Structural Organization of the Reduced Folate Carrier Gene in Chinese Hamster Ovary Cells*

Richard C. Murray, Frederick M. R. Williams, and Wayne F. Flintoff‡

From the Department of Microbiology and Immunology, University of Western Ontario, London, Ontario N6A 5C1, Canada

We have previously described the isolation of both a hamster genomic DNA cosmid clone (29) and a cDNA clone (pMtxT9) (28) that can, upon transfection, complement mutant Chinese hamster ovary (CHO) cell lines that are unable to take up folates. Mutant cells transfected with the cDNA are able to bind and take up Mtx and thus become sensitive to the cytotoxic action of the drug.

In our continuing studies to understand the molecular nature of the reduced folate carrier gene we have localized the gene for hamster rfc to chromosomes 1 and Z1 at position q2-q3 in CHO cells (30). The present study describes the organization of the hamster rfc gene, the identification of two alternatively spliced mRNA isoforms, and the mapping of the transcriptional start site.

EXPERIMENTAL PROCEDURES

Cell Lines, Cosmid Clones, and Plasmid Clones—The wild-type, Mtx-sensitive (Pro 3) and Mtx-resistant (Pro 3 MtxRII 5–3) cell lines and the maintenance have been previously described (31–32). The isolation of the cDNA clone, pMtxT9, and its properties have been described (28).

DNA Isolation—High molecular weight DNA was isolated from exponentially growing CHO cells by the procedure of Gross-Bellard et al. (33). Cosmid and plasmid clones were propagated in Luria-Bertani medium supplemented with ampicillin or ampicillin plus tetracycline. DNA was isolated from overnight cultures using the Qiagen plasmid kit according to the conditions of the supplier. The digested DNA was separated on a 0.8% agarose gel and transferred onto Biotrans nylon paper with 2–4 × SSC, 0.1% SDS at room temperature and three times for 15 min at 50°C. The blots were exposed to x-ray film at −70°C with intensifying screens.

Probe Labeling—The CHO rfc cDNA fragment from the plasmid pMtxT9 was used as probe and labeled with α-[32P]dCTP (ICN Biomedicals; 3000 Ci/mmol) by the random priming method described by Feinberg and Vogelstein (35). Labeled DNA was routinely obtained at a specific activity of 5–10 × 106 cpm/μg.

Oligodeoxynucleotides used as probes were end-labeled with T4 polynucleotide kinase (Pharmacia) and γ-[32P]ATP (ICN Biomedicals; 7000 Ci/mmol) to a specific activity of approximately 1 × 106 cpm/μg.

Restriction Endonuclease Mapping—The initial restriction endonuclease map of cosmid 100–2 was derived from partial endonuclease digestion and hybridization with either T3 or T7 end-labeled sequencing primers as described previously (29). Fine mapping of restriction endonuclease sites was done by hybridizing Southern blots with end-labeled oligodeoxynucleotide primers that span the cDNA sequence in pMtxT9.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by a grant from the Medical Research Council of Canada.

‡ The abbreviations used are: Mtx, methotrexate; CHO, Chinese hamster ovary; RT-PCR, reverse transcriptase-polymerase chain reaction; RACE, rapid amplification of cDNA ends; bp, base pair(s); kb, kilobase(s).

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 271, No. 32, Issue of August 9, pp. 19174–19179, 1996
© 1996 by The American Society for Biochemistry and Molecular Biology, Inc.
Volume 271, Issue 32, pp. 19174–19179, 1996
Printed in U.S.A.

19174
DNA Sequencing—Double-stranded DNA sequencing was performed by the dideoxy chain termination method using the T7 sequencing kit supplied by Pharmacia. Sequences were obtained using either T7 or Sp6 sequencing primer or synthetic oligodeoxynucleotide primers spanning the CDNA sequence in PMtxT9.

Cosmid clone fragments chosen for sequence analysis were isolated and subcloned into the vectors pGEM3 or pGEM4 and propagated in Escherichia coli strains SURE®, XL1-Blue, or JM109.

RNA Isolation—Poly(A) RNA was isolated from 2 × 10^9 exponentially growing Pro 3 cells using the Fast Track mRNA isolation system (Invitrogen). Approximately 30 μg of poly(A) RNA was obtained from this number of cells.

