Drosophila RNA polymerase II elongation factor DmS-II has homology to mouse S-II and sequence similarity to yeast PPR2

Tricia K. Marshall, Hongliang Guo and David H. Price*

Department of Biochemistry, University of Iowa, Iowa City, IA 52242, USA

Received July 24, 1990; Revised and Accepted September 28, 1990

EMBL accession no. X53670

ABSTRACT

DmSII is a *Drosophila* RNA polymerase II elongation factor which suppresses pausing by RNA polymerase II at specific sites on double stranded templates. Using antibodies produced against the purified protein, a *Drosophila* cDNA expression library was screened and a cDNA was isolated which encoded a portion of DmSII. When this cDNA was used to probe *K_c* cell mRNA the predominant species was found to be 1.4 kb in length. The original cDNA was used to screen a *Drosophila K_c* cell cDNA library resulting in the isolation of a 1.4 kb cDNA which was then sequenced. The deduced protein sequence for DmSII exhibited high similarity to mouse SII protein sequence. In addition, significant sequence similarity was found with the protein encoded by the yeast gene PPR2, which is involved in regulation of URA4 gene expression. The comparison of amino acid sequences suggests that DmSII is comprised of two domains homologous to mouse SII separated by a flexible, serine rich region of low homology. The shorter yeast protein has sequence similarity only to the carboxy terminal domain.

INTRODUCTION

Numerous factors involved in the control of elongation by RNA polymerase in *E. coli* have been identified and studied (1). Efforts to identify analogous eucaryotic factors have met with limited success. So far, only two types of factors have been described. One type, exemplified by the mouse protein SII, was found to stimulate the ability of RNA polymerase II to synthesize RNA in the presence of manganese cations (2). Later SII was shown to effect the elongation phase of transcription (3–7). Similar factors have been studied from yeast (8), and *Drosophila* (9). The second type of factor, exemplified by *Drosophila* factor 5 was found to be an initiation factor as well as a factor which was able to stimulate RNA polymerase II under more physiological conditions than SII (10,11). The analogous mammalian factor, TFIIF, (also called RAP30/74 (12) or beta/gamma (13)) has been shown to stimulate elongation (14).

Despite the fact that the mammalian factor, S-II (also called

TFIIS (4)), has been studied for over 10 years, it is not clear how it functions. When S-II is added into reactions containing partially purified initiation factors it enables RNA polymerase II to pass more freely through specific sites at which the polymerase has the propensity to pause (4,5). A similar factor, DmSII, was identified and purified from *Drosophila K_c* cell nuclear extracts (9). Pure DmSII has the intrinsic ability to suppress pausing by RNA polymerase II in a completely defined system (9). Mammalian S-II has been shown to interact with RNA polymerase II and can be found in elongation complexes (3,15). DmSII has only a weak affinity for RNA polymerase II and is not stably bound to elongation complexes in the absence of other factors (9). It is not clear if the differences are species specific or if the conditions used in the various studies are responsible for the differences. S-II has been shown to be phosphorylated, but an effect of phosphorylation on activity has not been demonstrated (16).

In order to resolve some of the observed differences between DmSII and the mammalian factor and to better understand the function of DmSII we have cloned cDNAs coding for the protein. Full length cDNA was obtained and the deduced protein sequence was compared to the mouse S-II protein sequence. The sequence comparison indicates that there are two regions of homology separated by a serine rich region. Sequence similarities were also found between DmSII and a yeast URA4 regulatory protein, PPR2.

MATERIALS AND METHODS

Materials

¹²⁵I-Protein A was provided by the Diabetes and Endocrinology Research Center at University of Iowa.

