Nucleic Acids Research, Vol. 19, No. 18 5037–5043

Comparison of several promoters and polyadenylation signals for use in heterologous gene expression in cultured *Drosophila* cells

Monica L.Angelichio*, Jeffery A.Beck, Hanne Johansen, and Mona Ivey-Hoyle+
Department of Gene Expression Sciences, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, PA 19406, USA

Received April 26, 1991; Revised and Accepted July 19, 1991

ABSTRACT

We have directly compared the ability of four promoters and three polyadenylation (poly(A)) signals to direct heterologous gene expression in stably transfected *Drosophila melanogaster* S2 cells. We compared two constitutive *Drosophila* promoters, the actin 5C distal promoter and the α1-tubulin promoter, with the tightly regulated *Drosophila* metallothionein (*Mtn*) promoter and the *Bombyx mori* fibroin promoter. We find that the actin 5C and induced *Mtn* promoters generate comparable high levels of RNA and protein in this system. The α1-tubulin promoter generates about four-fold lower levels, and the fibroin promoter shows no detectable activity in S2 cells. Interestingly, genes expressed from the constitutive actin 5C and α1-tubulin promoters are consistently present at three- to four-fold lower copy numbers than genes expressed from the inducible *Mtn* promoter or the inactive fibroin promoter. Poly(A) signals of both mammalian (SV40) and *Drosophila* (*Mtn*) origin efficiently directed stable RNA synthesis in S2 cells, and, as in mammalian cells, the SV40 late poly(A) signal was more efficient than the SV40 early poly(A) signal. Thus the process of polyadenylation appears to be conserved between mammalian and *Drosophila* cells.

INTRODUCTION

Drosophila melanogaster Schneider line 2 (S2) cells have become increasingly utilized over the past few years for the expression of heterologous proteins. High levels of protein expression can easily be achieved using a recently developed procedure whereby a plasmid encoding a gene of interest is cotransfected with a plasmid encoding resistance to hygromycin B. After three weeks of selection, a stable polyclonal cell line can be generated possessing an average of up to 1000 gene copies per cell (1) and expressing as high as 35 mg/l of protein using the *Drosophila Mtn* promoter (2). Proteins of prokaryotic, eukaryotic and viral origin have been expressed in this system and shown to be biologically and biochemically active and processed appropriately (1–4). *Drosophila* S2 cells have been valuable not only for the expression of large amounts of protein for biochemical studies (2, 5) but also for studies of regulation of gene expression *in vivo*. They have particularly been useful as a null background for *in vivo* studies of the activities of regulatory proteins whose functions are not normally expressed in S2 cells, e.g., the HIV-1 Rev protein, mammalian transcription factor Sp1, and *Drosophila* homeodomain proteins (3, 6, 7).

Although there is wide use of *Drosophila* S2 cells in the literature, there are few reports comparing or optimizing transcriptional regulatory signals that can be used to control expression. Mammalian promoters which have been studied in S2 cells function poorly, presumably due to the lack of appropriate mammalian transcription factor homologs (6, 8, 9). However, a few *Drosophila* promoters have been identified which work quite well in S2 cells. It has been stated that the *Drosophila* actin 5C distal promoter (10) is the strongest promoter among certain (unspecified) constitutive promoters in transient expression assays in S2 cells (11). In addition, the inducible *Drosophila* metallothionein (*Mtn*) promoter (12) has also been shown to produce high levels of protein expression in stably transfected S2 cells (2). The *Mtn* promoter allows tightly regulated expression, even at very high copy number, and is thus useful for expression of lethal gene products (4) or when an inducible system is desired for studying the regulation or kinetics of gene expression.

To expand our information about transcriptional regulatory signals that are useful for heterologous gene expression in *Drosophila* S2 cells, we have directly compared expression from four promoters of insect origin: the *Drosophila* actin 5C distal, *Mtn*, and α1-tubulin (13, 14) promoters, and the fibroin gene promoter from the silkworm *Bombyx mori* (15, 16). Since the α1-tubulin and fibroin genes, like actin 5C, encode abundant structural proteins (14, 15), the promoters for these genes, if they function in S2 cells, may be useful for the efficient synthesis of heterologous proteins. We also have investigated the effects of different polyadenylation (poly(A)) signals on heterologous gene expression since these signals are known to affect steady state RNA levels through their role in RNA 3′ end formation (17,

* To whom correspondence should be addressed

+Present address: Merck Sharp and Dohme Research Laboratories, West Point, PA 19486, USA

18). We have directly compared the SV40 early poly(A) signal, the SV40 late poly(A) signal, and the *Drosophila Mtn* poly(A) signal to determine if there are significant differences in the use of mammalian and insect poly(A) signals in S2 cells.

