 m*M* $MgCl_2$, 10 m*M* DTT.

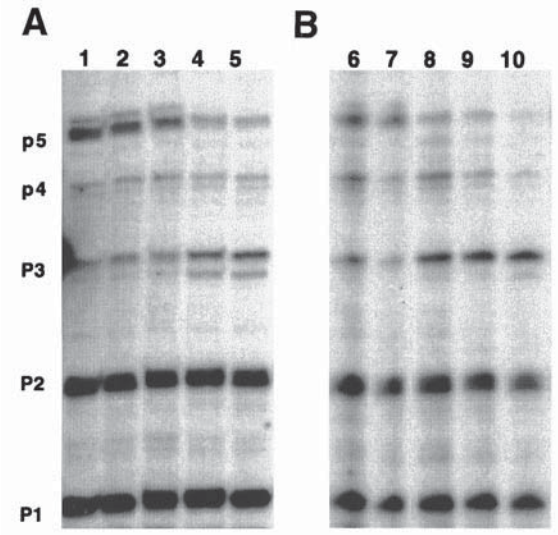


Fig. 2. Competitor oligonucleotides change the level of transcription from a nearby transcription initiation site. **(A)** pN-ras-CAT plasmid, 100 ng (37 fmol), was used as template and the transcripts were detected by reverse transcription using the CAT-primer (20 fmol). The AP1 oligonucleotide in 40- and 100-fold excess (1.5 pmol and 3.8 pmol in lanes 4 and 5) reduced the level of transcription from nearby site P5, and increased the level from P3, whereas the USF oligonucleotide in 40- and 100-fold excess (lanes 2 and 3) hardly altered the signal compared to control template without competitor (lane 1). Note that the signal from P1, P2, and P4 remained virtually the same. **(B)** The effect of Sp1 was studied using pN-ras-CAT (Sp1mut) in which the double Sp1 sites immediately upstream of fp-V were mutated. Compared to the wild-type template, pN-ras-CAT (200 and 100 ng in lanes 6 and 7, respectively), the signal from P5 was reduced from pN-ras-CAT (Sp1mut) (200 ng in lane 6 and 100 ng in lanes 7–10, respectively). Addition of the AP1 oligonucleotide in 40-fold excess (lane 10) further reduced the signals from P5 and P4, without altering the increased level of signal from the P3 site.

12. 10X kinase buffer: 500 m*M* Tris-HCl, pH 7.5, 100 m*M* $MgCl_2$, 50m*M* dithiothreitol, and 1 m*M* spermidine.
13. T4 polynucleotide kinase: 10 U/µL.
14. [γ-^{32}P] ATP: 5000 Ci/mmol, 10 mCi/mL.
15. Reverse transcription mix: prepare freshly each time by mixing 12.5 µL 2X Primer Extension Buffer, 9 µL 1 m*M* dNTPs, 0.5 U RNAsin (Promega) and 3 U AMV reverse transcriptase (Boehringer Mannheim) and H_2O to a final volume of 25 µL.
16. Phenol-chloroform (1:1).
17. TE: 10 m*M* Tris-HCl, pH 8.8, 1 m*M* EDTA.
18. 3 *M* NaOAc pH 5.3.
19. 100% ethanol.

20. 70% ethanol.
21. 5–6% sequencing gel.
22. Formamide/dye: 80%(v/v) formamide, 0.1%(w/v) xylene cyanol, 0.1%(w/v) bromophenol blue.
23. 10% acetic acid.
24. Sequencing apparatus and power supply.

3. Methods

3.1. Concentration of Nuclear Extract

High-quality nuclear extracts are required such as described originally by Dignam et al. ***(2)*** and modified by Prywes and Roeder ***(3)*** The buffers described can be made using HEPES-KOH, pH 7.9, instead of Tris-HCl, and a cocktail of protease inhibitors (Boehringer) is added to the hypotonic and nuclear extraction buffers. Typical preparation require on the order of 10^9 cells. Nuclear extracts can be concentrated step by ammonium sulfate precipitation decribed as follows ***(4)***.

1. Add ammonium sulfate (0.33g/mL) to the nuclear extract and rock the sample for 20 min after the ammonium sulfate is dissolved.
2. Collect the precipitate by centrifugation 85,000*g* for 20 min at 2°C,
3. Resuspended the pellet in 1.0 mL/10^9 cells of Buffer A-0.1 and dialyze twice for 90 min against >200 vol of Buffer A-0.1 at 4°C.
4. Measure the protein concentration. Typical preparations have a concentration of approx 10 mg/mL.
5. Stored extracts in small aliquots in liquid nitrogen or at < –70°C. In the latter case, the maximum storage time should not exceed 3–4 mo for reliable results.

3.2. Optimizing Transcription Conditions

It is necessary to systematically titrate the variables (*see* **Note 4**) such as the concentration of KCl, amount of DNA template and nuclear extract to obtain optimal signals from in vitro transcription. **Table 1** provides a set of conditions for this titration. The 20 µL reactions are set as follows.

1. To each microcentrifuge tube on ice, add H_2O, Buffer A-0.1/Buffer A-0, and creatine mix.
2. Add the NE, lightly vortex, and quickly spin the tube.
3. Allow to stand for 5 min at RT (21°C) for preincubation.
4. Add template DNA and leave for another 5 min at RT.
5. Add 1 µL of 8 m*M* rNTP mix and incubate at 25°C for 60 min for transcription (*see* **Note 5**).
6. Stop the reaction by adding 80 µL stop mix.
7. Incubate at 40°C for 30 min to allow Proteinase K digestion.
8. Add 100 µL TE and 200 µL phenol/chloroform. Vortex and spin by microfuge for 10 min.

Table 1
Suggested Conditions for Titration of DNA Template, Nuclear Extract, and Salt Level for Optimal Transcription (*see* Note 6)

Tube	DNA	Buffer A-0.1	Buffer A-0	Nuclear Extract[*8]	Final salt concentration
1	50 ng µL	2 µL	–	10 µL	60 m*M*
2	100 ng µL	4 µL	–	8 µL	60 m*M*
3	100 ng µL	2 µL	–	10 µL	60 m*M*
4	100 ng µL	–	2 µL	10 µL	50 m*M*
5	100 ng µL	–	–	12 µL	60 m*M*
6	150 ng µL	2 µL	–	10 µL	60 m*M*
7	200 ng µL	2 µL	–	10 µL	60 m*M*

9. Transfer the aqueous phase to a fresh tube and add 1/10 vol of 3 *M* NaOAc plus 2.5 vol. of cold ethanol (stored at –25°C) to precipitate RNA for 20 min at –25°C.
10. Centrifuge for 20 min and remove the supernatant.
11. Rinse the pellet with 100 µL 70% ethanol, centrifuge for 5 min and remove the supernatant.
12. Dry the pellet for 1 min by Speedvac. Avoid overdrying.

3.3. In Vitro Transcription with Competitor Oligonucleotides

Once the optimal levels of DNA, nuclear extract, and salt are determined, set up the assays as follows (*see* **Notes 7** and **8**).

1. To each Eppendorf tube on ice, add H_2O, Buffer A-0.1/Buffer A-0, and creatine mix.
2. To a set of tubes, add competitor oligonucleotides (*see* **Note 9**) to a level 50-, 100-, or 200-fold in excess of the template.
3. Add nuclear extract, lightly vortex, and quickly spin the tube.
4. Allow to stand for 5 min at RT (21°C) for preincubation.
5. Add template DNA and leave for another 5 min at RT and proceed as described in the **Step 6** of **Subheading 3.2.**

3.4. Primer Extension for Detecting Transcripts (*see* Note 10)

1. Label 2 pmol of primer using 0.5 µL of T4 polynucleotide kinase, 1 µL [γ^{-32}P]. ATP, 1.0 µL 10X kinase buffer in a total volume of 10 µL ***(1)***.
2. Incubate at 37°C for 15 min. Heat-inactivate the enzyme by heating at 90°C for 10 min.
3. Adjust the oligonucleotide concentration to 10 fmol/µL with TE.
4. Resuspend the pellet from **Subheading 3.2. Step 10** in 5 µL TE.
5. Add 2 µL labeled primer (~20 fmol) and 2 µL annealing buffer.
6. Vortex and spin briefly.

7. Incubate at 85°C for 3 min, and then at 65°C for 30 min (*see* **Note 11**).
8. Allow the tubes to cool down gradually to 37°C over 30 min to anneal.
9. Add 25 µL of Reverse Transcription Mix containing a freshly added AMV reverse transcriptase.
10. Incubate at 37°C for 60 min to allow primer extension.
11. Stop the reaction by adding 1/10 vol of 3 *M* NaOAc.
12. Add 2.5 vol of ethanol for precipitation, leave at –25°C for 20 min and centrifuge at 13,000 rpm for 15 min.
13. Remove the supernatant and rinse the pellet with 70% ethanol.
14. Centrifuge for 5 min. Remove the supernatant and dry in a Speedvac for 1 min.
15. Resuspend the pellet in 6 µL formamide/dye. Heat-denature at 95°C for 5 min and snap-chill in ice.
16. Resolve on a denaturing acrylamide gel (5–6%) with appropriate size markers (*see* **Note 12**).
17. Fix the gel in 10% acetic acid and dry for autoradiography.

4. Notes

1. Care should be taken to avoid nicking of plasmids in the presence of ethidium bromide during the CsCl gradient purification step, e.g., by covering tubes with aluminium foil to minimize exposure to ultraviolet (UV) light.
2. To allow a space in the reaction for competitor, high concentration of the stock solutions are used to minimize the total volume of the mixture. To reduce pipeting error in handling small volumes, prepare the mixture to cover all the reactions.
3. Glycogen can be used instead of tRNA as a carrier.
4. Other parameters to be optimized are the concentration of $MgCl_2$, incubation time, and temperature for a given promoter. In addition, template DNA concentration should be retitrated and optimized for each batch of nuclear extract.
5. The incubation temperature was reduced from 30 to 25°C for transcription reaction to reduce background RNAse activity.
6. The total volume occupied by Buffer A-0.1, A-0, and NE is adjusted to be 60% of the reaction to give the final concentration of 12 m*M* HEPES-KOH, pH 7.9, 0.6 m*M* EDTA, and 12% glycerol.
7. When the concentration of DNA template is too high, and the signal is strong, it may be difficult to detect competition effect. In such cases, reduce the concentration of DNA to the level at which a competition becomes detectable. The same applies for the concentration of NE.
8. When the test promoter has a single transcription initiation site, more convincing results may be obtained by including in the same reaction a control plasmid led by a promoter whose transcription activity is independent of specific transcription factors to be depleted by the competitor oligonucleotides. For example, AdMLP, SV40, or other viral promoters can be placed in front of the same reporter gene used for the test template. Transcripts derived from such control promoter–reporter gene constructs are detected using the same primer, and the size of the cDNA can serve to identify transcription from the specific promoter ***(5)***.