Purification of DmSII and production of antibodies

DmS-II was purified from *K_c* cell nuclear extracts as previously described (9) except that the phenyl superose step was omitted and a glycerol gradient step (11) was added. This method is not as effective at purification of DmSII as the previously reported method and the final preparation was about 50% pure. The fractions containing DmSII were electrophoresed in a preparative

* To whom correspondence should be addressed

12.5% SDS protein gel and blotted to nitrocellulose. The nitrocellulose containing about 50 ug of DmSII was excised and implanted subcutaneously into a rabbit. After 45 days the rabbit was boosted with about 25 ug of DmSII in Freund's incomplete adjuvant. Antisera were tested for reactivity to DmSII using partially purified DmSII transferred to nitrocellulose from an SDS protein gel. The preimmune antiserum reacted with several proteins in lanes of K_c cell nuclear extract. The immune serum also reacted with a number of proteins, but DmSII was the only protein specific for the immune serum. Using ^{125}I -Protein A and a 1:300 dilution of the immune serum, moderate signals were obtained to DmSII when 3 ul of K_c cell nuclear extract (150 ug protein) was probed.

cDNA libraries and cDNA cloning

Two *Drosophila* cDNA libraries were used. The first was obtained from Tao Hsieh (Duke University) and was constructed using mRNA from 0–12 hr embryos. The cDNA was cloned into lambda gt11 using Eco RI linkers (17). We constructed the second library using mRNA isolated from K_c cells. cDNAs were synthesized using the BRL (Bethesda Research Laboratories) cDNA kit and were cloned into lambda ZAP (Stratagene Inc.) using Eco RI linkers. The average size of the cloned cDNAs was about 1.5 kb.

Fifty plates containing a total of about 1,000,000 plaques of the embryonic cDNA library were screened using the DmSII antiserum at 1:10 dilution and ^{125}I -protein A. It was necessary to use a high concentration of antiserum in order to detect positive clones. The 800 bp cDNA obtained was used to screen the K_c cell cDNA library. Screening with antibody or DNA probes was carried out as described in Ausubel *et al.*, 1989 (18).

Sequence analysis

The GCG sequence analysis program was used to search the Genbank and NBRF databases for sequences similar to DmSII and to align the protein sequences (19). The FastA program was used to compare DmSII to other protein sequences in the NBRF database and TFASTA was used to compare DmSII to the set of amino acids obtained by translating the Genbank nucleic acid database (20). Protein sequence analysis utilized the BestFit (21) and PeptideStructure (22) programs of GCG.

RESULTS

In order to add to our understanding of the function of DmSII, we isolated cDNAs coding for the factor. Our approach was to use antibodies produced against the pure protein to screen a *Drosophila* cDNA expression library.

Cloning DmS-II cDNA

The purification of the DmSII from K_c cells and production of antibodies is described in the Materials and Methods section. Using a protein blot technique the crude antiserum was tested for reaction to purified DmSII and to DmSII present in crude K_c cell nuclear extracts. At 1:300 dilution of the antiserum a good signal was obtained for DmSII in both samples using ^{125}I -protein A to localize the rabbit antibodies (data not shown). DmSII was the only protein present in crude nuclear extracts which reacted with the immune serum and not the preimmune serum (data not shown). A lambda gt11 library containing early *Drosophila* embryonic cDNAs (17) was screened with the immune serum and several positive plaques were obtained.

During rescreening several plaques were found to react to both the preimmune and immune sera, but one was found to react with only the immune serum. Nitrocellulose was used to absorb proteins from large cleared plates of phage expressing fusion