MATERIALS AND METHODS

Plasmid Construction

Plasmid constructions were generated by standard cloning techniques (19). All constructs were analyzed for correct junction sequences by dideoxy sequencing, and PCR-generated fragments were sequenced in their entirety (Sequenase Version 2.0, U.S. Biochemical Corp.). Plasmids pM-tPA-E, pF-tPA-E, and pT-tPA-E were derived from pHsptPA. Plasmid pHsptPA contains an AatII-HindIII fragment containing a *Drosophila* heat shock promoter, a HindIII-SalI fragment from pBPVneotPA encoding human tissue plasminogen activator (tPA) and the SV40 early poly(A) signal (20), and a SalI-AatII fragment derived from pBR322 which confers ampicillin resistance. For the construction of pM-tPA-E, pHsptPA was digested with HindIII, treated with the Klenow fragment of DNA polymerase, and digested with AatII to remove the hsp promoter. An AatII-EcoRI synthetic linker and an EcoRI-StuI fragment from pDM131 (a kind gift of G. Maroni, [21]) containing the *Mtn* promoter (from −380 to +55 relative to the mRNA start site) were inserted. The plasmid pF-tPA-E was derived from pHsptPA by replacing the AatII-HindIII promoter fragment with a 259 bp AatII-HindIII fragment containing the *B.mori* fibroin promoter (position −238 to +16, relative to the mRNA start site, [22]). The AatII and HindIII sites were introduced upstream and downstream, respectively, of the fibroin promoter sequences in pfb5′Δ238 (a kind gift of Yoshiaki Suzuki, [22]) using synthetic oligonucleotides and PCR amplification. Plasmid pT-tPA-E was derived from pHsptPA in the same manner as pM-tPA-E except that an AatII-PstI linker and a 1.6 kb PstI-XmnI fragment from pDMtα1 (a kind gift of Pieter Wensink, [14]) containing the tubulin promoter (from −1300 to +120 relative to the mRNA start site, [13]) were inserted. Plasmid pM-tPA-M was constructed by first digesting pM-tPA-E with XhoI, treating with the Klenow fragment of DNA polymerase, and digesting with SalI, to remove the SV40 early poly(A) signal. The vector fragment was then ligated with a DraIII-SalI synthetic linker and a 231 bp fragment encoding the *Mtn* poly(A) signal which was isolated from pDM131 (a kind gift of G. Maroni [21]) by digesting with BglII, treating with the Klenow fragment of DNA polymerase, and digesting with DraIII. Plasmid pM-tPA-L was derived from pM-tPA-E by replacing a 170 bp XhoI-SalI fragment containing the SV40 early poly(A) signal with a 240 bp XhoI-SalI fragment containing the SV40 late poly(A) signal which was generated by PCR amplification from pDPSV (23). The XhoI and SalI sites were introduced 5′ to the unique BamHI site and 3′ to the unique BclI site, respectively, which flank the SV40 late poly(A) signal in SV40 DNA.

The construction of pM-120L-E was previously described (pMtBaL, [2]). This pBR322-based plasmid contains the *Mtn* promoter, the signal sequence from human tPA gene fused in frame with the coding region for the HIV-1 BaL external envelope glycoprotein gp120, and the SV40 early poly(A) signal. Plasmid pT-120L-E was derived by replacing an AatII-BglII fragment containing the *Mtn* promoter and tPA signal sequence, with a 1.6 kb AatII-BglII fragment containing the tubulin promoter and tPA signal sequence from pT-tPA-E. Plasmid pF-120L-E was also derived by replacing the AatII-BglII fragment from pM-120L-E with a 470 bp AatII-BglII fragment containing the fibroin promoter and the tPA signal sequence from pF-tPA-E. Plasmid pM-120L-L was derived from pM-120L-E by replacing an XhoI-SalI fragment containing the SV40 early poly(A) signal with the 240 bp XhoI-SalI fragment containing the SV40 late poly(A) signal described above. The construction of pM-120H-E was previously described (pMtBH10, [2]). This plasmid is essentially identical to pM-120L-E except that the gp120 coding region is from the HIV-1 isolate BH10 instead of BaL. BH10 sequences are 93% identical to BaL sequences in the gp120 coding region. The BH10 gp120 sequence was used instead of BaL to compare the *Mtn* and actin 5C promoters due to the greater ease of constructing the actin promoter expression unit with BH10 gp120 sequences (pA-120H-E). Plasmid pA-120H-E is essentially identical to pM-120H-E except that the *Mtn* promoter is replaced with a synthetic AatII-EcoRI linker and a 2.5 kb EcoRI-BamHI fragment from the clone pA29 (a kind gift of B. Bond-Matthews, [24]), containing the actin 5C distal promoter (from −2500 to +88 relative to the mRNA start site).

DNA Transfection and Cell Culture

D. melanogaster Schneider 2 cell culture and generation of stably transfected cell lines by a hygromycin B selection method have been described (4). Each plasmid (19 μg) was cotransfected using calcium phosphate co-precipitation (25) with 1 μg of pCOhygro (6), a plasmid containing the hygromycin B selectable marker. For analysis of expression, cells were seeded in fresh medium at 1.5×10^6 cells/ml in a T-flask and allowed to grow for 10 days. Expression from the *Mtn* promoter was induced on day 4 by the addition of 0.5 mM $CuSO_4$ to the medium. These conditions are optimal for expression from the *Mtn* promoter in our system (unpublished data).