Northern Blotting—Five μg of poly(A) RNA from Pro 3 cells were resolved on a 1.2% agarose gel containing 10 μM sodium phosphate (pH 7.0) as described previously (28). After electrophoresis, the RNA was transferred to Biotrans nylon paper (ICN Biomedicals) as described previously (28). Hybridization, washing, and autoradiography were performed as described previously for Southern blotting (29).

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)—First-strand cDNA was synthesized from 1 μg of poly(A) RNA using SuperScript II reverse transcriptase (Life Technologies, Inc.) according to the conditions recommended by the supplier. Sequences corresponding to PMtxT9 were amplified from the first-strand cDNA using a 20-mer anchor primer and a nested reverse primer designed at base pair 208 (P2: 5′-GACCGGAATGATCTTACAGT-3′) or base pair 2067 (P3: 5′-TGCTTGGCTGGCTG-GTCTG-3′) for 30 cycles of amplification using standard PCR conditions. The PCR products were separated by electrophoresis on 0.8% agarose gels or 5% acrylamide gels and visualized by ethidium bromide staining. The PCR products were subcloned using the TA cloning kit (Invitrogen) and were sequenced as described above.

Subcloning and Transfection of the Larger mRNA Isoform—The mRNA isoforms were amplified by RT-PCR from Pro 3 cells using the 20-mer forward primer P1 and a 20-mer reverse primer designed at base pair 1733 (P4: 5′-ACAGAGCTGGTGAA-GACATCCCTAAGG-3′) for 30 cycles of amplification using standard PCR conditions. The PCR products were separated by electrophoresis on 0.8% agarose gels or 5% acrylamide gels and visualized by ethidium bromide staining. The PCR products were directionally cloned into the pGEM3 or pGEM4 vectors and amplified in Escherichia coli strains SURE®, XL1-Blue, or JM109.

Subcloning of the LARGER M RNA ISOFORM—The mRNA isoforms were amplified by RT-PCR from Pro 3 cells using the 20-mer forward primer P1 and a 20-mer reverse primer designed at base pair 208 (P2: 5′-GACCGGAATGATCTTACAGT-3′) or base pair 2067 (P3: 5′-TGCTTGGCTGGCTG-GTCTG-3′) for 30 cycles of amplification using standard PCR conditions. The PCR products were separated by electrophoresis on 0.8% agarose gels or 5% acrylamide gels and visualized by ethidium bromide staining. The PCR products were subcloned using the TA cloning kit (Invitrogen) and were sequenced as described above.

Transfection of the PCR product clones was carried out by the polybrene procedure as described previously (36). Ten μg of DNA per 1–2 × 10^5 cells were transfected into the mutant Pro 3 MtxRII 5–3 cells. After transfection, cells were selected by growth in low levels of folic acid (2 nm) (37). Cells able to grow in the selective medium were isolated and tested for sensitivity to Mtx as described previously (31). The cDNA contains the sequence AGA beginning at base pair 1499, which codes for arginine in the predicted amino acid sequence. The cDNA sequence from the Pro 3 cell line, has the sequence GGA at this position, which codes for a glycine residue. This difference appears to be a polymorphism, because other cDNA clones from the CHO-K1 library contain the AGA sequence at this position, while cDNA clones and RT-PCR sequence data from the Pro 3 cell line all contain the GGA sequence. Although the significance of this change is not clear, both cDNA from Pro 3 and cDNA from CHO-K1 or Pro 3 (see below) are able to complement the Mtx transport defect in the mutant cells, which indicates that this change may not be functionally important.

The 5′-RACE product clones were sequenced by restriction endonuclease digestion and DNA sequencing to identify clones of the larger DNA isoform.

The RACE product clones were screened by restriction endonuclease digestion and DNA sequencing to identify clones of the larger DNA isoform.

The cDNA sequence in PMtxT9 was amplified from the first-strand cDNA using a 20-mer forward primer P1 and a 20-mer reverse primer designed at base pair 1733 (P4: 5′-ACAGAGCTGGTGAA-GACATCCCTAAGG-3′) for 30 cycles of amplification using standard PCR conditions. The PCR products were separated by electrophoresis on 0.8% agarose gels or 5% acrylamide gels and visualized by ethidium bromide staining. The PCR products were subcloned using the TA cloning kit (Invitrogen) and were sequenced as described above.