Fig. 1 RNA blot. The indicated amounts of K_c cell cytoplasmic poly (A) plus RNA were electrophoresed on a 1% agarose/formaldehyde gel. The RNA was blotted on nitrocellulose and probed with the 800 bp cDNA insert from pDmSII-1 which was labeled by random priming. RNA markers were obtained from BRL (Bethesda Research Laboratories).

```

1  GGAATTCGGTATTCTTGACAAGCCACCGCCGGTAACTAAACGCGTTAGGGTGT
61  CTGTCAACTTGTCTGCGAATACAGAAACGAAAAATAGTAAACACATTTATTGGAGG
121 GAAAAAAGAAATAGCCCGGCTCCGCGATAAACACTGTGCGAGACCCGAGCCCGATT
181 AAAGCCGACAGCCGACGCCAGTTAACCTAACCTCTGATCAAAATTAAGTGGGCCAAATGA
241 GCGTGGAGGAGGAAGTGTTCGAATCCAAAAAGAGATGAGCAAGATGGCCAGCGACGGCA
301 CAGGACAGGATCAGGCTCTGGACCTGTGAAGGCCCTGCAACGCTTAACATCAATCTCG
361 ACATTCTGACCAAAACGCGCATCGGCATGACGCTAAACGAGCTGCGAAGAGTAGCAAGG
421 ACGACGAGGTGATCGCTCTGGCCAGACGCTGATCAAGAACTGGAACGATTCTCTGCCA
481 GCCCGGCCCAACCCACCGAAGACAGCTCCGCTAAGGAGGGCTCTCCAAACACAGCA
541 GTGCTCGAAGTCCACAGTGGCGCAAGTCGTCTGCTCAATCTCGGCAAGGATAAGT
601 CCAGTTCATCCAGCTCATCCAAGGATAAAGAGAAGAGGCTCGACATCTCTGTCACAGA
661 CCTCGTTTCCCTCCGGCGGATGACAGATGCGGTCCGCATCAATGCCGCAAGTGTGG
721 CCACCGCACTGAAGATTGGCGAAGTGGCGAGGATGCGGTGAGCCGAGGAAATGGCCG
781 CCGAACTGGAAGATGCCATTTACTCCGAGTTCAATAACACGGATATGAAGTACAAGATC
841 GCATTAGTCAAGCGTGGCCAACTCTGAAGAGCCCAAGAAATCTGGATTGCGCGCAACT
901 TTATGTGCGGCGCGCTCACTGCCAAGCAGTGGCCAAATGACGCCGAGGAGATGGCCA
961 GCGATGAGATGAAGAAGTCGCGGAGAGTTCTGCTAAGGAGGCCATCAACGATGCCAGC
1021 TGCCACCGTGCAGGCGACCAAGACCGATCTCTCAAGTGGCCAAAGTGAAGAAGCGCA
1081 ACTGCACCTACACAGCTGACGACTCGTTCGCGGATGAACCTATGACCACTTCGTCTG
1141 TGTGCAACGAGTGGCGCAACCGATGGAAGTCTGCTCACTCCCGAAATTTGTTCCAAAT
1201 ATAAATCAATCCAAATGAGCACACACACATCTTATCTTAAATGAACATTG
1261 TAAAGACACGACGCGAGAAACATTTTGTATCTGCTGCCCACTGGGAGCGGTTTAC
1321 TCTGACCACTTAAACCTTAAACGAGAGAAAGAAATTTCAATAACATTAACTATTGATA
1381 AAACGGAAAAAAA

```

Fig. 2 The sequence of the 1.4 kb DmSII cDNA. The 1.4 kb insert in pDmSII-2 was sequenced using double stranded DNA and the Sanger/dideoxy method. The coding strand is presented with the open reading frame underlined.

proteins which reacted with either immune serum only (positive) or preimmune serum (control). The filters were incubated with the immune serum under conditions used for the original screening and then were washed extensively. The selected antibodies were eluted as described in Kelly *et al.*, 1986 (23) and used to probe protein gel blots of authentic DmSII. A strong signal was obtained with antibodies selected with the positive clone and no signal was obtained with antibodies selected with the control clone (data not shown).

The selected lambda gt11 clone contained an 800 bp Eco RI insert which was subcloned into pSP73 and called pDmSII-1. pDmSII-1 was partially sequenced and was found to have sequence similarity to the previously isolated mouse S-II (24) when the amino acid sequences were compared. Considering this sequence similarity to the mouse SII and the antibody selection results, we concluded that the cDNA encoded DmSII. In order to determine the size of DmSII mRNA an RNA blot was performed in which pDmSII-1 was used to probe *Drosophila* K_c cell mRNA. The predominant mRNA was 1.4 kb in size (Fig. 1).