Protein Analysis

Samples (4 μl and 12 μl of each culture supernatant) were denatured in Laemmli sample buffer (26) at 90°C for 2 minutes prior to separation on a 10% polyacrylamide-SDS gel and transfer onto nitrocellulose for Western immunoblot analysis (27). Proteins were detected using rabbit antiserum to purified recombinant human tPA (kindly supplied by Zdenka Jonak) or rabbit antisera to purified, *Drosophila*-produced, HIV-1 BH10 gp120 and ^{125}I-labeled protein A (New England Nuclear). Molecular weight markers were prestained protein high molecular weight standards from Bethesda Research Laboratories.

Northern Blot Analysis

Total RNA was prepared as previously described (3). RNA from 1×10^6 cells was separated on a 1% formaldehyde-agarose gel along with RNA size standards (Bethesda Research Laboratories). Hybridization (19) was carried out with tPA or gp120-specific restriction fragments labeled using the Pharmacia Oligolabeling Kit. To compare the amount of RNA loaded in each lane, these probes were removed from the nitrocellulose membranes by treatment in 0.1×SSPE, 0.1% SDS at 90°C, and the membranes were incubated with a 1 kb HindIII fragment encoding the *Drosophila* sgs-5 gene to detect endogenous glue gene RNA. This fragment was derived from pDM384 (a kind gift of Greg Guild, [28]).

Southern Blot Analysis of Integrated DNA

Total DNA was prepared by direct lysis of cells as previously described (4). The DNA was quantitated by measuring

absorbance at 260 nm. The tPA coding sequence was excised from 1 μg of DNA on a 1.2 kb BglII-SacI fragment. The gp120 coding sequence was excised on a 917 bp StuI-SacI fragment. The digested DNA was separated on a 1% agarose gel and transferred to nitrocellulose for hybridization (19). Probes for tPA and gp120 sequences were the same as those used for Northern blot analysis. To compare the amounts of DNA loaded in each lane, the tPA and gp120 probes were removed from the nitrocellulose as described above and the filter was incubated with a 2.5 kb EcoRI-BamHI fragment encoding the actin 5C promoter to detect endogenous actin gene sequences.

Laser Densitometry

Autoradiograms were scanned with a LKB Ultrascan XL laser densitometer.

RESULTS

Expression vectors for comparison of promoters and poly(A) signals

We wished to directly compare the effects of four different promoters and three different poly(A) signals on the expression of heterologous genes in stably transfected *Drosophila* S2 cells. The promoters are the *Drosophila* metallothionein (*Mtn*), actin 5C distal, and α1-tubulin promoters, and the *B.mori* fibroin promoter (Fig. 1). The three poly(A) signals derive from the SV40 early region, the SV40 late region, and the *Drosophila* metallothionein (*Mtn*) gene. The boundaries of the DNA fragments used for the promoters and poly(A) signals were chosen so as to include all of the regulatory sequences known to be involved in the function of each (See Materials & Methods). We have chosen the *Mtn* promoter used in conjunction with the SV40 early poly(A) signal as the standard for our comparisons since all previous expression work using this system employs this combination of signals (2, 3).

We examined the effects of these signals on more than one reporter gene in independent cell lines in order to establish the reproducibility of any differences observed. The use of two different reporter genes also serves to distinguish differences in expression levels which might be gene specific from those which are inherent properties of the expression signals. The reporter genes used encode human tissue plasminogen activator (tPA) or one of two similar human immunodeficiency virus (HIV-1) external envelope glycoproteins (gp120). As shown in Figure 1, five constructs encoding tPA and six constructs encoding gp120 were made. Plasmid constructs are named according to the promoter, reporter gene, and poly(A) region present. Constructs containing the same reporter gene were transfected in parallel experiments using the same amount of plasmid DNA to generate stable polyclonal cell lines. We chose to analyze polyclonal cell lines so that any differences in expression caused by specific chromosomal position effects from one cell to another would be averaged out.