9. As an alternative to the inclusion of a control promoter, oligonucleotides can serve as a control. In this case, a good control is an oligonucleotide, which is identical to the test sequence, but mutated in the cognate region, so that it does not bind the specific transcription factor.
10. Detection of transcripts by primer extension using a specific primer gives results with relatively clean background in autoradiography, compared to the method using conventional run-off transcription (i.e., not using G-less cassette). However, titration of the primer concentration (e.g., 5–30 fmol) is necessary to obtain the optimal ratio of signal to background. For the same reason, the concentration of the carrier tRNA in the stop mix was halved in this protocol compared to the standard one.
11. The brief incubation at high temperature is to remove the secondary structure in RNA products. The second incubation time at 65°C is longer than some protocols suggested for annealing, but it is because of the low concentration of a primer used. The temperature may be optimized for annealing a primer to the target RNA.
12. For gel electrophoresis, it is helpful to run labeled size markers (e.g., HpaII digested pBR322) or a sequence ladder for more accurate determination of the initiation site.

References

1. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. 2nd ed., Cold Spring Harbor Lab. Press.
2. Dignam, J. D., Lebovits, R. M. L., and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11,** 1475–1489.
3. Prywes R. and Roeder, R. G. (1986) Inducible binding of a factor to th c-fos enhancer. *Cell* **47,** 777–784.
4. Shapiro, D. J., Sharp, P. A., Wahli, W. W., and Keller, M. J. (1988) A high-efficiency HeLa cell nuclear transcription extract. *DNA* **7,** 47–55.
5. Watt , F. and Molloy, P. L. (1993) Specific cleavage of transcription factors by the thiol protease, m-calpain. *Nucleic Acids Res.* **21,** 5092–5100.
6. Nicolaiew, N., Triqueneaux, G., and Dautry, F. (1991) Organisation of the human N-ras locus: characterisation of a gene located immediately upstream of N-*ras*. *Oncogene* **6,** 721–730.
7. Watt et al., to be submitted.
8. Thorn, J. T., Todd, A. V., Warrilow, D., Watt, F., Molloy, P. L., and Iland, H. J. (1992) Characterisation of the human N-ras promoter region. *Oncogene* **6,** 1843–1850.
9. Unpublished results.

22

Computer Software for Eukaryotic Promoter Analysis

Dan S. Prestridge

1. Introduction

1.1. Background

Since the initiation of the Human Genome Project ***(1–3)***, project originators have envisioned a need for the development of computer software programs to analyze and annotate the massive amount of sequence data expected to be produced as a result of sequencing efforts ***(4)***. Efforts directed toward this goal include the development of computer software for gene and gene feature recognition ***(5)***, including the recognition and functional characterization of eukaryotic promoter sequences.

Because the development of computer methods to recognize genes is of a more obvious and immediate need, the development of computer methods to recognize and characterize eukaryotic promoter sequences, although less obvious ***(6)***, is nonetheless just as important. These reasons include: 1) aiding in the identification of gene sequences, 2) aiding in the identification of the regulatory properties of promoters that control the expression genes downstream of them, and 3) aiding in the identification of networks of genes that respond to a specific regulatory stimulus (for example, the addition of a hormone), in both lateral and hierarchical regulatory control networks.

The development of such computer algorithms has, however, been a very difficult and long process because of the very complex and individual structural characteristics of eukaryotic promoter sequences ***(7–16)***. Eukaryotic promoter sequences, unlike their much simpler prokaryotic counterparts, do not contain any sequence elements that are consistently shared between all promoter sequences. Instead, each eukaryotic promoter contains a unique set of transcriptional elements that differentiates its regulatory properties from most

From: *Methods in Molecular Biology, vol. 130, Transcription Factor Protocols*
Edited by: M. J. Tymms

other promoter sequences in the genome, allowing the multitude of unique regulatory programs to exist that are required by highly complex, differentiated, multicellular organisms *(9,10,17)*.

Eukaryotic Pol II promoters, the promoters of all known protein coding genes, often contain one or both of two basic sequence elements: the TATA box *(18–20)* and the Initiator (INR) *(11,14,20–22, for review)*. Approximately 70–80% of pol II promoter sequences contain a TATA or TATA-like element, with a consensus of TATAA *(23)*, about 25–30 bases upstream of the transcription start site (TSS). Many promoters also contain another element at or near the TSS, the INR. While both elements can promote transcription, the TATA box seems to be the element most important in determining the location of the TSS. In addition to these two common transcriptional elements (TEs), each promoter contains a unique combination of other TEs from the hundreds that are now known and described *(10,13)*. These elements give each promoter sequence its own unique program of expression in different tissues, cell types, and under different developmental, physiological, or environmental conditions.

Most promoter analysis programs start with an attempt to recognize TEs. One group of programs only attempts to locate TE binding sites, making no attempt to carry the analysis any further. A second group of programs extends the analysis further by attempting to recognize patterns of TEs that are consistent with promoter sequences. Finally, a third group of programs attempts to functionally classify predicted promoter sequences, or locate specific functional classes of promoters (**Table 1**).

Attempting to find putative TE binding sites in regulatory DNA sequences began with a database of transcription factor binding site sequences *(24)* and a program developed to scan a DNA sequence against that database *(25)*. From that time, several transcription factor binding site sequence databases have been developed along with a suite of programs to analyze novel DNA sequences against these databases (**Subheading 2.**). Today, because there are several TE databases and an abundant number of TE analysis programs to utilize these databases, none has been shown to be clearly superior to the others, and all of them function to give an investigator a large list of putative TE binding sites; most of them false positive sites that must be functionally investigated before being accepted as functional sites.

At about the same time, efforts began on developing computer programs to predict pol II promoter sequences *(23)*, or specific types of promoter sequences *(26,27)*. These early efforts were based upon the recognition of individual elements that are common to many pol II promoters, such as the TATA box *(23)*, or upon the recognition of a specific functional class of promoter sequences using a simple pattern of elements *(26,27)*. Today, there are many promoter prediction programs available, based upon one or more of several recognition

Table 1
Databases and Programs Available for Analyzing Eukaryotic Promoter Sequences

Transcription Factor Binding Site and Promoter Databases		
Database	Access	Internet address
TFD	FTP	ncbi.nlm.nih.gov, repository/TFD directory
TRANSFAC	WWW	transfac.gbf-braunschweig.de
IMD	FTP	beagle.colorado.edu, pub directory
COMPEL	FTP	transfac.gbf.de/pub/databases, compel directory
TRRD	FTP	transfac.gbf.de/pub/databases, trrd directory
EPD	FTP	ncbi.nlm.nih.gov, repository/EPD directory
Transcription Factor Binding Site Analysis Programs		
Program	Access	Internet address
SIGNAL SCAN*a*	FTP	biosci.cbs.umn.edu, pub/sigscan directory
PatSearch	WWW	transfac.gbf-braunschweig.de/cgi-bin/patSearch patsearch.pl
TESS	WWW	agave.humgen.upenn.edu/utess/tess
MATRIX SEARCH	FTP	beagle.colorado.edu, pub directory
MatInspector	WWW	www.gsf.de/biodv/matinspector.html
ConsInspector	WWW	www.gsf.de/biodv/consinspector.html
TFSEARCH	WWW	www.genome.adjp/SIT/TFSEARCH.html
Promoter Prediction and Characterization Programs		
Program	Access	Internet address
GRAIL*a*	WWW	avalon.epm.ornl.gov/Grail-1.3/
NNPP	WWW	www-hgc.lbl.gov/projects/promoter.html
PROMOTER SCAN*a*	WWW	biosci.umn.edu/software/proscan promoterscan.htm
TSSG/W	WWW	dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html
FunSiteP/AutoGene	FTP	ftp.bionet.nsc.ru, pub/biology/aug directory
Promoterl/GeneID	Email	geneid@darwin.bu.edu
PromFind	FTP	iubio.bio.indiana.edu, molbio/ibmpc directory
PromFD	WWW	beagle.colorado.edu/~chenq
GenomeInspector	FTP	ariane.gsf.de, pub/unix/genomeinspector directory
Model Generator/ Inspector	WWW	www.gsf.de/biodv/modelinspector.html

*a*Other internet sites are available—see text.

strategies. Recognition and functional classification of pol II promoter sequences has improved significantly over the past few years; however, false

positive rates are still very high in all of them, limiting their present usefulness. However, progress has been rapid in this field, and if false positive rates can be dropped with an increase in the ability of these programs to functionally classify promoter sequences, they may prove to be very valuable tools in discovering novel genomic sequences of targeted genes with desired regulatory properties.

1.2. Analyzing DNA

1.2.1. Analyzing Known Promoter Sequences

If an investigator has a known promoter sequence, but knows little else about the sequence, there are three characteristics that can be investigated by analysis of the promoter sequence: 1) locating the transcription start site, 2) identifying transcription factor binding sites that may be contained in the sequence, and 3) characterizing the regulatory properties of the promoter sequence (**Fig. 1**).

The first analysis (transcription start site prediction) can be done using programs that attempt to predict promoter sequences (**Subheading 3.**). These programs generally require 100–500 bases of sequence data, depending on individual program requirements. These programs are largely in experimental stages, and none of them produce reliable results yet; although some at least yield results that may be worth testing further. Some of these programs attempt to make very accurate predictions of the start site, whereas others only predict regions of sequences that are promoter-like, but do not indicate a TSS. Those programs that are better at identifying the actual start site of transcription are trained to recognize a TATA box and/or an initiator region. These programs, however, whereas better at locating the start site, are generally poorer at making overall promoter predictions due to very high false positive reporting rates.