The 3′ RACE product was cloned directly on the cosmid DNA or on subcloned fragments to determine the intron/exon boundaries. These analyses revealed 7 exons and 6 introns that span approximately 15.3 kb (Fig. 1A). The donor and acceptor sequences (Table I) conform to consensus GT/AG splice site sequences with the exception of the donor splice site of intron 3. This site has a C replacing the consensus T at the +2 position of the intron. Although it is rare, this substitution has been observed in other donor splice sites (39).

The exon sequences were in agreement with the previously reported cDNA sequence (28) with one exception. The original cDNA was isolated from a CHO-K1 cDNA expression library. The CDNA contains the sequence AGA beginning at base pair 1499, which codes for arginine in the predicted amino acid sequence. The cDNA sequence from the Pro 3 cell line, has the sequence GGA at this position, which codes for a glycine residue. This difference appears to be a polymorphism, because other cDNA clones from the CHO-K1 library contain the AGA sequence at this position, while cDNA clones and RT-PCR sequence data from the Pro 3 cell line all contain the GGA sequence. Although the significance of this change is not clear, both cDNA from Pro 3 and cDNA from CHO-K1 or Pro 3 (see below) are able to complement the Mtx transport defect in the mutant cells, which indicates that this change may not be functionally important.

The 3′ RACE product clones were sequenced by restriction endonuclease digestion and DNA sequencing to identify clones of the larger DNA isoform.

The cDNA sequence in PMtxT9 was amplified from the first-strand cDNA using a 20-mer forward primer P1 and a 20-mer reverse primer designed at base pair 1733 (P4: 5′-ACAGAGCTGGTGAA-GACATCCCTAAGG-3′) for 30 cycles of amplification using standard PCR conditions. The PCR products were separated by electrophoresis on 0.8% agarose gels or 5% acrylamide gels and visualized by ethidium bromide staining. The PCR products were subcloned using the TA cloning kit (Invitrogen) and were sequenced as described above.

The cDNA sequence from the Pro 3 cell line, has the sequence GGA at this position, which codes for a glycine residue. This difference appears to be a polymorphism, because other cDNA clones from the CHO-K1 library contain the AGA sequence at this position, while cDNA clones and RT-PCR sequence data from the Pro 3 cell line all contain the GGA sequence. Although the significance of this change is not clear, both cDNA from Pro 3 and cDNA from CHO-K1 or Pro 3 (see below) are able to complement the Mtx transport defect in the mutant cells, which indicates that this change may not be functionally important.

The other bands in the Southern blot analysis, including the 4.3-kb BamHI doublet and the 6.4-kb HindIII band, were comparable between the cosmids 100–2 and genomic sequences.

Identification of Alternatively Spliced Messages—To deter-
mine the nature of the messenger RNA produced in Pro 3 cells, a Northern blot of poly(A⁺) selected mRNA was probed with the cDNA sequence (Fig. 3A). The Northern analysis indicated that there were two mRNA species approximately 2.5 kb and 2.6 kb in size. The two mRNAs were present in similar amounts in Pro 3 cells, since the band intensities are comparable (Fig. 3A). Amplification of the hamster rfc gene using first-strand cDNA synthesized from Pro 3 cells confirmed the presence of two mRNA species that differed in size by 121 bp (Fig. 3B). Using PCR primers specific for the 5' end of pMtxT9, the 121-bp difference was localized to this portion of the pMtxT9 sequence (Fig. 3C). The PCR products shown in Fig. 3, B and C, were cloned and sequenced. The sequence of the lower molecular weight species was identical to the sequence of pMtxT9. The sequence of the higher molecular weight species contained an additional 121 bp inserted at position 168 to 178, which was known to be a splice junction (Fig. 4). An oligodeoxynucleotide specific for the 121-bp sequence was used to map the new exon relative to the genomic DNA in the second exon of the cosmid 100–2. This exon was labeled exon 2 and the numbering of the 5'-untranslated region of the previously reported pMtxT9 sequence (28) was changed to include the new exon (Fig. 4). Therefore the hamster rfc gene codes for two mRNA species (Fig. 1B), rfc mRNA1, which contains exons 1 through 7, and rfc mRNA2, in which exon 2 is absent.