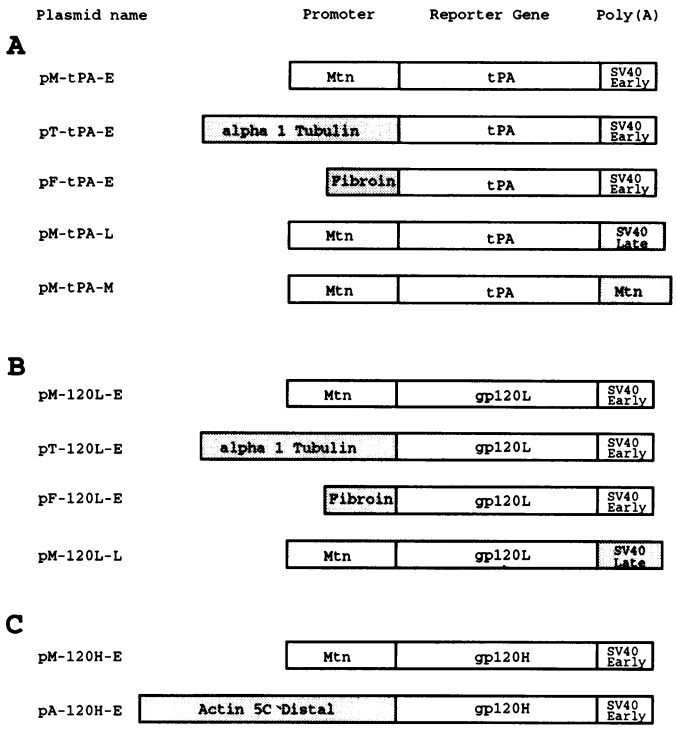

Figure 1. Diagram of transcription units for comparison of different promoters and poly(A) signals. Each of the expression units shown is present on a pBR322-based plasmid. Plasmid names (left) consist of three designations: a letter designating the promoter (M = *Mtn*, T = α1-tubulin, F = fibroin, A = actin 5C distal); the identity of the reporter gene (tPA = tissue plasminogen activator, 120L = HIV-1 BaL gp120, 120H = HIV-1 BH10 gp120); and a letter designating the poly(A) signal (E = SV40 early, L = SV40 late, M = *Mtn*). Plasmids are grouped into three classes, A, B, and C, based on the reporter gene present. Sequences that are different from the prototype constructs shown at the top of each group are indicated by shading.

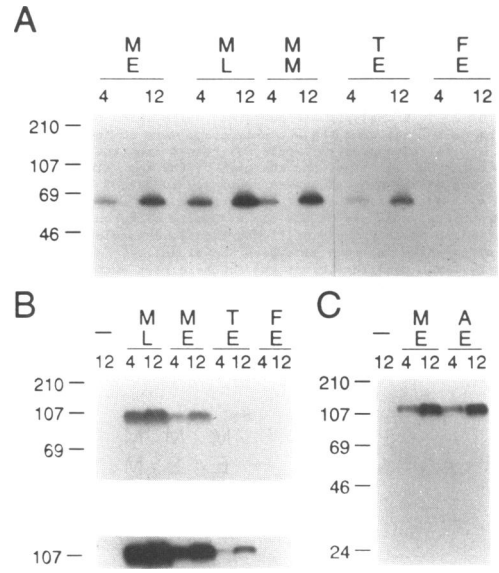

Figure 2. Detection of tPA and gp120 proteins by Western blot analysis. Two different volumes of each culture supernatant were loaded on the gels, as indicated by the numbers above each lane (4 μl and 12 μl). Apparent molecular masses (in kilodaltons) are shown to the left of each panel. A. Culture supernatants from cells transfected with pM-tPA-E (ME), pM-tPA-L (ML), pM-tPA-M (MM), pT-tPA-E (TE), or pF-tPA-E (FE) were analyzed using rabbit polyclonal antiserum to recombinant human tPA. Alternating unlabeled lanes contain 12 μl of culture supernatant from nontransfected cells. B. Culture supernatants from nontransfected cells (−) or cells transfected with pM-120L-L (ML), pM-120L-E (ME), pT-120L-E (TE), or pF-120L-E (FE) were analyzed using rabbit polyclonal antiserum to recombinant HIV-1 gp120. A longer exposure of the same blot is shown at the bottom for better visualization of the gp120 from the pT-120L-E construct. C. Culture supernatants from nontransfected cells (−) or cells transfected with pM-120H-E(ME) or pA-120H-E(AE) analyzed as in B.

Comparison of promoters

Expression of the secreted tPA and gp120 proteins from the cell lines transfected with the different promoter constructs was assayed by Western blot analysis of culture supernatants (Fig. 2), and the data is summarized in Table 1. The results show that the cell line expressing tPA from the *Mtn* promoter produced approximately three-fold more tPA than the cell line employing the α1-tubulin promoter. Similarly, the cell line expressing gp120 from the *Mtn* promoter produced five-fold more gp120 than the cell line using the α1-tubulin promoter. The cell line using the actin 5C distal promoter produced approximately the same level of gp120 as the one using the *Mtn* promoter. No tPA or gp120 protein was observed from the cell lines transfected with the fibroin promoter constructs. Thus regardless of the reporter gene used, the relative levels of protein production follow the same pattern with the different promoter constructs: *Mtn* = actin 5C distal > α1-tubulin > fibroin. These relative protein levels were reproducible not only with different reporter genes in independent cell lines but also with repeated harvests from the cell lines from one week to the next (See legend to Table 1).

TABLE 1. Comparison of Promoters and Polyadenylation Signals

		Relative protein levels[1]		Relative RNA levels[2]		Relative copy number[3]		Relative protein level per gene copy[4]	
		tPA	gp120	tPA	gp120	tPA	gp120	tPA	gp120
Prom.	Poly(A)								
M	E	1.0	1.0	1	1	1.0	1.0	1.0	1.0
T	E	0.37	0.19	0.2	<0.1	0.23	0.23	1.6	0.83
F	E	N.D.	N.D.	N.D.	N.D.	1.3	1.2	N.D.	N.D.
A	E	-	0.93	-	0.8	-	0.34	-	2.7
M	L	1.9	3.2	3	5	0.71	0.75	2.7	4.3
M	M	1.3	-	1	-	0.91	-	1.4	-