The second type of analysis (locating putative transcription factor binding sites) may be done by using sequence analysis software along with a database of binding site sequences (**Subheading 2.**). There are several databases of transcription factor binding sites and software packages in existence that can use them in analyzing sequences. There are also programs (such as in the GCG and Staden packages) that can perform this type of analysis, but are not associated with corresponding databases, are more difficult to use, and produce results that are more difficult to interpret. These programs will not be discussed here; only those programs that contain precompiled databases, and are customized for this type of analysis will be reviewed. Results of these types of programs will give the user a list of putative binding sites, often a very long list. These programs, like the promoter prediction programs, produce results indicating many more false positive sites than true sites. This problem limits the use of these types of programs at present to situations where *(28)*:

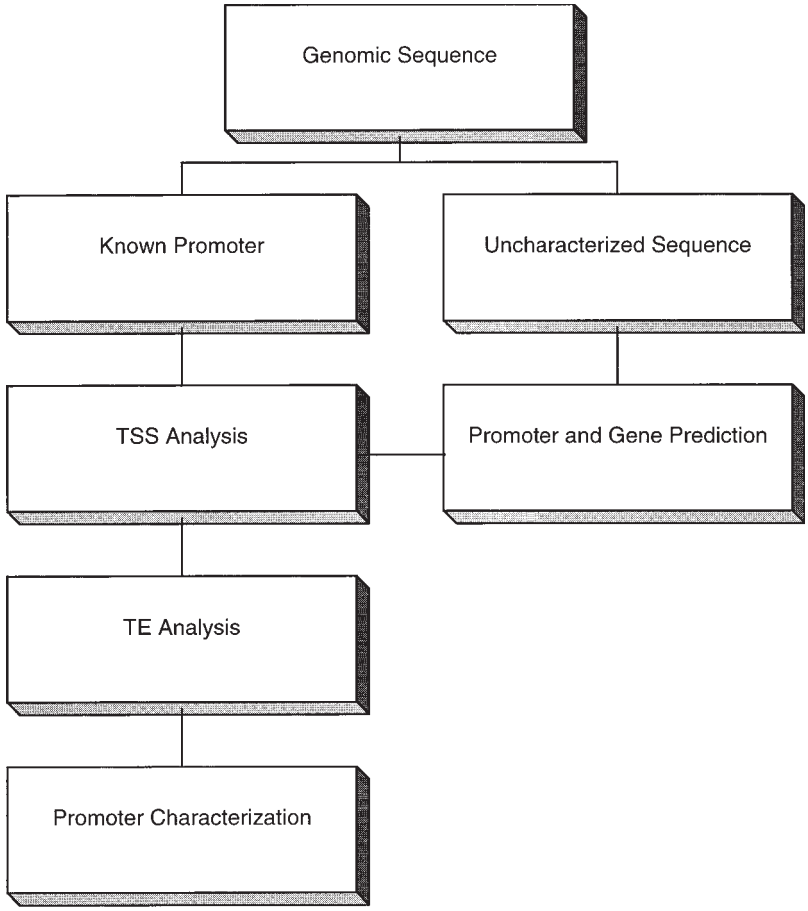


Fig. 1. Promoter analysis flow chart. Beginning wilh genomic sequence data, if the sequence has not been characterized, the analysis begins with promoter and prediction. Once a putative promoter has been identified, analysis can proceed to identify a putative transcription start site, and putative transcription factor binding sites. The final step would be the characterization of putative regulatory properties of the novel promoter sequence; however, programs with this capability are only now being developed.

1. an investigator knows a specific site in a sequence is bound by an unknown factor, but does not know what the factor is;
2. an investigator has a known promoter sequence and wishes to identify putative regulatory sites;
3. an investigator has some evidence of specific regulatory properties of a sequence but has not located the suspected regulatory elements.

The third type of promoter analysis program makes an attempt to further characterize the functional regulatory properties of a promoter sequence (**Subheading 3.3.**). These programs have only very recently become available and are in early stages of development; however, they may prove to be the most valuable of all promoter analysis programs if they prove to be successful in characterizing the regulatory properties of promoter sequences or aid in locating promoters with desired regulatory properties in a databank.

1.2.2. Analyzing Uncharacterized Genomic Sequences

The analysis of a previously uncharacterized genomic sequence is usually aimed at finding genes and regulatory sites that control the gene's activity (promoters and enhancers). There are, at present, no programs specifically designed to identify enhancers. There are, however, several promoter prediction programs available (**Subheading 3.**). For analyzing genomic length sequences, usually hundreds of thousands of bases to megabases in length, the promoter prediction programs that attempt to locate promoters are based only upon recognition of a TATA box and an initiator site are unsuitable. Because these programs may be best for locating the exact start of transcription, they are the poorest programs for locating promoters in large genomic length sequences due to their very high false positive rates.

Because of the very high false positive rates reported by the transcription factor binding site analysis programs, it makes little sense to perform this type of analysis on genomic length sequences.

2. Transcription Factor Binding Site Recognition

2.1. Transcription Factor Databases

2.1.1. Methods for Representing Binding Site Sequences

Computational methods for recognizing transcriptional regulatory elements, most especially transcription factor binding sites, is still a field in a great deal of flux ***(29)***. Ways in which binding sites are represented, databases of binding sites, and software used to analyze DNA sequences for binding sites and promoter sequences, have only recently been developed and are undergoing a rapid degree of change and evolution.

Before any complex analysis can be done to locate functional transcriptional elements or a promoter sequence (a functional unit of transcription factor binding sites) in uncharacterized DNA sequences, effective methods must be developed in effectively representing binding sites, and developing databases of these sites. Collections of specific binding sites, or representatives of these collections of specific sites, consensus sequences, and position weight matrices, have been developed and are all being used (**Fig. 1**). Each type of repre-

sentation has advantages and disadvantages; as yet there is no model available that can represent binding sites with perfect recognition of functional sites with no false positives.

2.1.1.1. Specific Binding Sites

A single segment of DNA in a promoter or enhancer sequence that is bound by a particular transcription factor is called a *specific binding site* (**Fig. 2A**). It is unambiguous and represents only that one specific protein binding site, although it is usually very similar to other specific binding sites bound in other locations in the DNA by the same transcription factor. One way to represent these binding sites is to simply record each individual binding site in a database, and use this database of specific binding sites to scan DNA sequences that match these sites.

One advantage to using specific binding sites is that any site in the scanned DNA that matches a site in the database is ensured of matching exactly a site of known functionality in the database. However, finding a site that matches one in the database does not immediately imply functionality, and this applies to all present methods of representing binding sites, including those methods described below. In order for a transcription factor to be considered functional, it must be verified by experiment. There are many possible reasons why a site that matches one in a database may not be functional. It may be possible that there are other sites in close proximity that interfere with factor binding, or the site may be in the wrong sequence context. Transcription factors must usually interact with other transcription factors in order to function. A factor binding to a site that is out of context will not be functional.

As many earlier studies of transcription factor binding sites have shown, transcription factors do not bind only a specific binding site sequence, but instead a family of closely related sequence sites. A major disadvantage to using only specific binding sites in a database is the high probability that not all possible binding sites that can be bound by a transcription factor have been reported or are known. Hence, many sites that might indeed prove to be functional sites would not be reported using a specific binding site database to scan a DNA sequence. To solve this problem, two other methods of representing binding sites have been adopted: 1) consensus sequences and 2) weighted matrices ***(30–32 for review)***.

2.1.1.2. Consensus Sequences

Because any transcription factor will bind a family of closely related binding site sequences, it would be useful to include the entire family of sites in one descriptor (**Fig. 2B**). One way of doing this is by combining the individual site information into a consensus sequence. Ideally, a consensus sequence would

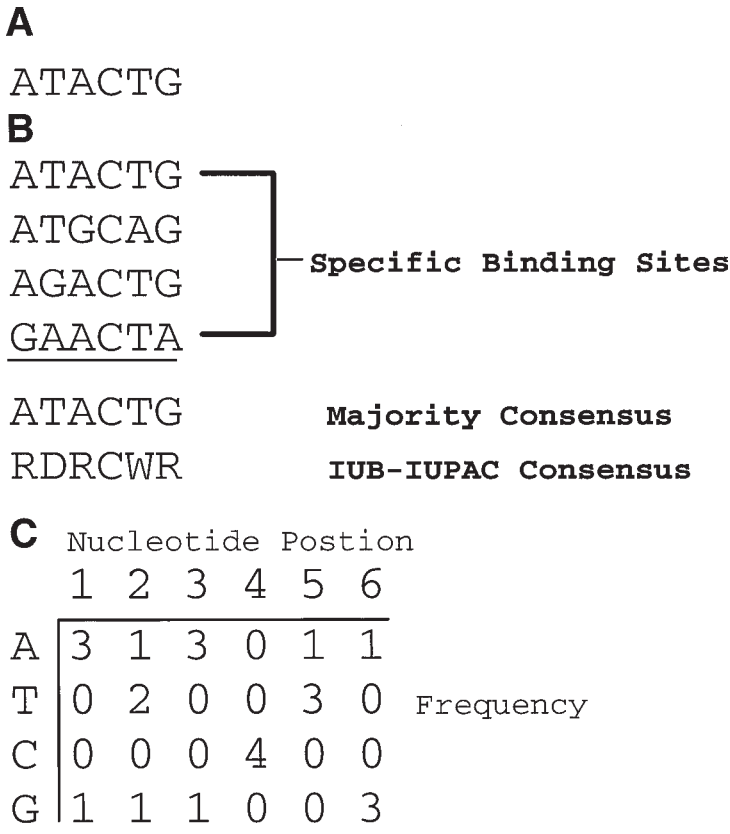


Fig. 2. Representing binding sites. A. Example of a specific binding site. B. Example of a family of closely related binding sites bound by the same factor, and the development of a consensus sequence by majority representation or by an all-inclusive IUB-IUPAC ambiguous nucleotide representation. C. Building a position weight matrix for the same family of binding sites used in B.

be constructed so that it would recognize every site used to build the consensus, could be used to recognize all sites bound by the factor yet are slightly different than any of the sites used in constructing the consensus, and yet would be able to discriminate against those sites that are closely related on a sequence level but are not bound by the factor. Whereas such an ideal consensus binding site does not exist for any factor, consensus sequences are still useful as an aid to identifying possible sites.

There are two ways in which to construct a consensus sequence. Both involve aligning all of the sites to be included in the consensus, but differ in the way they assign a descriptor to represent each column of nucleotides. One is

based upon a "majority wins" rule, in which the majority of identical nucleotides in a column of aligned sites is used in the consensus. Using this method to recognize novel sites will result in fewer reported sites. A higher number of true positives to false positives will be found, but it will miss many functional sites, even some used in building the consensus itself. Indeed, it is possible, using this method, to build a consensus by majority that will not recognize a single site used in building the consensus.

The second method that may be used in building a consensus sequence is by utilizing the IUB-IUPAC ambiguous nucleotide codings ***(33)*** to represent each column of aligned sites. This method is usually all-inclusive and can be used to recognize every site used in building the consensus. When used to recognize novel sites, a IUB-IUPAC-based consensus will find more true sites, but with a much higher false positive rate.