In order to obtain full length cDNA a library was constructed in lambda ZAP with cDNA made to *Drosophila* K_c cell mRNA. 5 × 10⁵ plaques were screened with the 800 bp insert of pDmSII-1 and about 25 positive plaques were found. Half of these positive phage rescreened and 5 were purified and converted into plasmids. Two plasmids contained a 1.4 kb Eco RI insert. One of these was selected and called pDmSII-2. Both pDmSII-1 and pDmSII-2 were sequenced using the dideoxy method with double stranded DNA. The sequence of both clones contained a poly (A) sequence in similar positions indicating that the shorter of the two clones, pDmSII-1, lacked the 5' end of DmSII mRNA. The sequence of the 1.4 kb insert in pDmSII-2 is presented in Fig. 2. The sequence of the Eco RI linker added during the cloning of the cDNA is present at the 5' end of the sequence. Both strands were sequenced except for a small portion of the 5' untranslated region which was only sequenced in one direction.

Sequence similarity between DmS-II, mouse S-II, and yeast PPR2

Two methods of sequence comparison were used to locate other cloned genes with sequence similarity to DmSII. When the NBRF protein database (updated 12/89) was searched using the GCG FastA method the only significant similarity found was to the mouse S-II sequence. Using a TFastA search of the Genbank DNA database (updated 3/90) another protein, PPR2 (25), was picked up as being significantly similar. The GCG BestFit program was used to align the protein sequences for DmSII, mouse S-II, and PPR2 (Fig. 3). Examination of the similarities between DmSII and SII indicated that there were two regions of high similarity separated by a serine rich region of lower similarity (Fig.3). The serine rich region of DmSII extends from amino acid 84 to 142 (59 residues) and for SII from 81 to 128 (48 residues). The *Drosophila* protein is slightly longer (313 amino acids) than the mouse protein (301 amino acids) and the difference in size of the two serine rich regions accounts for most of the difference in size of the two proteins. The yeast protein is only 128 amino acids and was similar to the carboxy terminus of the two elongation factors (Fig.3).

The protein sequence of DmSII and SII were further analyzed using the PeptideStructure program of GCG. The surface probability plots for both proteins were very similar when either the amino terminal or the carboxy terminal domains were compared (Fig. 4). The serine rich domains of both proteins

contained the region of highest surface probability, but the location of this region within the serine rich region was different in the two proteins. The comparison of the amino terminal and carboxy terminal domains of the two proteins in flexibility probability plots also indicated similarity (Fig. 4). The most striking result was the presence of the highest flexibility probability in the serine rich domain of both proteins.

DISCUSSION

We have cloned cDNAs encoding DmSII, a *Drosophila* RNA polymerase II elongation factor. One cDNA, pDmSII-2, is the same length as the K_c cell mRNA and contains the entire coding region. The deduced protein sequence is 313 amino acid residues in length. The cDNA has only one large open reading frame which codes for a 34.3 kDa protein. This size is smaller than the 36 kDa estimated for the pure protein in SDS protein gels (9).

Drosophila K_c cells express only one predominant mRNA (1.4 kb) coding for DmSII (Fig. 1). In contrast, mouse SII cDNAs detected at least four different mRNAs (2.8, 1.4, 2.4, and 1.7 kb in order of decreasing abundance) when Ehrlich ascites tumor cell mRNA was probed (24). This difference could be due to