N.D., none detected.
Each construct is represented with the promoter and poly(A) signal shown on the left (designations as in Figure 1) and the reporter gene shown at the top.
1. Protein production from the cell lines was compared by western blot analysis, as shown in Fig. 2. This analysis was performed twice using two independent harvests of the cell lines, and the autoradiograms were scanned by laser densitometry. The relative protein levels determined from the two sets of harvests were averaged for each cell line, and the mean of the two values is reported. The standard deviation of the two values was routinely 10-20% of the mean value.
2. Estimated by laser densitometry of the Northern blot autoradiograms in Fig. 3 and other appropriate exposures of the blots. Due to background in the lanes, values are estimated to only one significant figure. The relative RNA levels have been normalized to the endogenous glue gene RNA signal.
3. Estimated by laser densitometry of the Southern blot autoradiograms in Fig. 4 and other appropriate exposures of the blots. The relative copy numbers have been normalized to the endogenous actin gene signal.
4. Relative protein levels divided by the relative copy number.

If the different levels of protein observed reflect the differential abilities of the promoters to direct stable RNA synthesis, then these same relative differences should be observed at the level of RNA. To examine steady state RNA levels, total RNA was prepared from cells that were harvested at the same time that protein production was analyzed, and the RNA was examined by Northern blot analysis (Fig. 3). The relative amounts of gp120- and tPA-specific RNA, normalized to the amount of endogenous glue gene mRNA, are shown in Table 1. For both tPA and gp120 reporters, the cell lines with the *Mtn* promoter clearly produced higher levels of RNA than the cell lines with the α1-tubulin promoter. Similar levels of gp120 RNA are observed from the *Mtn* and actin 5C promoters. However, no RNA was detected from the fibroin promoter for either tPA or gp120 constructs. It is clear from the data summarized in Table 1 that in all cases, relative RNA levels at the time of cell harvest correlate very well with the corresponding protein levels.

Finally, to determine to what extent the RNA and protein levels in these cell lines reflect the differential abilities of the promoters to direct stable RNA synthesis, the copy numbers for the reporter genes were compared by Southern blot analysis (Fig. 4). With the transfection procedure used to generate these cell lines, multiple copies of the gene usually integrate into the chromosome in a head-to-tail array (1). For the genes from the prototype plasmids, copy numbers were approximately 700 for pM-tPA-E, 1000 for pM-120L-E, and 700 for pM-120H-E. Since each cell line represents a polyclonal population in which a range of copy numbers are present, these numbers represent an average copy number over the population. The relative gene copy numbers for all of the cell lines, normalized to the endogenous actin 5C gene signal, are shown in Table 1. Interestingly, the gene copy numbers in both of the cell lines transfected with the α1-tubulin promoter constructs are about four-fold lower than those in cell lines transfected with *Mtn* promoter constructs. Thus, this lower copy number for α1-tubulin promoter constructs appears reproducible regardless of the reporter gene. Similarly, the gp120 gene from the actin 5C promoter construct was present at approximately three-fold lower copy number than the gene from the corresponding *Mtn* promoter construct. In contrast, cell lines transfected with fibroin promoter constructs possessed gene copy numbers very similar to those with the *Mtn* promoter constructs

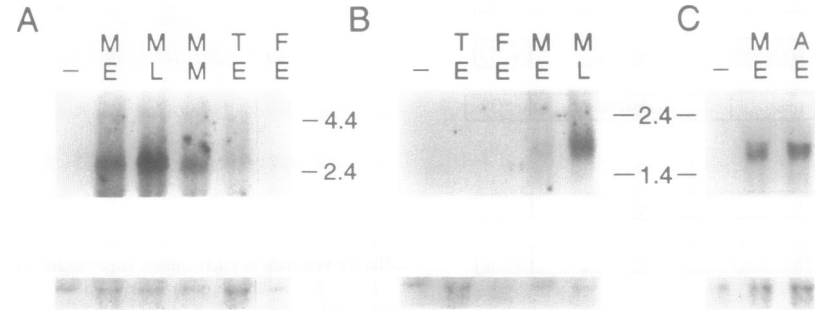

Figure 3. Detection of tPA and gp120 mRNAs by Northern blot analysis. In the top panels, tPA- or gp120-specific probes were used. To compare the amount of RNA loaded in each lane, all three blots were incubated a second time with a probe specific for endogenous glue gene RNA, as shown in the bottom panels. Sizes of RNA markers are given in kb. A. RNA from nontransfected cells (–) or cells transfected with pM-tPA-E (ME), pM-tPA-L (ML), pM-tPA-M (MM), pT-tPA-E (TE) or pF-tPA-E (FE) was analyzed using a 2.0 kb BglII-SalI fragment from pF-tPA-E containing the entire tPA coding region. The predicted size of the tPA RNA (excluding the poly(A) tail) is 2.3 kb. B. RNA from nontransfected cells (–) or cells transfected with pT-120L-E (TE), pF-120L-E (FE), pM-120L-E (ME) or pM-120L-L (ML) was analyzed using a 1.3 kb BglII-StyI fragment from pM-120L-E containing the BaL gp120 coding sequences. The predicted size of the gp120 RNA (excluding the poly(A) tail) is 1.7 kb. C. RNA from nontransfected cells (–) or cells transfected with pM- 120H-E (ME) or pA-120H-E (AE) were analyzed using a 1.3 kb NdeI-StyI fragment from pM-120H-E containing the BH10 gp120 coding region.

for both tPA and gp120. Thus, gene copy number is consistently three- to four-fold lower when a constitutive promoter is used than when an inducible (*Mtn*) or apparently nonfunctional (fibroin) promoter is used.