Whereas many, if not most, consensus sequences have been developed by manual alignment of known binding sites, software is also available for defining these sites automatically. The advantages of using these programs to develop consensus sequences are that the sites either need only be roughly aligned ***(34)*** or not aligned at all initially ***(35,36)***. Additionally, the consensus may be better defined than those that are manually developed ***(34)***.

The consensus method is usually chosen by the original investigator who describes the binding site and not the database staff. This leads to another problem in using such consensus sequences: a problem of quality control. Not only are the various consensus sequences developed using different methods, but because they are developed in the laboratories in which they are investigated and described, each consensus is built upon a unique number of sites of varying quality using independent (and nonstandardized) binding site verification methods. The creation of consensus sequences in this way results in databases of these types of sites whose accuracy can never be verified or measured in any reliable and consistent way. Although databases of consensus sequences and specific binding sites are filled with such variable and unreliable data, when used with these deficiencies in mind, they can, and have proved, to be very useful in locating putative transcription factor binding sites. In an effort to better, and more accurately, describe these families of binding sites, position-weighted matrices (PWMs) have been developed.

2.1.1.3. Position-Weighted Matrices

PWMs have several advantages over the use of consensus sequences and specific binding sites, and some deficits. PWMs are also developed by multiple alignment of the known binding sites. Next, however, the frequency of each possible nucleotide (A, T, C, or G) is calculated for each nucleotide position in the binding site. This results in a matrix of frequencies of four rows (one

for each nucleotide) and a number of columns equal to the length of the binding site (**Fig. 2C**). These base frequencies may then be used directly in scoring against a DNA sequence, or they may be further converted into weights by using one of several methods. A DNA sequence can then be scored against the matrix. If the score is high enough, above a given cutoff value (usually experimentally or statistically determined), then the sequence is considered a match to the binding site.

Advantages of PWMs over other methods include the use of a cutoff value (which can be adjusted to various levels of sensitivity). PWM scores have been correlated with binding site strength ***(37,38)***, and higher overall sensitivity ***(39)***. Disadvantages of using PWMs include the following: 1) a fewer number of matrices available today than specific binding site and consensus sequences, 2) they are more difficult to develop by primary investigators and hence are usually constructed by database staff, 3) they require a greater number of sites to produce a reliable PWM, 4) PWMs are not sensitive to the possibility of base dependence (when a change in one nucleotide in one position may affect the nucleotide in another position within the binding site), and 5) PWMs do not usually take into consideration the possibility of subtle subfamilies of binding sites (the last two items are also a problem with consensus sequences).

Overall, the use of a well-developed PWM based upon a careful selection of binding sites is better than using specific binding sites or consensus sequences ***(40)***. However, in practice, there are only a few factors that have a large enough number of binding sites available to produce a quality PWM, and often the data are not prescreened well enough before inclusion into the matrix. Both problems can lead to collections of poor quality matrices that are not any better than using specific binding sites or consensus sequences. The development and availability of PWM databases has been relatively recent, and should improve over time with an increase in the availability of primary binding site data and the improvement of methods for developing better matrices ***(41)***.

2.1.2. Databases

Several databases of transcription factors and their corresponding binding sites have been compiled and are described here. These databases are, however, of little use without corresponding sequence analysis software (**Subheading 2.2.**) except for information retrieval, when that capability exists. In such cases, the database may be searched by factor name, binding sites, or other search term.

2.1.2.1. Transcription Factor Database (TFD)

The Transcription Factor Database ***(24,42–44)*** was the first public, computer readable database of transcription factor binding sites. It is currently,

however, not maintained, having been released last in 1993. The database is, however, still very useful, although aging rapidly. The TFD has recently been updated: Ghosh, D. (1999) Object Oriented Transcription Factor Database (ooTFD). Nucleic Acids Res. 27, 315–317.

The TFD is a relational database (a database of logically linked tables) consisting of tables containing information on binding sites (SITES table), transcription factors (FACTORS), binding domains (DOMAINS table; which contains zinc finger domains, leucine zipper domains, etc.), cDNAs of cloned transcription factors (CLONES), deduced transcription factor peptide sequences (POLYPEPTIDES), methods, and references. The SITES database contains both specific and consensus binding sites. The latest version can be obtained via anonymous FTP to ncbi.nlm.nih.gov, in the repository/TFD directory. However, these files are not directly useful without a database program or analysis program that is compatible with the TFD. Database programs such as PARADOX or SYBASE (or any other relational database management system) can be used if the user is familiar with QBE or SQL database query languages. A program that was specifically designed to provide investigators with a user friendly interface to the TFD, InfoTrac TFD ***(45)***, seems to be, at the time of this writing, no longer available (as the author could not be reached). The difficulty in accessing the TFD has mostly limited its use to sequence analysis programs. Only one program is specifically designed to utilize the TFD, and this is the SIGNAL SCAN program ***(25,28,46)***, however, some commercial sequence analysis packages contain utilities to make use of the TFD SITES table (including GCG, Intelligentics GeneWorks, and MacVector) to do trivial sequence scans.

2.1.2.2. TRANSFAC AND TRANSFAC MATRIX DATABASES

TRANSFAC is a more recent, currently up-to-date transcription factor database that has much better software support for both database retrieval and sequence analysis. The TRANSFAC database ***(10,47–50,55)*** began as a simple compilation of transcription factors and their binding sites ***(47)***, later progressing into computer readable format ***(51)***. The structure of the database is very similar to that of the TFD, consisting of eight relational tables: sites, factors, cells (cell and tissue types), methods, references, classes (similar to TFD DOMAINS), matrix (compiled position weight matrices for factors with abundantly available binding site data), and gene (containing data about the promoter or gene containing the element) tables.

The TRANSFAC database, like the TFD, contains both specific binding site and consensus sequences; however, the consensus sequences are generated from data within the TRANSFAC database from a consensus-generating program called ConsIndex ***(34)***. In addition to these, TRANSFAC also contains a

library of PWMs for the binding sites of some of the more major transcription factors. The TRANSFAC database can be found on the web at transfac.gbf-braunschweig.de.

The TRANSFAC database has its own web-based data retrieval program that is both easy to use and easy to access (at the above web address). The TRP (TRANSFAC Retrieval Programs) programs ***(50)*** are no longer supported (E. Wingender, personal communication).

Several programs are available that use the TRANSFAC database to scan sequences for putative binding sites. Programs that can scan sequences using the TRANSFAC specific and consensus binding sites include PatSearch, SIGNAL SCAN, ConsInspector, and TESS. Programs that can utilize the TRANSFAC matrices include MatInspector and TFSEARCH. All of these are discussed in **Subheading 2.1.2.3.**

2.1.2.3. Information Matrix Database (IMD)

The IMD is a recently developed database of transcription factor binding site weight matrices ***(51)***. It is very similar to the TRANSFAC matrices, except the matrices are developed using a different algorithm. Unlike the TRANSFAC database, it is a flat-file database containing only the matrices, factor names, and reference citations. The matrices are divided into files containing matrices of seven different organism classes. The TFD and TRANSFAC databases were both used in developing the IMD database. At present, there is no utility or program that can be used to browse the database, however, there is an associated sequence scanning analysis program called MATRIX SEARCH ***(51)***. The MATRIX SEARCH program has also been incorporated as part of the SIGNAL SCAN program (Version 4).

2.1.2.4. COMPEL, TRRD, and EPD

The Composite Element (COMPEL), Transcription Regulatory Region Database (TRRD), and the Eukaryotic Promoter Database (EPD) are all specialty databases often used by investigators developing promoter recognition algorithms. However, occasionally, they may be of use to other investigators interested in gene regulation and promoter structure.

The COMPEL database ***(50,54)*** consists of composite elements: two or more transcriptional elements that act together in a synergistic or antagonistic relationship. These elements are located on the same DNA helix and may be located from overlapping positions to about 80 base pairs distant. At present, there are no analysis programs that use the COMPEL database to analyze DNA sequences.

The TRRD database ***(50,56)*** is a higher order database consisting of a hierarchy of regulatory features from an upper level represented by regulatory

regions that regulate the activity of several genes, to lower levels of regions that regulate single genes, to promoters, enhancers, transcription factor binding sites, and other regulatory entities. Like the COMPEL database, there are no programs that are capable of using this data to analyze sequences at present.

The EPD ***(23,57)*** is a compilation of well-characterized eukaryotic promoter sequences. It also contains data on known transcriptional elements contained within the promoter sequences, and information on some regulatory properties. If homology to one or more of the promoters contained in the EPD can be found against a sequence of interest, then the EPD may provide clues to some putative regulatory properties of the sequence under investigation. For this purpose, the EPD can be converted into a GCG readable database so that the GCG homology search programs (such as Fasta) can be used to compare a sequence of interest to the EPD. There is no specialized software currently available with which to browse the EPD or to use the database to analyze DNA sequences.

2.2. Transcription Factor Binding Site Analysis

There are two basic types of binding sites analysis programs, based upon the type of database they use. One class of programs, an older class, utilizes the specific binding site and consensus sequence databases. These are the SIGNAL SCAN, ConsInspector, TESS (String-based search), and PatSearch programs. The second class of programs are based upon the binding site weighted matrix databases including MATRIX SEARCH, ConsInspector, MatInspector, TFSEARCH, and TESS (Matrix-based search). TESS and ConsInspector are capable of both types of analyses, and the latest version of SIGNAL SCAN contains an integrated version of MATRIX SEARCH.

There is, at this time, no clearly superior method of searching for binding sites. There are advantages and disadvantages for both general approaches, and for each program. Prediction of binding sites using matrix databases are usually considered to be of a higher quality. However, there are a limited number of matrices available and many transcription factor binding sites that are not yet represented by a weighted matrix. In addition, the quality of the analysis results are strongly influenced by the quality of the matrix itself ***(58,59)***: how the matrix was constructed, and what specific sites were used in development of the weight matrix. Matrices, and the methods used to create them, are improving rapidly; however, until there are tested, high-quality weighted matrices available for all transcription factor binding sites, sequence analysis programs that use specific binding site and consensus sequence databases will continue to be useful.

Programs that use matrices also have the advantage over programs that use specific binding sites or consensus sequences by being able to score for the quality of the match. It is possible to score a match to a specific site or consensus sequence by assigning a probability for a random match, but this indicates only the probable rarity of the match, and does not indicate how well the sequence matches the pattern. A match to a specific site or to a consensus is essentially a hit or no hit; however, when a sequence is scored and compared to a matrix, the score can indicate the quality of the match. This is important for two reasons: 1) the user can select a match cutoff score, which can give the user more power over what is an "acceptable" site and 2) there is some evidence that a matrix score is correlated with a site's biological functionality ***(37)***.