Fig. 3 Comparison of *Drosophila* DmSII, mouse S-II and yeast PPR2. (A) The amino acid sequences of the three proteins were compared using the BestFit program of the GCG sequence analysis package. D, DmSII; M, mouse S-II; Y, yeast PPR2. The extent of similarity is based on a transformation of the Dayhoff table which takes into account the frequency of specific amino acid substitutions found in evolutionarily related proteins (32). The normalized version of the table is available in the GCG program in the file SWGapPep.CMP. Solid lines indicate identities (table value of 1.5). Colons indicate highly conservative changes (values between 0.5 and 1.5). Periods indicate semi-conservative changes (values between 0 and 0.5). Amino acids that are identical in all three proteins are in bold face type. Six cysteine residues conserved between DmSII and SII in the carboxy terminal region are underlined. (B) The two homology domains revealed from the sequence comparison are diagrammed. N, amino terminal domain; C, carboxy terminal domain. The two domains are separated by a serine rich region (SR) which is not highly conserved. The yeast protein is similar in sequence to the carboxy terminal domain.

the unusual structure of the cDNA pS-II-2 used in the mouse study. pS-II-2 was shown to have sequences around the translation start site that differed with other SII cDNAs. It is possible that pS-II-2 was derived from a cloning artifact in which an S-II cDNA was inserted into another cDNA. These other sequences may have been responsible for the 2.8, 2.4, and 1.7 kb mRNAs. We obtained one unusual cDNA which was 1.7 kb in length and diverged from the sequence found for the other cDNAs several nucleotides downstream from the termination codon (data not shown). We presume that this cDNA was due to a cloning artifact since no 1.7 kb mRNA was found (Fig. 1). Preliminary *Drosophila* genomic DNA blots suggest that DmSII is coded for by a single copy gene (data not shown) as was found for SII when a cDNA containing only S-II sequences was used (24).

The deduced protein sequence for DmSII indicates that the protein exhibits high sequence similarity to mouse S-II (24). DmSII is highly charged and serine rich. Of 313 residues 51 (16.3%) are basic and 39 (12.4%) are acidic. The calculated pI is 9.62. Comparison of the protein sequences for the two RNA polymerase II elongation factors indicates that there is 56% identity and 72% similarity using the BestFit program of GCG with a gap weight of 3 and a length weight of 0.1. Alignment of the two sequences indicates that there are two domains of sequence similarity separated by a serine rich region (Fig. 3). The carboxy terminus contains the highest degree of similarity with the last 25 amino acids of both sequences being identical

except for one conservative change. The amino terminus is quite similar until amino acid 83 of DmSII. The region from amino acid 84–142 of DmSII is the least well conserved and is comprised of 42% serine residues. Although the analogous region of the mouse protein has low similarity to DmSII it is serine rich. The amino terminal region in DmSII and SII are very similar in size and likewise the carboxy terminal region is very similar in size with the differences being only 2 or 3 residues. The serine rich regions differ in length (59 residues for DmSII and 48 for SII) and account for the overall difference in size of the two proteins. It has been shown that the two RNA polymerase II elongation factors, DmSII and mouse SII, have analogous activities in that both stimulate the overall incorporation of nucleotides in the presence of manganese (2,9) and both suppress pausing at specific sites in the presence of magnesium (4–7,9). Taking into account the similarities in sequence, size, and function, we conclude that DmSII is the *Drosophila* homologue of mouse SII.

The sequence similarity with PPR2, a yeast protein involved in regulation of the URA4 gene (25), is also striking with 57% similarity and 38% identity. The most similar region between PPR2 and DmSII is close to the carboxy terminus. The amino acids which are identical to DmSII are almost entirely residues which are also identical between S-II and DmSII. The three proteins share 42 amino acids. The sequence similarity between mouse S-II and PPR2 was not noticed when the mouse sequence

Fig. 4 Surface probability and flexibility probability of DmSII and SII. Using the GCG PeptideStructure program the amino acid sequences of DmSII and SII were analyzed. The vertical dotted lines indicate the borders of the sequence homology domains shown in Fig. 3.

was published probably because the PPR2 sequence was not present in available protein databases. This similarity was discovered by other workers and reported recently (26). We picked up the sequence similarity to PPR2 only after comparing our protein sequence with the translation of the Genbank nucleic acid database. PPR2 does not contain the entire carboxy terminal domain found in DmSII and SII (Fig. 3).