The addition of this exon occurs 5' to the putative translational start site. Thus it was anticipated that there should be no functional difference between the protein products encoded by the two messages. To confirm this, the two mRNA species were amplified by RT-PCR and cloned into the expression vector pCR3 (Invitrogen). The resulting RT-PCR clones were sequenced, and three clones of the large mRNA isoform, which contain the 121-bp second exon, were transfected into the mutant Pro 3 MtxRII 5–3 cell line and selected for the ability to complement the mutant phenotype. All three clones were able to rescue the phenotype at a frequency comparable to that of the cDNA clone pMtxT9, which corresponds to the mRNA2 sequence. This indicates that there is no apparent functional difference between the mRNA isoforms with respect to their ability to complement the mutant phenotype.

To ensure that DNA fidelity was maintained during the amplification process, the sequence of the RT-PCR clones was...
compared to the cDNA sequence. Three sequence alterations that were presumably due to polymerase error were noted in the three RT-PCR clones. One of the three clones had no sequence alterations, while another clone had a silent mutation at base pair 780. The third clone had two point mutations, a silent mutation at base pair 1648 and a mutation at base pair 853 that resulted in the substitution of methionine for leucine at amino acid 285 in the predicted amino acid sequence. This change is relatively conservative and did not appear to affect the ability of the clone to rescue the mutant phenotype.

The 5'-region of the Hamster Mtx Transport Gene—The transcriptional start site of the hamster rfc gene was mapped using 5'-RACE and RNase protection analysis. The results of the 5'-RACE analysis are indicated in Fig. 4. Thirty-three clones were obtained and sequenced. The majority of the clones mapped to six sites near the beginning of the pMtxT9 sequence with two clones mapping to a site at base pair 53 of the putative open reading frame.

RNase protection was used to confirm the location of the transcription start sites. Six RNase protection products were detected (Fig. 5) whose positions corresponded well with the 5'-RACE products near the beginning of the pMtxT9 sequence (Fig. 4). The predominant RNase protection product appears to correspond to the first base pair of the pMtxT9 clone (Fig. 4). Two minor bands were seen in the RNase protection approximately 50 bp upstream of the major RNase protection product. These bands were not very intense but may indicate that there is a low level usage of an upstream transcription start site. No 5'-RACE products corresponding to an mRNA of this size were detected.

Two hundred sixty-one base pairs of sequence upstream of the major transcriptional start site from the cosmid 100–2 are also shown in Fig. 4. This sequence is consistent with a “house-keeping” gene promoter sequence in that it is fairly GC-rich, does not contain a TATA box element, and contains a consensus binding site for the Sp1 transcription factor at −346 (40).

**DISCUSSION**

This report describes the organization of the hamster rfc gene, the characterization of two alternatively spliced transcripts, and the identification of the transcriptional start sites. The gene is organized into seven exons and six introns. The sequences at the splice site junctions conform to the consensus GT/AG with the exception of the donor splice site of intron 3, which contains a C instead of a T at the +1 position of the intron. Although this splice site does not conform to the consensus, it appears to function efficiently in CHO cells. We have not been able to detect a splicing error at this intron/exon junction in mRNA from Pro 3 cells using RT-PCR analysis. Splicing errors at other splice site junctions do seem to occur.
in the hamster rfc gene. In screening the mRNA from Pro 3 cells by RT-PCR, we were able to detect a potential splice error at the splice acceptor site of exon 6. Two RT-PCR clones have been isolated that contain either a 7- or 29-bp deletion at the beginning of exon 6. The two deletions both begin at the first base pair of exon 6 and end after the next two downstream AG sequences, indicating that they may be cryptic splice acceptor sites. The deletions in these RT-PCR products correspond to the 29-bp deletion previously reported in Williams et al. (28) for clones pMtxT5 and pMtxT7. Since the 29-bp deletion is found in both a cDNA library and by RT-PCR analysis, there must be a proportion of mRNAs in the cell that contain the deletion. Although our PCR conditions were not quantitative, the majority of PCR products isolated did not contain the 29- or the 7-bp deletion.