It is difficult to determine which databases and programs are best in this type of analysis. There are several problems in testing them against each other. It is easy to determine the true positive reporting rate for these programs; however, it is impossible to tell if a site reported by a program in a test sequence where no such site has been experimentally determined is a false-positive report, or a report of a true site that has not yet been discovered. In addition, the quality of consensus sequences and weight matrices can vary substantially within a database. One database may have a better representation of specific site, and poorer representation of another site. There are, at this time, no standards or universal methods for determining how consensus sequences or weighted matrices are constructed. There is some evidence, however, that the quality of the database or specific matrix is more important than the chosen analysis program or method of scoring ***(58,59)***. This supports the idea that database or matrix selection might be more important than the selection of the analysis program.

The best criteria for selecting an analysis program are: 1) the databases it can access, 2) its availability and ease of setup, 3) web access or the availability of compatible computer platforms (UNIX, PC, Macintosh), 4) the quality of the user interface, 5) the ability to select organism class-specific databases, 6) the ability to select a specific site or matrix, and 7) the ability to use user-supplied binding sequences or matrices. All of the programs reviewed below are readily available, either as web-based programs or easily installed by anyone with common knowledge of computers.

2.2.1. Specific Binding Site and Consensus Programs

2.2.1.1. SIGNAL SCAN

SIGNAL SCAN ***(25,28,46)*** was the first program that was specifically designed to analyze DNA sequences for transcription factor binding sites, and the first to include a precompiled database. The latest version of the program,

version 4 ***(46)***, includes both specific and consensus binding site databases and programs as well as a weight matrix database and program (MATRIX SEARCH described below). Both the TFD and TRANSFAC databases have been included, and have been partitioned into organism-class-specific databases. SIGNAL SCAN also includes a utility that allows the user to update either database. In addition, users can create a database of their own binding sites to scan sequences. Search results can be displayed in order of sequence location, grouped by factor name, or mapped to the sequence. The program accepts GCG, Staden, and Fasta formatted sequences. The SIGNAL SCAN program comes in both IBM PC and UNIX compatible versions (anonymous FTP to biosci.cbs.umn.edu, in directory pub/sigscan) or can be accessed at three different web sites (http://bimas.dcrt.nih.gov/molbio/signal/, http://www.tigem.it/TIGEM/HTML/sigscan.html, an http://ch.nus.sg:80/bio/sigscan/signal.html). The web versions of SIGNAL SCAN were developed independently by Robin Hart and Rao Parasa at the National Institutes of Health (United States) and Meena Sakharkar at the National University of Singapore. Whereas the downloadable and web versions are largely the same, there are some slight differences. For example, the PC and UNIX versions can perform sequence searches utilizing user-supplied binding site sequences, whereas the web versions cannot. However, in the web versions, there are links to database entries for reported binding sites.

2.2.1.2. PatSearch

PatSearch (or PatternSearch) ***(50,55)*** is one of the latest entries in the field. It is only available through the web (http://transfac.gbf-braunschweig.de/cgi-bin/patSearch/patsearch.pl), and can access only the TRANSFAC database. However, it is the only program at present that can take advantage of using the special notation used in TRANSFAC to indicate more significant nucleotides in binding sites, which are represented as uppercase letters in the database. Additionally, PatSearch provides a probability score based upon the likelihood of a random match (this is not the same as scores provided by matrix matching programs, and not as meaningful). The user can select a percent mismatch for both core and flanking regions (nucleotides represented as lower case in TRANSFAC binding site definitions), as well as some other user-modifiable parameters. However, a search for vertebrate binding sites on a randomly generated 170 bp sequence, using the program defaults which allows a 50% mismatch in flanking regions (0% mismatch in the core), found 257 binding sites—far too many to give any useful binding site information (by comparison, the SIGNAL SCAN program, using all TRANSFAC vertebrate binding sites, finds 37 exact matching sites). If the flanking region mismatch percentile

is lowered to 0%, then only six binding sites are found; a great disparity in results. This proves to be a problem with any such program that allows mismatches: allowing mismatches where exact matches to specific and consensus sequences already leads to an almost unacceptable number of false positives, dramatically increases the number of reported sites. Allowing mismatches leads to results so populated with false positive sites that results may be meaningless to most investigators. Allowing mismatches while scanning for single sites, or a small family of sites, would be a useful feature; however, in this version of PatSearch, there is no capability to limit your search to specific binding sites, except by organism class. The results are linked to the corresponding TRANSFAC database entries, but do not indicate the nucleotides matching core nucleotides, and sites are not ordered in any way (such as by score, factor, or occurrence in sequence).

Because of its ability to use TRANSFAC significant nucleotide notation, the capability to allow some user supplied parameters, and the calculation of random probability of matches, the PatSearch program may prove to be a better search tool than the simpler programs such as SIGNAL SCAN. However, at the present time, because of the added complexity and overly relaxed defaults in the program parameters, its use may be better limited to more experienced investigators.

2.2.1.3. TESS—String-based Search

TESS (http://agave.humgen.upenn.edu/utess/tess) consists of three search programs: String-based search, Filtered String-based search, and matrix-based search programs (Schug, and Overton, unpublished). The matrix-based search program will be discussed in a later section.

TESS is TRANSFAC-based. Like PatSearch, TESS also allows the user to set the percentage of mismatches (the default is 10%); however, this is a mismatch over the whole length of the binding site, not separate core and surrounding nucleotide mismatch parameters as in PatSearch. TESS ignores the special TRANSFAC notation for significant nucleotides. However, unlike PatSearch, while the capability to allow mismatches would result in an extreme number of false positives, users can limit the number of binding sites reported by the ability to set two additional score cutoff parameters based upon log likelihoods. Only those sites that are within the percentage mismatch and greater than the minimum likelihood cutoffs will be reported to the user. However, since the cutoffs are based upon log likelihoods, it is a bit difficult for most nonmathematically oriented molecular biologists to understand and use correctly (versus something like percentage match or a simple probability). However, these values can just be modified by the user until the desired results are obtained.

The same random sequence that PatSearch reported 257 binding sites (and SIGNAL SCAN reported 37), TESS reports a much more manageable 16 binding sites using default settings (using the Filtered search, and selecting binding sites from the vertebrate database). The default values in TESS seem to be much more realistic than those in PatSearch. The results can be displayed either in text format or by using Java. The Java output is a map, and pointing to factors using a computer mouse will highlight the factor binding site; however, the sites in the Java-based output are not linked to database entries as they are in the text output. The text output may be sorted in any number of different ways.

The String-based search is of little use because the whole TRANSFAC database is searched (default output on the test sequence is 124 binding sites). Most of these binding sites would be from organism classes far removed from that being investigated, and the corresponding binding site information would usually be of marginal interest. The Filtered string-based search is much better. In this search, the user can limit the factors used in the search by the factor selection attribute (several selectable fields including organism class, accession number, name, interacting factors, and other fields), and a search pattern (the actual accession number, class, name, etc.). Selecting the attribute to search is easy, however, entering the search pattern is not. For example, to include all vertebrate factors, the user must know to search for "vertebrate." There is, however, help on this, that includes a list of search terms that may be used. This capability is the most powerful and flexible way to limit searches of any program in the field. It is the only program in this category that can, for example, search for single binding sites.

2.2.2. Position Weight Matrix Programs

2.2.2.1. MATRIXSEARCH/SIGNAL SCAN 4

MATRIXSEARCH ***(53)*** is available as a stand-alone program (UNIX version only) or as a program integrated into the SIGNAL SCAN program (UNIX and IBM PC compatible versions). The latest version is the stand-alone program and is available by anonymous FTP to beagle.colorado.edu in the "pub" directory as a file called imd.tar. The program is easy to set up and has a simple text-based user interface modeled after the SIGNAL SCAN program. There is no web version presently available.

MATRIXSEARCH will accept Fasta or Staden formatted sequences and comes with its own precompiled transcription factor binding site matrix database (the IMD, >350 matrices) that is subdivided into several organism classes. The only menu options are to scan a sequence for binding sites or to view matrix information for selected matrices. Whereas the program does not allow

the user to enter any kind of score cutoffs, it does offer the user the choice of including matrices that produce high numbers of false positives or not. The output is sorted only by factor name and includes a matching quality score for each site. There is no capability to select individual matrices for searching.

2.2.2.2. MatInspector

The MatInspector ***(37)*** program comes with a precompiled database of more than 250 matrices, also partitioned into organism classes, and is available for the widest number of computer platforms. The program is available in Macintosh, DOS, Windows 95 and NT, and UNIX versions as well as being accessible by the web. All of these are available through a common web site: http://www.gsf.de/biodv/matinspector.html.

MatInspector has a well-designed user interface, and allows the user flexibility in choosing core (the four best conserved nucleotides) and matrix similarity cutoff scores. The user can also select individual matrices or selected groups of matrices. The output can be sorted by matrix name or by sequence location and matching matrices are linked to factor entries in the TRANSFAC database. The output also contains scores for both core and matrix matching, and sequence matches are displayed with the core sequence in upper case letters. MatInspector will accept Staden, Fasta, IG, EMBL, GCG, or GCG-RSF sequence formats.

2.2.2.3. ConsInspector

ConsInspector ***(34,60)*** is a program related to MatInspector, except that it also analyzes DNA sequence surrounding the binding site; however, the matrix library that is included with the program has only 27 matrices (only 10 vertebrate and five yeast matrices), so it is of little use for general transcription factor binding site searches. However, ConsInspector might prove to be useful if one of the included matrices represents one of particular interest to the investigator. ConsInspector also can allow the user to update a matrix with a user-supplied binding site sequence, the only such program that has this capability.

The program is very similar in look and operation to the MatInspector program, except in the case of the matrix similarity cutoff parameter where there is a region similarity cutoff parameter. Search results are output to separate files for each matrix match, a feature that even with the inclusion of a file viewer, makes analyzing the results difficult, especially when the match results are at the end of each file. The program is available in VMS, UNIX, DOS, Windows 95, and Macintosh versions through the web at http://www.gsf.de/biodv/consinspector.html.