It is not clear if the PPR2 gene is the yeast analogue of the two RNA polymerase II elongation factors. Although the sequence similarity is quite striking, PPR2 is less than half the size of DmSII and S-II. A protein with similar properties to DmSII and SII has been isolated from yeast and is 37 kDa in size (8). Therefore, it is not likely that PPR2 is the yeast homologue of DmSII and SII. It has been shown that S-I(b), a 24 kDa protein, which is now believed to be similar or identical to the carboxy terminus of mouse S-II (24) has the ability to stimulate RNA polymerase II in a nonspecific transcription assay in which manganese is used. In addition, a 21 kDa proteolytic fragment derived by chymotrypsin cleavage of SII or the 24 kDa protein also has RNA polymerase II stimulatory activity (27). Although PPR2 is smaller than the SII fragments it is possible that PPR2 might have some elongation factor activity. It would be interesting to determine if expression of DmSII or the carboxy terminus of DmSII could rescue yeast deficient in PPR2.

The carboxy terminal domain of both SII and DmSII contains a conserved set of 8 cysteine residues. Six of these are spaced in a repeating motif at the extreme carboxy terminus of the protein. The repeat is comprised of 3 cysteines separated from each other by 2 and 4–6 residues. Acidic residues proceed the second cysteine in each repeat and are present between the second and third cysteine in each repeat. The repeats are separated by 19 residues. The cysteine repeats do not conform exactly to any of the consensus zinc finger motifs found in many DNA binding proteins, but most closely resembles the C₆ consensus (28) found in the DNA binding protein, GAL4 (29). Although the elongation factors suppress pausing at specific sites (4–7,9), no sequence specific binding has been reported for either SII or DmSII. SII does have the ability to interact with DNA, at least non-specifically, as shown by its differential sensitivity to proteolytic digestion in the presence of DNA (30). It is possible that the specificity of the action of the elongation factor is conferred by the interaction of the factor with paused RNA polymerase II. SII has been shown to interact with RNA polymerase II (3) and with a domain of the large subunit of RNA polymerase II (31).

Fragments of mouse SII containing the carboxy terminal domain have RNA polymerase II stimulatory activity (27), but the function of the amino terminal domain and the serine rich region remains unknown. The region calculated to have the greatest probability of being on the surface of both DmSII and SII was found in the serine rich region (see Fig. 4). In addition flexibility probability calculations indicate that the most flexible region of both DmSII and SII is the serine rich region (see Fig. 4). With the information now available, we would like to propose a model for the structure of DmSII in which two conserved domains are connected by a flexible, serine rich hinge. A model was proposed earlier in which mouse SII was divided into two domains (24). Our model suggests that these two domains are separated from each other (see Fig. 3, B). The sequence alignment indicates fairly precise boundaries for each domain (Fig. 3, A). SII has been shown to be a phosphoprotein (16) with the amino terminal half containing the phosphates. It is not clear if the serine rich region or the amino terminal domain is

phosphorylated. It is possible that phosphorylation of the serine rich region or the amino terminal region is responsible for mediating the action of the carboxy terminal domain. Studies using DmSII expressed in bacteria are currently underway which will address functional aspects of the domain structure of DmSII.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institutes of Health (RO1 GM35500) and a grant from the Diabetes and Endocrinology Center at University of Iowa (DK25295). We would like to thank Dan Weeks and Marc Wold for critical review of the manuscript. Ann Sluder assisted in purification of DmSII which was carried out in the laboratory of Arno Greenleaf at Duke University (9).