Another possible splicing error that involves the splice site junctions that flank exon 4 has been reported in the literature. The hamster rfc cDNA clone pMtxT5 (28) contains a 757-bp Sac I fragment that contains the 53-bp first exon, 109 bp of 5'-flanking sequences, and approximately 240 bp of intron 1 was used to generate a riboprobe for RNase protection. The probe was incubated with 3 μg of yeast RNA (lane 1) and 2.5 μg of Pro 3 mRNA (lane 2), digested with RNase, separated on a 20% acrylamide gel, and autoradiographed as described under “Experimental Procedures.” The arrow indicates the major RNase protection product and the arrowheads indicate the other products. The two bands near the top of the figure are of lower intensity than the other bands and appear consistently in other experiments. ●, the first base pair of the cDNA clone pMtxT9.

The original cDNA clone pMtxT9, indicating that there is no significant functional difference between them in this assay system. Although this is the first report of an alternative splice in the rfc gene, we also see a doublet band on a Northern blot of mouse RNA. It will be interesting to see if the two mRNA isoforms show any differences in spatial/temporal expression or translational regulation in an intact organism.

We have mapped the transcriptional start sites of the rfc gene to several positions approximately 200 bp upstream of the translational start site. The RACE data and RNase protection data taken together indicate that there are approximately six transcription start sites close to the beginning of the pMtxT9 cDNA clone. The size of transcripts generated from these start sites agrees with the observed size of the rfc mRNA on a Northern blot. The use of imprecise transcriptional start sites is common in TATA-less “housekeeping” genes (40).

Two additional bands in the RNase protection assay indicate that there may be a low level usage of an upstream transcriptional start site. We have not been able to detect any 5'-RACE products that would correspond to these bands, nor do we see a larger mRNA on a Northern blot. The possible use of an upstream transcriptional start site is currently being investigated using RT-PCR.

Brigle et al. (41) have shown an alignment of the translated nucleotide sequence upstream of the predicted ATG initiation codon from the CHO and L1210 rfc cDNAs. The authors point out that the degree of similarity and lack of an upstream
in-frame termination codon may indicate that these sequences encode additional protein information. However, an analysis of the putative 5'-untranslated region of the human cDNA sequence, which does not contain an in-frame stop codon, revealed very little homology with the hamster or mouse sequences. This is in contrast with the amino acid sequence downstream of the putative ATG initiation codon, which is highly conserved among hamster, mouse, and human (23). These observations combined with the transcriptional start site mapping data presented in this paper and the size of the messenger RNA on a Northern blot imply that the predicted ATG initiation codon is the major translational start site.

The information obtained in these studies should prove useful for determining the state of the rfc gene in both normal and transport-deficient cell systems.

Acknowledgments—We thank T. M. Underhill and P. L. Ferguson for helpful discussions.