2.2.2.4. TFSEARCH

TFSEARCH (by Yutaka Akiyama, Kyoto University) is available only through the web (http://www.genome.ad.jp/SIT/TFSEARCH.html). At present, there is no publication available describing the program. Several organism classes of matrices are available, all based upon the TRANSFAC matrices. On the web page, it appears that the user might be able to modify a threshold value; however, at this time, this value is preset and not modifiable. The output is mapped to the sequence, and the sites are linked to the corresponding TRANSFAC database entries.

2.2.2.5. TESS—Matrix-Based Search

The Matrix-based Search version of TESS is also available only through the web (http://agave.humgen.upenn.edulutess/tess). TESS uses the TRANSFAC matrix database, divided only into vertebrate, invertebrate, and fungi matrix classes. There is no filter-based searching capability as there is in the Filtered string-based search version of TESS. Other than the use of matrices, the program is very similar to the previously discussed String-based version.

3. Recognition of Eukaryotic Pol II Promoters

The development of programs to recognize and characterize eukaryotic pol II promoters is a new and very active field of investigation, but is still in its infancy. Currently, the programs are on the verge of being useful, but still report far too many false positives and miss too many true positives to be truly useful. There seems to be a nearly linear tradeoff in the currently available programs between the capability to recognize promoters correctly and reporting false positives ***(61)***. At one end of the spectrum, there are programs that report much fewer false positives, but do not recognize most promoter sequences. At the other end of the spectrum, there are programs that recognize a majority of promoter sequences, but report so many false positives that the results are of little or no use (except for providing further evidence of an already suspected promoter). The existing programs fall roughly into two groups: 1) programs that attempt to recognize promoters based upon core promoter elements (TATA box, INR, CAAT box, and possibly some other elements found frequently in promoter sequences) and 2) those that attempt to use the whole ensemble of elements found in promoter sequences (either from a transcription factor binding site database or by statistically significant oligonucleotides, which probably for the most part, represent binding sites).

All of the programs reviewed below are readily available, either as web-based programs or easily installable by anyone with common knowledge of computers.

3.1. Core Promoter-Based Programs

3.1.1. GRAIL

The GRAIL program *(62)* includes a pol II promoter prediction module *(63)*, however, because part of its algorithm is integrated with coding region predication in GRAIL, it cannot be used independently of the rest of the GRAIL program. The promoter prediction module is used as an additional algorithm to aid in the prediction of genes in an attempt to demarcate separate transcription units. Because it is not available as a stand alone program, it is difficult to compare it to other promoter prediction programs. The program scores DNA sequences to TATA, GC box, CAAT box, cap site, and translation start site signals using matrices and distance constraints through a neural network. This promoter prediction program, or module, is trained to recognize only TATA-containing promoters. In addition to potential promoter sequence data, coding sequence data must be supplied in addition to the putative promoter sequence, as the coding region recognition is part of the promoter recognition algorithm. If, in the analysis, GRAIL finds a putative promoter, it is displayed graphically along with any putative core promoter features. GRAIL is available both on the web and as a stand alone program for UNIX computers (XGRAIL). The web version can be found at http://avalon.epm.ornl.gov/Grail-1.3/, and XGRAIL may be obtained by anonymous FTP to ftp://arthur.epm.ornl.gov/pub/xgrail/.

3.1.2. NNPP (Promoter Prediction by Neural Network)

The NNPP program (Reese, http://www-hgc.lbl.gov/homes/reese/doc/euprom-Jmb96-abstract.html) is another neural net-based program only upon the TATA box and Initiator elements. The program is available only through a web page (http://www-hgc.lbl.gov/projects/promoter.html). NNPP, like the GRAIL associated program, is strongly biased toward TATA-containing promoters. It allows users to choose between prokaryotic and eukaryotic promoters, and to set a promoter cutoff score. Results indicate 50 base promoter predictions and a predicted TSS, along with scores for each prediction.

3.2. Transcription Factor Binding Site and Overrepresented Oligonucleotide-Based Programs

3.2.1. PROMOTER SCAN

The PROMOTER SCAN program *(6)* is, based upon the uneven distribution of transcription factor binding sites in promoter vs nonpromoter sequences *(64)* with an additional scoring for the presence of a TATA box based upon the Bucher TATA box weight matrix *(23)*. The binding sites are derived from the last version of the TFD. The latest version of PROMOTER SCAN (version 2.0) is available only through the web (http://biosci.umn.edu/software/proscan/

promoterscan.htm) or can be set up as a local web page (for a fee). The program can identify TATA-containing as well as TATA-less promoters, however, it may favor TATA-containing promoters. Sequences may be submitted in GCG, Staden, or Fasta format and, unfortunately, the analysis is limited to 10,000 bases per scan at this time. This can be very prohibitive in the analysis of large genomic-sized segments of DNA. There are no user modifiable parameters. Results include predicted promoter regions (250 base regions) and a putative transcription start site if a TATA box is found, as well as a relative score.

This version of PROMOTER SCAN also includes an optional module for performing further analysis on predicted promoter sequences. If the program predicts a sequence to contain a promoter, it will then 1) compare the predicted promoter to the EPD to identify other promoters with a similar pattern of binding sites and 2) provide a list of binding sites that are common between the predicted promoter sequence and similar promoters in the EPD. This part of the program is still in early experimental stages, but may be useful in identifying significant regulatory features in novel promoter sequences.

3.2.2. TSSG/W

The TSSG and TSSW programs (V. Solovyev, http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html) are based upon the methods used by the PROMOTER SCAN program except for an improved method for locating the transcription start site using a linear discriminant function to combine sequence motifs and sequence composition of the start site (http://dot.imgen.bcm.tmc.edu:9331/gene-finder/Help/tssg.html). The original PROMOTER SCAN program did not include a function to more closely delineate the TSS (it did not score for a TATA box, as it does now), so the TSSG and TSSW programs were developed to aid in locating the TSS with better precision. There is no information available at this time on which method might be better: The Bucher TATA box weight matrix used by the current version of PROMOTER SCAN or the linear discriminant function used by the TSSG/W programs. The TSSG program is based upon the TFD, whereas the TSSW program is based upon the TRANSFAC database.

The TSSG/W programs can be accessed through the web or by an E-mail server (service@bchs.uh.edu or services@bioinformatics.weizmann.ac.il with the subject line "tssg"). Sequences must be in Staden format, and the web version is limited to only 7000 bases. In addition to promoter predictions, the results include a list of transcriptional elements in the predicted promoters.

3.2.3. FunSiteP

Like GRAIL, AutoGene (version 1.1) contains a module for the prediction of promoter sequences ***(65)***, called FunSiteP. AutoGene is available via anony-

mous FTP to ftp.bionet.nsc.ru, in directory pub/biology/au. The program is a little difficult to install, needing both the aug.zip and install.exe programs. It comes without any installation instructions. The aug.zip must be unzipped using the "pkunzip -d" command, then the install.exe program must be run to complete installation of the program.

The FunSiteP module is also based upon the uneven distribution of binding sites in promoter versus nonpromoter sequences, but uses binding sites derived from a text compilation of binding sites by Faisst and Meyer ***(66)*** rather than one of the more established binding site databases.

The program will accept Staden, GenBank, or EMBL formatted sequences. Once the analysis has been completed for gene analysis by Autogene, the promoter prediction results can be viewed either in graphic or text format. The results include only the positions of the predicted promoter sequences and scores. The authors report to be working on a promoter classification scheme ***(63)***, but this has not yet been implemented in this AutoGene module.

3.2.4. Promoter1.0

Promoterl.0 (S. Knudsen, unpublished) is now associated with the GeneID program ***(67)***. The program was available in the past as a stand-alone UNIX version, but is apparently no longer so, being available currently only with the associated GeneID program via the GeneID E-mail server. Access to the program can be gained only by sending a sequence in Fasta format to geneid@darwin.bu.edu. Turn around time on the analysis results may take several hours to overnight. A web version is reportedly being developed.

Promoterl.0 uses an unusual and unique algorithm in which simulated DNA binding proteins are created *in computero* and computationally evolved by mutation to bind promoter sequence elements whereas not binding nonpromoter sequences (S. Knudsen, unpublished) based upon molecular dynamics principles. No real transcription factors are used. Whereas an interesting and potentially promising algorithm, it appears to have a very high false positive rate in its present form ***(6)***, and the program appears to be no longer supported.

3.2.5. PromFind

The PromFind program ***(68)*** takes yet another unusual approach in an attempt to solve the problem of relying on the quality and quantity of transcription factor binding sites in databases by basing promoter prediction on hexamer frequencies in promoter sequences.

The program is available only in an MS-DOS version available by anonymous FTP (iubio.bio.indiana.edu, in directory molbio/ibmpc, file profinll.exe). The program has a primitive command line interface and sequences are

accepted in EMBL, GenBank, or Fasta formats. Results give only the location of the putative promoter in the input sequence.

PromFind is also unusual because it always predicts one, and only one, promoter in a given input sequence (or two promoters, one for each strand, if both stands are analyzed; by default, only the top strand is analyzed), whether there is one or more real promoters in the sequence or none at all. This is owing to one of the theoretical aspects of the program that assumes, in essence, that the highest scoring region of any given input sequence is a promoter sequence. This makes it both difficult to judge what length the input sequence should be in an analysis (for example, in a genome sequencing project), and makes it very difficult to compare to other promoter prediction programs. However, it might prove to be most useful if an investigator suspects that there is a promoter in a given sequence by some other evidence, and uses PromFind to locate the promoter.

3.2.4. PromFD

PromFD ***(69)*** is a program that uses a combined method of looking for both over-represented binding sites (based upon the IMD) and looking for overrepresented 5–l0 bp oligonucleotides in DNA sequences to predict promoter sequences. The program is available by anonymous FTP as a UNIX source code version (beagle.colorado.edu, in directory pub, file PromFD.tar) and work on a web version that is reportedly being completed at this time (http://beagle.colorado.edu/~chenq). There is a difference between the UNIX and web versions: the web version does not perform the IMD part of the analysis, basing its prediction only on overrepresented oligonucleotides. However, the author states that this has little effect on the results. In the web version, it is not clear what sequence formats are accepted because the web page links to the sample sequences did not work, and it is not indicated in any instructions anywhere else. In addition, the web version comes as a Java-based program, and you must set up a special directory on your own system to hold the sequence information. In addition, the system must be a web server, and you must give the URL address of your sequences file to the program. With this program, you cannot just enter sequence data via a web page as with all other such programs. This makes the program much more inconvenient and difficult to use than similar promoter prediction web pages. At the time of this writing, the web version was not working, but is expected to be working soon.