Comparison of poly(A) signals

The same series of analyses that was used to compare the promoters was used to compare the poly(A) signals. Protein production from the cell lines transfected with the different poly(A) constructs was assessed by Western blot analysis (Fig. 2, Table 1). The cell line using the *Mtn* poly(A) signal produced approximately the same level of tPA as the corresponding cell line using the SV40 early poly(A) signal. However, the cell lines using the SV40 late poly(A) signal produced approximately two-fold more tPA and three-fold more gp120 than the analogous cell lines using the SV40 early poly(A) signal.

These relative differences were also seen in the steady state levels of RNA from these cell lines as detected by Northern blot analysis (Fig. 3, Table 1). Both tPA and gp120 RNA levels in the cell lines with the SV40 late poly(A) constructs were significantly higher (three to five fold) than those in the cell lines with the SV40 early poly(A) constructs. RNA levels from the *Mtn* poly(A) and SV40 early poly(A) cell lines were approximately the same, as were the protein levels.

With regard to gene copy number in these cell lines (Fig. 4, Table 1), there is less variation than was seen with the different promoter constructs. Cell lines transfected with the SV40 late poly(A) constructs show a slightly lower copy number than with the SV40 early poly(A) constructs with both the tPA and the gp120 series. There is little difference in the gene copy number between the *Mtn* and SV40 early poly(A) constructs. Thus for these cell lines, the differences in RNA and protein levels observed reflect primarily differences in the relative abilities of the poly(A) signals to direct stable RNA synthesis.

DISCUSSION

We have directly compared the usefulness of several promoters and poly(A) signals for expression of heterologous gene products in stably transfected *Drosophila* S2 cells. Among the combinations of functional promoters and poly(A) signals tested, up to ten-fold differences in the level of protein production were observed. Two lines of evidence indicate that these different protein levels reflect differences in the inherent abilities of the expression signals to direct stable RNA synthesis. First, in every case examined, we found that the relative effects of the different promoter and poly(A) signals were the same regardless of whether tPA or gp120 was used as the reporter gene. Furthermore, in independent experiments using other reporter genes, we have also seen equivalent production from the actin 5C distal and *Mtn* promoters, and from the SV40 early and *Mtn* poly(A) signals. Thus, the effects of the expression signals on production are reproducible. Second, there was a very good correlation between the relative protein levels and relative RNA levels in every case. This data suggests that the different protein and RNA levels observed likely stem in large part from differential efficiencies with which the promoters and poly(A) signals are used for transcription initiation and RNA 3′ end formation, respectively. However, we note that the various promoter and poly(A) fragments used do contribute different 5′ and 3′ untranslated sequences at the ends of the mRNAs which could potentially affect mRNA stability differently. Thus our steady state RNA and protein levels measure the overall effects of the different promoter and poly(A)-encoding fragments on RNA production. Differences in gene copy number also likely play a role in the different RNA and protein levels observed in some cases, as discussed below.

Promoters

Among the promoters examined, we find that overall production was highest in the cell lines transfected with the actin 5C distal and *Mtn* promoter constructs, which generated comparable levels of RNA and protein in our system. However, since fewer copies of the actin 5C promoter construct were integrated relative to the *Mtn* construct, the actin promoter may be inherently stronger (Compare relative protein level per gene copy in Table 1.). The use of these two efficient promoters together in a single cell line would provide an attractive system to study the regulation of an RNA or protein (expressed constitutively from the actin 5C promoter) by another protein whose expression could be induced from the *Mtn* promoter (3).

Our results also show that the promoter from the α1-tubulin gene of *Drosophila* can be used for constitutive heterologous gene expression in S2 cells. The overall production from the cell lines transfected with the α1-tubulin promoter constructs is three- to