REFERENCES

1. Antony, A. C., Kane, M. A., Portillo, R. M., Elwood, P. C., and Kolhouse, J. F. (1985) J. Biol. Chem. 260, 14911–14917
2. Elwood, P. C., Kane, M. A., Portillo, R. M., and Kolhouse, J. F. (1986) J. Biol. Chem. 261, 15416–15423
3. Kane, M. A., Elwood, P. C., Portillo, R. M., Antony, A. C., and Kolhouse, J. F. (1986) J. Biol. Chem. 261, 15625–15631
4. Kane, M. A., Portillo, R. M., Elwood, P. C., Antony, A. C., and Kolhouse, J. F. (1986) J. Biol. Chem. 261, 44–49
5. Henderson, G. B., Tsuiji, J. M., and Kumar, H. P. (1988) Arch. Biochem. Biophys. 274, 327–337
6. Deutsch, J. C., Elwood, P. C., Portillo, R. M., Macey, M. G., and Kolhouse J. F. (1989) Arch. Biochem. Biophys. 274, 327–337
7. Jensen G., Kathmann, I., Rödel, M., and Kolhouse, J. F. (1989) Biochim. Biophys. Acta 914, 432–439
8. Kane, M. A., and Waxman, S. (1989) J. Biol. Chem. 264, 14862–14865
9. Fan, J., Vitols, K. S., and Huennekens, F. M. (1991) J. Biol. Chem. 266, 14911–14917
10. Matherly, L. H., Czajkowski, C. A., and Angeles, S. M. (1991) Cancer Res. 51, 3420–3426
11. Matherly, L. H., Angeles, S. M., and Czajkowski, C. A. (1992) J. Biol. Chem. 267, 3253–3260
12. Matherly, L. H., and Angeles, S. M. (1994) Biochem. Pharmacol. 47, 1094–1098
13. Williams, F. M. R., and Flintoff, W. F. (1995) J. Biol. Chem. 270, 2974–2979
14. Coney, L. R., Tomassetti, A., Carayannopoulos, L., Frasca, V., Kamen, B. A., Colnaghi, M. I., and Zurawski, V. R. (1991) Cancer Res. 51, 6125–6132
15. Rappoport, J., Senger, G., Trowsdale, J., and Campbell, I. G. (1992) Genomics 14, 423–430
16. Sadasivan, E., Cedeno, M., and Rothenberg, S. P. (1992) Biochim. Biophys. Acta 1131, 91–94
17. Sadasivan, E., Cedeno, M. M., and Rothenberg, S. P. (1994) J. Biol. Chem. 269, 4725–4735
18. Price, E., Ratnam, M., Rodeman, K. M., and Freiheit, J. H. (1988) Biochemistry 27, 7853–7858
19. Farn, J., Vitols, K. S., and Huennekens, F. M. (1991) J. Biol. Chem. 266, 14862–14865
20. Matherly, L. H., Angeles, S. M., and Czajkowski, C. A. (1992) J. Biol. Chem. 267, 3253–3260
21. Matherly, L. H., Angeles, S. M. (1994) Biochem. Pharmacol. 47, 1094–1098
22. Williams, F. M. R., and Flintoff, W. F. (1995) J. Biol. Chem. 270, 2974–2979
23. Moskow, J. A., Gong, M., He, R., Saglas, M. K., Dixon, K. H., Anzick, S. L., Metzler, P. S., and Cowan, K. H. (1995) Cancer Res. 55, 3790–3794
24. Prasad, P. D., Ramamoorthy, S., Leibach, F. H., and Ganapathy, V. (1995) Biochem. Biophys. Res. Commun. 206, 681–687
25. Wong, S. C., Proefke, S. A., Bhushan, A., and Matherly, L. H. (1995) J. Biol. Chem. 270, 17468–17475
26. Dixon, K. H., Lapirht, B. C., Chiu, J., Kelley, K., and Cowan, K. H. (1994) J. Biol. Chem. 269, 17–20
27. Williams, F. M. R., Murray, R. C., Underhill, T. M., and Flintoff, W. F. (1994) J. Biol. Chem. 269, 5810–5816
28. Underhill, T. M., Williams, F. M. R., Murray, R. C., and Flintoff, W. F. (1992) Somat Cell Mol. Genet. 18, 337–349
29. Chan, F. P. H., Williams, F. M. R., Rogers, K. A., and Flintoff, W. F. (1995) Cytogenet. Cell Genet. 71, 148–150
30. Flintoff, W. F., Davidson, S. V., and Siminovitch, L. (1976) Somat Cell Genet. 2, 245–261
31. Flintoff, W. F., Spindler, S. M., and Siminovitch, L. (1976) In Vitro 12, 749–757
32. Gross-Bellard, M., Oudek, P., and Chambon, P. (1978) Eur. J. Biochem. 83, 32–38
33. Souleman, E. M. (1975) J. Biol. Chem. 98, 503–517
34. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13
35. Underhill, T. M., and Flintoff, W. F. (1989) Mol. Cell. Biol. 9, 1754–1758
36. Underhill, T. M., and Flintoff, W. F. (1989) Somat Cell Mol. Genet. 15, 49–59
37. Leibbrandt, M. E. I., Khokha, R., and Koropatnick, J. (1994) Cell Growth & Differ. 5, 17–25
38. Shapiro, M. B., and Senapathy, P. (1987) Nucleic Acids Res. 15, 7155–7174
39. Azizkhan, J. C., Jensen, D. E., Pierce, A. J., and Wade, M. (1993) Crit. Rev. Eukaryotic Gene Expression 3, 229–254
40. Azizkhan, J. C., Jensen, D. E., and Chambon, P. (1993) Crit. Rev. Eukaryotic Gene Expression 3, 229–254