The UNIX version uses a primitive command line interface, and accepts sequence data only in Staden format with a minimal length of 500 bases. The UNIX version source code needed some modifications to compile on our system (Sun Solaris 2.5.1), and at this time is still not working properly.

An additional problem with PromFD is the lack of a strand-specific promoter prediction. The program will predict a promoter in a region of sequence, but without reference to the strand. This leads to a problem of deciding on which strand the promoter is located, and in what orientation transcription would start (and hence the location of any associated gene). In effect, the program makes two transcription start site predictions for each promoter prediction, one start site on each end of the predicted promoter region. Because of this problematic method of reporting predictions, it is not possible to compare its results with any other current promoter prediction program ***(61)***. If any comparison is done with other programs, each PromFD promoter prediction should count as two promoter predictions, one on both strands. This would drop the reported PromFD true positive and false positive rates to half of the reported rates ***(69)***.

3.3. Predicting Promoter Regulation and Recognizing Specific Types of Promoters

Being able to identify novel pol II promoters in genomic sequences is certainly valuable from the perspectives of gene identification and demarcating the 5' end of a gene; however, it would be much more valuable to be able to characterize the putative regulatory properties of predicted promoter sequences. If we had such a capability, it would aid in the identification of specific target genes and regulatory families of genes (genes regulated in a similar manner), and aid in the elucidation of regulatory networks (genes that are coregulated or participate in a regulatory cascade). However, at this time, research and development of software in this area is only beginning. Presently, there are only a couple programs available that attempt promoter regulatory classification, and a few, generally unavailable, programs that can identify one very specific type of promoter sequence (these programs were developed only for theoretical tests, and never fully developed into useful programs).

The first example of programs able to identify specific regulatory types of promoter sequences were those programs developed by Claverie and Sauvaget ***(26)*** to identify glucocorticoid and heat-shock responsive promoters. The programs were given sequence signals (binding sites), distances between signals, and homology levels to create a search pattern. The pattern was then used to search sequence databanks. A more complex pattern-searching program was developed by Staden ***(27)*** and tested on identifying globin family promoters. The program could also use a number of small motifs (such as binding sites and repeat elements, like the Claverie and Sauvaget model), but could be combined using a more complex set of logical operators: AND, OR, and NOT. Also in the pattern was a distance range between each motif; up to 20 motifs could be used to build a single pattern. This program is included in the Staden Package of sequence analysis programs (http://www.mrc-lmb.cam.ac.uk/

pubseq/). The major disadvantages of both the Claverie and Sauvaget, and Staden programs is that the user must manually define each search pattern (each signal [consensus or matrix], range between signals, and logical relationships [AND, OR, NOT; Staden]). No precompiled libraries of patterns are available.

A more recent effort using this approach has been the development of binding site weight matrices for muscle transcription factors MEF2 and myogenin ***(40,70)*** (binding site matrices and the C source code are available from the author). Fickett found that, at least in the case of MEF2 and myogenin binding sites, the distance relationship between the elements appears to be related to the helical alignment of the elements on the DNA helix; this was used in addition to the individual weight matrices to predict muscle-specific promoters.

If programs of this class, programs that rely on specific patterns of elements, are to be generally useful in the search for novel promoter sequences, then reliable and tested libraries of these patterns must first be developed. At this time, there is no such library available. However, software to aid in finding and developing such patterns is beginning to emerge. The GenomeInspector program ***(71,72)***, (with associated program modules such as ConsInspector and MatInspector) can analyze promoter sequences for overrepresentation of, and spatial relationships between, selected binding sites and other genomic features. Using the software to identify distance-correlated sequence motifs, a model can be developed to recognize specific classes of promoter sequences (although functional tests on this have been very limited to date). However, GenomeInspector is not designed to look for these patterns in databanks. Its primary purpose is to develop these patterns in *silico* or to verify experimentally derived patterns. GenomeInspector is only available in a UNIX X Windows version by anonymous FTP to ariane.gsf.de in directory pub/unix/genomeinspector.

Another more advanced set of programs is capable of semiautomatically generating a pattern of binding sites given at least 10 similar promoter sequences and an initial simple model (ModelGenerator) and then using the pattern to scan sequences (ModelInspector) for promoters containing the search pattern ***(73)***. The model that is developed is more advanced than previous models used by other programs: the model contains elements that are *determining*, those found in most or all sequences in the set; and elements that are *nondetermining*, those that are found in only a subset of sequences. Distance information is also included in the model. ModelInspector then uses this pattern by scoring first for determining elements, and those subsequences that score higher than a set threshold are then further scored for nondetermining elements to give a total score. Both programs are available for popular UNIX platforms at http://www.gsf.de/biodv/modelinspector.html, but only one model, for retroviral LTRs, is included.

One potential problem for all of the programs above is that they assume that promoters with a similar regulatory function always have transcriptional elements in the same relative order and spacing. This may be true for some promoters with a shared evolutionary history, but may not hold for all classes of promoters with a similar regulatory regime. It may be that the order of elements is maintained for historical reasons and not functional reasons. It also may be true that element order may be crucial for the function of some pairs or groups of elements, but not the whole functional set of elements in promoters. A second problem faced by all of these programs is that without an extensive precompiled library of patterns, only a few and specific types of promoters may be recognized and classified. It was with these two problems in mind that the approach taken in the PROMOTER SCAN program was developed (http://biosci.umn.edu/software/proscan/promoterscan.htm). The program assumes only a similar composition of binding sites, irrespective of order and spacing of sites. When a promoter is predicted by the basic program, as an option, it will compare the composition of binding sites of the predicted promoter with all known promoters in the EPD, resulting in a set of promoters that are similar in binding site composition. Then a list of common binding sites is displayed. The program is, however, still in preliminary stages.

4. Summary and Prospects

The development of software to analyze eukaryotic pol II promoters is still in its infancy. Software for recognizing individual transcription factor binding sites is abundantly available; yet all programs, regardless of methods or databanks used, result in an unacceptably high number of false positive sites reported. This is most likely a result of attempting to recognize binding sites independently of their context ***(74)***. Much needs to be done in understanding the context of functional binding sites, from an experimental perspective, and including this context information in binding site recognition software. Functional context might include the presence of other binding sites, relative position and helical alignment to other binding sites ***(70)***, DNA structural characteristics ***(75,76)*** such as curvature ***(77,78)***, or other local (or even distant) sequence characteristics.

Promoter recognition software has only recently become available in the last couple of years, and the present programs represent the initial efforts of the various research groups involved in the effort. An interesting result by the recent comparison of the various available programs by Fickett and Hatzigeorgiou ***(61)*** although based on a very limited test set, indicate that all of today's programs fall along the same line in which there is a tradeoff in the increase of true positive recognition for an increase in the number of false positives reported. Because one can usually vary this tradeoff within each pro-

gram by adjusting acceptable promoter cutoff scores and other parameters, it seems from the results reported by Fickett and Hartzigeorgiou that the specific algorithm used in the analysis may not be very significant. In addition, the available programs use a variety of training-set data based upon various transcription binding sites databases, various size oligonucleotides, and simulated transcription factors. It seems that results are comparable between the current programs no matter the algorithm or the training sets used. This result suggests the possibility that more than just the primary DNA sequence alone is involved in the recognition of promoters by the in vivo transcription machinery. Other elements important to promoter recognition may include DNA secondary and tertiary structure, nucleosome binding or the binding of other proteins that may complex with DNA and make it available or unavailable for binding of factors, or some other mechanism such as genomic imprinting that may determine the availability of promoters for active transcription. Some of these elements might be predictable (DNA secondary structure), whereas others may not be predictable (genomic imprinting).

The most valuable aspect of promoter recognition, that of functional classification and characterization, is just beginning. There are a few programs available that are taking the first immature steps in this direction, but at this time there is no software available that can classify promoters effectively.

References

1. Watson, J. D. (1990) The human genome project: past, present, and future. *Science* **248,** 44–49.
2. Cantor, C. R. (1990). Orchestrating the human genome project. *Science* **248,** 49–51.
3. Watson, J. D. and Cook-Deegan, R. M. (1991). Origins of the human genome project. *FASEB* **5,** 8–11.
4. DeLisi, C. (1988) Computers in molecular biology: current applications and emerging trends. *Science* **240,** 47.
5. Burks, C., Myers, E., and Stormo, G. (1990) Final report: workshop on detecting genes and other components of genomic structure. Unpublished Rep., Los Alamos National Lab., LA-UR-91-1713.
6. Prestridge, D. S. (1995) Predicting Pol II promoter sequences using transcription factor binding sites. *J. Mol. Biol.* **249,** 923–932.
7. Dynan, W. S. and Tjian, R. (1985) Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature* **316,** 774–778.
8. Jones, N. C., Rigby, P. W. J., and Ziff, E. B. (1988) Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. *Genes Dev.* **2,** 267–281.
9. Mitchell, P. J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245,** 371–245.

10. Wingender, E. (1990) Transcription regulating proteins and their recognition sequences. *CRC Crit. Rev. Euitaryotic Gene Express.* **1,** 11–48.
11. Wasylyk, B. (1988) Transcription elements and factors of RNA polymerase B promoters of higher eukaryotes. *Crit. Rev. Biochem.* **23,** 77–120.
12. McKnight, S. and Tjian, R. (1986) Transcriptional selectivity of viral genes in mammalian cells. *Cell* **46,** 795–805.
13. Latchman, D. S. (1991) Eukaryotic transcription factors, Academic, New York.
14. Smale, S. T. (1997) Generality of a functional initiator consensus sequence. *Gene* **182,** 13–22.
15. Stargell, L. A. and Struhl, K. (1996) Mechanisms of transcriptional activation in vivo: two steps forward. *Trends Genet.* **8,** 311–315.
16. Novina, C. D. and Roy, A. L. (1996) Core promoters and transcriptional control. *Trends Genet.* **9,** 351–355.
17. Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987) Regulation of inducible and tissue-specific gene expression. *Science* **236,** 1237–1245.
18. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C., and Chambon, P. (1980) Promoter sequence of eukaryotic protein-coding genes. *Science* **209,** 1406–1414.
19. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) The ovalbumin gene -sequence of putative control regions. *Nucleic Acids Res.* **8,** 127–142.
20. Penotii, F. (1990) Human DNA TATA boxes and transcription initiation sites. *J. Mol. Biol.* **213,** 37–52.
21. Kaufmann, J. and Smale, S. T. (1994) Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev.* **8,** 821–829.
22. Smale, S. T. and Baltimore, D. (1989) The "initiator" as a transcriptional control element. *Cell* **57,** 103–113.
23. Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212,** 563–578.
24. Ghosh, D. (1990) A relational database of transcriptional factors. *Nucleic Acids Res.* **18,** 1749–1756.
25. Prestridge, D. S. (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Comput. Applic. Biosci.* **7,** 203–206.
26. Claverie, J. and Sauvaget, I. (1985) Assessing the biological significance of primary structure consensus patterns using sequence databanks. I. Heat-shock and glucocorticoid control elements in eukaryotic promoters. *Computer. Applic. Biosci.* **1,** 95–104.
27. Staden, R. (1988) Methods to define and locate patterns of motifs in sequences. *Comput. Applic. Biosci.* **4,** 53–60.
28. Prestridge, D. S. and Stormo, G. (1993) SIGNAL SCAN 3.0: new database and program features. *Comput. Applic. Biosci.* **9,** 113–115.
29. Bucher, P., Fickett, J. W., and Hartzigeorgiou, A. (1996) Computational analysis of transcriptional regulatory elements: a field in flux. *Comput. Applic. Biosci.* **12,** 361–362.