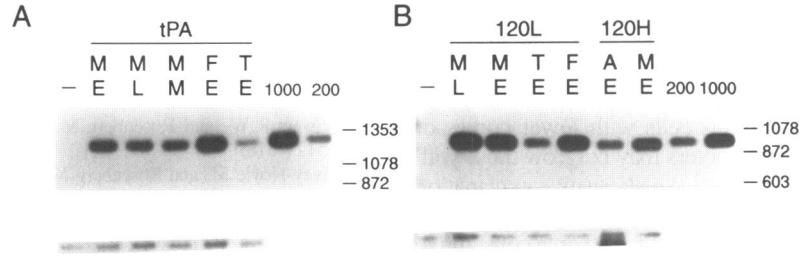

Figure 4. Southern blot analysis of tPA and gp120 gene copy number. Genomic DNA was isolated from nontransfected cells (–) and transfected cell lines (designated as in Figs. 2 and 3), and 1 μg was digested to excise a 1.2 kb tPA-encoding fragment (panel A) or a 917 bp gp120-encoding fragment (panel B). The samples were electrophoresed and blotted to nitrocellulose for analysis with either a tPA-specific probe (top panel A) or a mixture of BaL gp120- and BH10 gp120-specific probes (top panel B). The probes were the same as those used in Fig. 3. The standard lanes contain digested pM-tPA-E (panel A) or pM-120L-E (panel B) plasmid DNA loaded in amounts equivalent to 200 and 1000 gene copies per cell. To compare the amount of DNA loaded in each lane, the blots were also incubated with an actin promoter-specific probe to detect endogenous actin 5C gene sequences, as shown in the bottom panels. The band appearing below the endogenous actin gene band in the pA-120H-E cell line originates from integrated pA-120H-E DNA which contains actin 5C promoter sequences. Sizes of DNA markers (in base pairs) are shown to the right of each panel.

five-fold lower than that from cell lines using the *Mtn* or actin 5C distal promoters. As with the actin 5C distal promoter, the α1-tubulin promoter constructs emerge after selection at lower copy numbers than the analogous *Mtn* promoter constructs. If considered on a per copy basis (Table 1), the α1-tubulin promoter is only two- to three-fold less productive than the actin 5C distal promoter and is comparable to the induced *Mtn* promoter.

In contrast to the *Drosophila* promoters examined, the fibroin promoter from the silkworm *B.mori* failed to show any detectable activity in S2 cells, even though the gene copy numbers for analogous fibroin and *Mtn* promoter constructs were similar. It is known from *in vivo* studies and *in vitro* studies using a variety of cell extracts that fibroin promoter activity is restricted to the silk gland of *B.mori* (16, 29). This tissue specificity correlates with the presence or absence of certain proteins in these extracts which bind to the fibroin promoter *in vitro*. It has been suggested that these factors may be homeodomain proteins analogous to the *Drosophila* EVE and ZEN proteins, and indeed EVE and ZEN are able to bind to sequences in the fibroin promoter which are homologous to their consensus binding site sequences (16). Since homeodomain proteins including EVE and ZEN are known to be absent in *Drosophila* S2 cells (7), the failure of the fibroin promoter to function detectably in these cells likely results from the absence of appropriate transcription factors. The *Drosophila* S2 cell line therefore provides a convenient null background in which to coexpress these and other factors along with a fibroin promoter-reporter gene construct to assess the role of the factors in fibroin promoter function *in vivo*. The usefulness of S2 cells as a null background for transfection assays to study transcription factors whose functions are absent from S2 cells is well established, not only for *Drosophila* factors such as homeodomain proteins (reviewed in 7), but also for factors from other species such as mammalian factor Sp1 (6).

It is interesting that all three of the constructs in our study which utilized a constitutive promoter (actin 5C distal or α1-tubulin) were present at significantly lower copy numbers than the analogous constructs with the *Mtn* promoter, which is inactive during the growth of the culture (in the absence of copper) and the fibroin promoter, which is nonfunctional in our system. This effect was reproducible in the independent cell lines regardless of the reporter gene used. One explanation for this consistent pattern is that high copy numbers of certain constitutively active promoters may be disadvantageous for the cell. A vast excess of these constitutive promoters may compete for limited transcription factors that are required for function of the endogenous actin and tubulin promoters or other constitutive promoters. Since the products of these genes are major structural proteins involved in cytoskeleton and microtubule formation (14, 30), such a competition for limited transcription factors could conceivably slow cell growth. Thus cells with fewer copies of the transfected actin or tubulin promoters may outgrow those with higher copy numbers during selection. An alternative explanation for the lower copy numbers observed with the constitutive promoters is that overproduction of tPA or gp120 on a constitutive basis may be detrimental to the cells to the extent that cells with lower copy number would outgrow the higher copy number population. However this is clearly not the explanation for the lower copy number of the tubulin promoter constructs, since cells producing five times more protein (from the actin 5C distal promoter) grow quite well. These two explanations can be distinguished experimentally by analyzing copy number in S2 cells transfected with constructs encoding the different promoters but lacking a reporter gene.

Poly(A) signals

From our comparison of the different poly(A) signals, we have been able to improve the levels of protein production in our *Drosophila* expression system two- to three-fold by replacing the SV40 early poly(A) signal with the SV40 late poly(A) signal. In addition, two aspects of our comparison reveal that the process of polyadenylation is conserved between mammalian and *Drosophila* cells. First, both types of poly(A) signals are utilized with comparable efficiencies in S2 cells. We observe no significant difference in RNA or protein production between the use of the SV40 early poly(A) signal and the *Drosophila Mtn* poly(A) signal, and the SV40 late poly (A) signal was actually utilized approximately three-fold more efficiently than the others. Second, the relative strengths of the two SV40 poly(A) signals is the same in *Drosophila* S2 cells as it is in human and monkey cells where the SV40 late poly(A) signal has also been shown to generate higher levels (five-fold) of steady state mRNA than the SV40 early poly(A) (18). Thus the factors and nucleic acid signals involved in polyadenylation appear to be conserved in *Drosophila* and mammalian cells to a large extent.