REFERENCES

- Yager, T.D., and von Hippel, P.H. (1987) In Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., and Umberger, H.E., (eds.), *Escherichia coli* and *Salmonella typhimurium*. American Society for Microbiology, Washington, D.C., pp. 1241–1275.
- Sekimizu, K., Nakanishi, Y., Mizuno, D., and Natori, S. (1979) *Biochemistry* **18**(8), 1582–1587.
- Horikoshi, M., Sekimizu, K., and Natori, S. (1984) *J. Biol. Chem.* **259**(1), 608–611.
- Reinberg, D., and Roeder, R.G. (1987) *J. Biol. Chem.* **262**, 3331–3337.
- Rappaport, J., Reinberg, D., Zandomeni, R., and Weinmann, R. (1987) *J. Biol. Chem.* **262**, 5227–5232.
- Reines, D., Chamberlin, M.J., and Kane, C.M. (1989) *J. Biol. Chem.* **264**, 10799–10809.
- SivaRaman, L., Reines, D., and Kane, C.M. (1990) *J. Biol. Chem.* **265**, 14554–14560.
- Sawadogo, M., Sentenac, A., and Fromageot, P. (1980) *J. Biol. Chem.* **255**(1), 12–15.
- Sluder, A.E., Greenleaf, A.L., and Price, D.H. (1989) *J. Biol. Chem.* **264**, 8963–8969.
- Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1987) *J. Biol. Chem.* **262**, 3244–3255.
- Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989) *Mol. Cell Biol.* **9**, 1465–1475.
- Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J. (1988) *Molecular and Cellular Biology* **8**(4), 1602–1613.
- Conaway, J.W., and Conaway, R.C. (1989) *J. Biol. Chem.* **264**, 2357–2362.
- Flores, O., Maldonado, E., and Reinberg, D. (1989) *J. Biol. Chem.* **264**, 8913–8921.
- Sopta, M., Carthew, R.W., and Greenblatt, J. (1985) *J. Biol. Chem.* **260**(18), 10353–10360.
- Sekimizu, K., Kubo, Y., Segawa, K., and Natori, S. (1981) *Biochemistry* **20**, 2286–2292.
- Nolan, J.M., Lee, M.P., Wyckoff, E., and Hsieh, T.-S. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3664–3668.
- 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*, Greene Publishing Associates and Wiley-Interscience, New York.
- Devereux, J., Haeblerli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
- Pearson, W.R., and Lipman, D.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
- Smith, T.F., and Waterman, M.S. (1981) *Advances in Applied Mathematics* **2**, 482–489.
- Jameson, B.A., and Wolf, H. (1988) *CABIOS* **4**(1), 181–186.
- Kelly, J.L., Greenleaf, A.L., and Lehman, I.R. (1986) *J. Biochem* **261**(22), 10348–10351.
- Hirashima, S., Hirai, H., Nakanishi, Y., and Natori, S. (1988) *J. Biol. Chem.* **263**(8), 3858–3863.
- Hubert, J.-C., Guyonvarch, A., Kammerer, B., Exinger, R., Liljelund, P., and Lacroute, F. (1983) *EMBO J.* **2**, 2071–2073.
- Davies, C.J., Trgovcich, J., and Hutchison, C.A. (1990) *Nature* **345**, 298–298.
- Horikoshi, M., Sekimizu, K., Hirashima, S., Mitsuhashi, Y., and Natori, S. (1985) *J. Biol. Chem.* **260**(9), 5739–5744.
- Evans, R.M., and Hollenberg, S.M. (1988) *Cell* **52**, 1–3.

29. Corton, J.C., and Johnston, S.A. (1989) *Nature* **340**, 724–727.
30. Horikoshi, N., Sekimizu, K., and Natori, S. (1990) *J. Biol. Chem.* **265**, 11854–11857.
31. Rappaport, J., Cho, K., Saltzman, A., Prenger, J., Golomb, M., and Weinmann, R. (1988) *Mol. Cell Biol.* **8**, 3136–3142.
32. Schwartz, R.M., and Dayhoff, M.O. (1979) In Dayhoff, M.O. (ed.), *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington, D.C., pp. 353–358.