30. Staden, R. (1984) Computer methods to locate signals in nucleic acid sequences. *Nucleic Acids Res.* **12,** 505–519.
31. Stormo, G. D. (1988) Computer methods for analyzing sequence recognition of nucleic acids. *Ann. Rev. Biophys. Biophys. Chem.* **17,** 241–263.
32. Stormo, G. D. (1990) Consensus patterns in DNA. *Methods Enzymol.* **183,** 211–220.
33. Cornish-Bowden, A. (1985) Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Res.* **16,** 3021–3030.
34. Frech, K., Herrmann, G., and Werner, T. (1993) Computer-assisted prediction, classification, and delimitation of protein binding sites in nucleic acids. *Nucleic Acids Res.* **21,** 1655–1664.
35. Stormo, G. D. and Hartzell, G. W. (1989) Identifying protein-binding sites from unaligned DNA fragments. *Proc. Natl. Acad. Sci. USA* **86,** 1183–1187.
36. Hertz, G. Z., Hartzell, G. W., and Stormo, G. D. (1990) Identification of consensus patterns in unaligned DNA sequences known to be functionally related. *Computer. Applic. Biosci.* **6,** 81–92.
37. Berg, O. G. and von Hippel, P. H. (1987) Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters. *J. Mol. Biol.* **193,** 723–750.
38. Berg, O. G. and von Hippel, P. H. (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J. Mol. Biol.* **200,** 709–723.
39. Quandt, K., Frech, K. Karas, H., Wingender, E., and Werner, T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23,** 4878–4884.
40. Fickett, J. W. (1996a) Quantitative discrimination of MEF2 sites. *Mol. Cell Biol.* **16,** 437–441.
41. Claverie, J. M. and Audic, S. (1996) The statistical significance of nucleotide position-weight matrix matches. *Comput. Applic. Biosci.* **12,** 431–439.
42. Ghosh, D. (1991) New developments of a transcription factors database. *TIBS* **16,** 445–447.
43. Ghosh, D. (1992) TFD: The transcription factors database. *Nucleic Acids Res.* **20,** 2091–2093.
44. Ghosh, D. (1993) Status of the transcription factors database. *Nucleic Acids Res.* **21,** 2091–2093.
45. Hoeck, W. G. (1994) InfoTrac TFD: a microcomputer implementation of the Transcription Factor Database with a graphical user interface. *Comput. Applic. Biosci.* **10,** 323–327.
46. Prestridge, D. S. (1996) SIGNAL SCAN 4.0: additional databases and sequence formats. *Comput. Applic. Biosci.* **12,** 157–160.
47. Wingender, E. (1988) Compilation of transcription regulating proteins. *Nucleic Acids Res.* **16,** 1879–1902.
48. Wingender, E. (1994) Recognition of regulatory regions in genomic sequences. *J. Biotechnol.* **35,** 273–280.

49. Wingender, E., Dietze, P., Karas, H., and Knuppel, R. (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* **24,** 238–241.
50. Wingender, E., Kel, A. E., Kel, O. V., Karas, H., Heinemeyer, T., Dietze, P., Knuppel, R., Romaschenko A. G., and Kolchanov N. A. (1997) TRANSFAC, TRRD and COMPEL: towards a federated database system on transcriptional regulation. *Nucleic Acids Res.* **25,** 265–268.
51. Wingender, E., Heinemeyer, T., and Lincoln, D. (1991) Regulatory DNA sequences: predictability of their function. *Adv. Mol. Gen.* **4,** 95–108.
52. Knuppel, R., Dietze, P., Lehnberg, W., Frech, K., and Wingender, E. (1994) TRANSFAC retrieval program: a network model database of eukaryotic transcription regulating sequences and proteins. *J. Comput. Biol.* **1,** 191–198.
53. Chen, Q., Hertz, G. Z., and Stormo, G. D. (1995) MATRIX SEARCH 1.0: a computer program that scans DNA sequences for transcriptional elements using a database of weight matrices. *Comput. Applic. Biosci.* **13,** 29–35.
54. Kel, O. V., Romaschenko, A. G., Kel, A. E., Wingender, E., and Kolchanov, N. A. (1995a) A compilation of composite regulatory elements affecting gene transcription in vertebrates. *Nucleic Acids Res.* **23,** 4097–4103.
55. Wingender, E., Karas, H., and Knuppel, R. (1997) TRANSFAC database as a bridge between sequence data libraries and biological function. *Proc. 1997 Pacific Symp. Biocomput., World Scientific, Singapore,* pp. 477–485.
56. Kel, O. V., Romaschenko, A. G., Kel, A. E., Naurnochkin, A. N., and Kolchanov, N. A. (1995b) *Proc. 28th Ann. Hawaii Int. Conf. System Sciences HICSS7,* Biotechnol. Comput., IKE Computer Society Press, Los Alamitos, CA, **5,** 42–51.
57. Bucher, P. and Trifonov, E. N. (1986) Compilation and analysis of eukaryotic Pol II promoter sequences. *Nucleic Acids Res.* **14,** 10,009–10,026.
58. Frech K., Quandt, K., and Werner T. (1997a) Finding protein-binding sites in DNA sequences: the next generation. *Trends Biochem Sci.* **22,** 103–104 .
59. Frech K., Quandt, K., and Werner T. (1997b) Software for the analysis of DNA sequence elements of transcription. *Comput. Applic. Biosci.* **13,** 89–97.
60. Frech, K., Dietze, P., and Werner, T. (1997c) ConsInspector 3.0: new library and enhanced functionality. *Comput. Applic. Biosci.* **13,** 109–110.
61. Fickett, J. W. and Hartzigeorgiou, A. G. (1997) Eukaryotic promoter recognition. *Genome Res.* **7,** 861–878.
62. Uberbacher, E. C. and Mural, R. J. (1991) Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc. Aatl. Acad. Sci.* **88,** 11,261–11,265.
63. Matis, S., Xu, Y., Shah, M., Guan, X., Einstein, J. R., Mural, R., and Uberbacher, E. (1996) Detection of RNA polymerase II promoters and polyadenylation sites in human DNA sequence. *Comput. Chem.* **20,** 135–140.
64. Prestridge, D. S. and Burks, C. (1993) The density of transcriptional elements in promoter and non-promoter sequences. *Human Mol. Genet.* **2,** 1449–1453.

65. Kondrakhin, Y. V., Kel, A. E., Kolchanov, N. A., Romashchenko, A. G., and Milanesi, L. (1995) Eukaryotic promoter recognition by binding sites for transcription factors. *Comput. Applic. Biosci.* **11,** 477–488.
66. Faisst, S. and Meyer, S. (1992) Compilation of vertebrate-encoded transcription factors. *Nucleic Acids Res.* **20,** 3–26.
67. Guigo, R., Knudsen, S., Drake, N., and Smith, T. (1992) Prediction of gene structure. *J. Mol. Biol.* **226,** 141–157.
68. Hutchinson, G. B. (1996) The prediction of vertebrate promoter regions using differential hexamer frequency analysis. *Comput. Applic. Biosci* **12,** 391–398.
69. Chen, Q. K., Hertz, G. Z., and Stormo, G. D. (1997) PromFD 1.0: a computer program that predicts eukaryotic pol Il promoters using strings and IMD matrices. *Comput. Applic. Biosci.* **13,** 29–35.
70. Fickett, J. W. (1996b) Coordinate positioning of MEF2 and myogenin binding sites. *Gene* **172,** GC19–GC32.
71. Quandt, K., Grote, K., and Wemer, T. (1996) GenomeInspector: a new approach to detect correlation patterns of elements on genomic sequences. *Comput. Applic. Biosci.* **12,** 405–413.
72. Quandt, K., Grote, K., and Werner, T. (1996b) GenomeInspector: basic software tools for analysis of spatial correlations between genomic structures within megabase sequences. *Genomics* **33,** 301–304.
73. Frech, K. and Werner, T. (1997) Specific modelling of regulatory units in DNA sequences. *Proc. 1997 Pacific Symp. Biocomput., World Scientif., Singapore,* pp. 151–162.
74. Trifonov, E. N. (1996) Interfering contexts of regulatory sequence elements. *Comput. Appl. Biosci.* **12,** 423–429.
75. Karas, H., Knuppel, R., Schulz, W., Sklenar, H., and Wingendler, E. (1996) Combining structural analysis of DNA with search routines for the detection of transcription regulatory elements. *Comput. Applic. Biosci.* **12,** 441–446.
76. Benham, C. J. (1996) Computation of DNA structural variability—a new predictor of DNA regulatory regions. *Comput. Applic. Biosci.* **12,** 375–381.
77. Shpigelman, E. S., Trifonov, E. N., and Bolshoy, A. (1993) CURVATURE: software for the analysis of curved DNA. *Comput. Applic. Biosci. 9,* 432i-440.
78. Ponomarenko, M. P., Ponomarenko, J. V., Kel, A. E., and Kolchonov, N. A. Search for DNA conformational features for functional sites. Investigation of the TATA box. *Proc. 1997 Pacific Symp. Biocomput., World Scientific, Singapore,* 340–351.
79. Genetics Computer Group (1996). Program Manual for the GCG Package, Version 9. 575 Science Drive, Madison, WI 53711.
80. Schug, J. and Overton, G. C. (unpublished) TESS: Transcription Element Search Software on the WWW. Tech. Rep. CBIL-TR-1997-1001-v0.0, of the Computat. Biol. Informatics Lab., School of Medicine, Univ. Pennsylvania, 1997 URL: http://agave.hurngen.upenn.edu/tess/index.html.