CONCLUSIONS

In summary, we find significant differences in the level of RNA and protein produced from different promoter and poly(A) sequences in stably transfected *Drosophila* S2 cells. In addition to identifying promoters (actin 5C distal and *Mtn*) and poly(A) signals (SV40 late) that are the most productive for heterologous gene expression in a high copy number stable transfection system, our results also reveal more general information about expression in S2 cells. Our results suggest that the copy number at which a population of cells maintains a particular gene can vary considerably depending on the active or inactive nature of the promoter during cell growth. In addition our data provide evidence that the nucleic acid signals and factors involved in the process of polyadenylation are conserved in insect and mammalian cells.

ACKNOWLEDGEMENTS

We thank Dr. Martin Rosenberg for helpful discussions and critical comments on the manuscript and Deborah Jiampetti for assistance with laser densitometry.

REFERENCES

1. van der Straten,A., Johansen,H., Rosenberg,M. and Sweet,R.W. (1989) *Current Methods in Molecular Biology* **1**, 1–8.
2. Ivey-Hoyle,M., Culp,J.S., Chaiken,M.A., Hellmig,B.D., Matthews,T.J., Sweet,R.W. and Rosenberg,M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 512–516.
3. Ivey-Hoyle,M. and Rosenberg,M. (1990) *Mol. Cell. Biol.* **10**, 6152–6159.
4. Johansen,H., van der Straten,A., Sweet,R., Otto,E., Maroni,G. and Rosenberg,M. (1989) *Genes and Development* **3**, 882–889.
5. Culp,J.S., Johansen,H., Hellmig,B., Beck,J., Matthews,T.J., Delers,A. and Rosenberg,M. (1991) *Bio/Technology* **9**, 173–178.
6. Courey,A.J. and Tjian,R. (1988) *Cell* **55**, 887–898.
7. Hayashi,S. and Scott,M.P. (1990) *Cell* **63**, 883–894.
8. van der Straten,A., Johansen,H., Sweet,R. and Rosenberg,M. (1987) In Mitsuhashi,J. (ed.), Invertebrate Cell Systems Applications. CRC Press, Boca Raton, Vol. 1, pp. 131–134.
9. DiNocera,P.P. and Dawid,I.B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7095–7098.
10. Chung,Y.-T. and Keller,E.B. (1990) *Mol. Cell. Biol.* **10**, 6172–6180.
11. Han,K., Levine,M.S. and Manley,J.L. (1989) *Cell* **56**, 573–583.

12. Otto,E., Allen,J.M., Young,J.E., Palmiter,R.D. and Maroni,G. (1987) *Mol. Cell. Biol.* **7**, 1710–1715.
13. Theurkauf,W.E., Baum,H., Bo,J. and Wensink,P.C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8477–8481.
14. Kalfayan,L. and Wensink,P.C. (1981) *Cell* **24**, 97–106.
15. Suzuki,Y. Takiya,S., Hara,W., Obara,T., Suzuki,T. and Hui,C.-c. (1987) In Iwai,K. (ed.), Gunma Symp. Endocrino. Center for Academic Publications Japan, Tokyo and VNU Science Press BV, Utrecht, Vol. 24, pp. 13–26.
16. Hui,C.-c., Matsuno,K. and Suzuki,Y. (1990) *J. Mol. Biol.* **213**, 651–670.
17. Levitt,N., Briggs,D., Gil,A. and Proudfoot,N.J. (1989) *Genes and Development* **3**, 1019–1025.
18. Carswell,S. and Alwine,J.C. (1989) *Mol. Cell. Biol.* **9**, 4248–4258.
19. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Pfarr,D.S., Sathe,G. and Reff,M.E. (1985) *DNA* **4**, 461–467.
21. Maroni,G., Otto,E. and Lastowski-Perry,D. (1985) *Genetics* **112**, 493–504.
22. Tsuda,M. and Suzuki,Y. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7442–7446.
23. Gimmi,E.R., Soprano,K.J., Rosenberg,M. and Reff,M.E. (1988) *Nucl. Acids Res.* **16**, 8977–8997.
24. Bond-Matthews,B. and Davidson,N. (1988) *Gene* **62**, 289–300.
25. Wigler,M., Pellicer,A., Silverstein,S., Axel,R., Urlaub,G. and Chasin,L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1373–1376.
26. Laemmli,U.K. (1970) *Nature* **227**, 680–685.
27. Gross,M., Sweet,R.W., Sathe,G., Yokoyama,S., Fasano,O., Goldfarb,M., Wigler,M. and Rosenberg,M. (1985) *Mol. Cell. Biol.* **5**, 1015–1025.
28. Shore,E.M. and Guild,G.M. (1986) *J. Mol. Biol.* **190**, 149–158.
29. Suzuki,Y., Tsuda,M., Takiya,S., Hirose,S., Suzuki,E., Kameda,M. and Ninaki,O. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9522–9526.
30. Fyrberg,E.A., Bond,B.J., Hershey,N.D., Mixter,K.S. and Davidson,N. (1981) *Cell* **24**, 107–116.