

# Molecular cloning of the human CTP synthetase gene by functional complementation with purified human metaphase chromosomes

Masatake Yamauchi, Noriko Yamauchi and Mark Meuth

Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Herts EN6 3LD, UK

Communicated by T.Lindahl

**Successive rounds of chromosome-mediated gene transfer were used to complement a hamster cytidine auxotroph deficient in CTP synthetase activity and eventually to clone human genomic and cDNA fragments coding for the structural gene. Our approach was to isolate human *Alu*<sup>+</sup> fragments from a tertiary transfectant and to utilize these fragments to screen a panel of primary transfectants. In this manner two DNA fragments, both mapping within the structural gene, were identified and used to clone a partial length cDNA. The remaining portion of the open reading frame was obtained through the RACE polymerase chain reaction technique. The open reading frame encodes 591 amino acids having a striking degree of similarity to the *Escherichia coli* structural gene (48% identical amino acids with 76% overall similarity including conservative substitutions) with the glutamine amide transfer domain being particularly conserved. As regulatory mutations of CTP synthetase confer both multi-drug resistance to agents widely used in cancer chemotherapy and a mutator phenotype, the cloning of the structural gene will be important in assessing the relevance of such phenotypes to the development of cellular drug resistance.**

**Key words:** CTP synthetase/chromosome transfer/drug resistance/mutator/nucleotide synthesis

## Introduction

The stringent regulation of intracellular levels of cytidine nucleotides in mammalian cells is essential for maintaining the accuracy of DNA replication and governing sensitivity to several cytotoxic drugs (Meuth, 1989). These nucleotides are synthesized in both prokaryotic and eukaryotic cells by the amination of UTP in the reaction  $UTP + ATP + glutamine \rightarrow CTP + ADP + P_i + glutamate$ . The enzyme catalysing this step, CTP synthetase, is also a key regulatory point as its activity is subject to activation by GTP and inhibition by CTP (McPartland and Weinfeld, 1979) and mutant cell strains altered in this regulation (insensitive to inhibition by CTP) display a complex phenotype: (i) increased intracellular pools of CTP and dCTP (Meuth *et al.*, 1979, 1985; de Saint Vincent and Buttin, 1980), (ii) multi-drug resistance—arabinosyl cytosine, 5-fluorouracil, thymidine, bromodeoxyuridine and DNA alkylating agents (Meuth, 1981; Meuth *et al.*, 1982; Aronow *et al.*, 1984; Kaufman, 1986), and (iii) an increased rate of spontaneous

mutation (Meuth *et al.*, 1979; Aronow *et al.*, 1984; Chu *et al.*, 1984).

Since many nucleoside and base analogues are widely used in the chemotherapy of leukaemias and other tumours, the development of such genetically codominant mutations in target cell populations could have important consequences for the outcome of the treatment. We attempted to define the nature of the regulatory mutations at the DNA sequence level through the cloning of wild type and mutant CTP synthetase genes. To accomplish this, we first isolated hamster cell strains deficient in CTP synthetase activity that are dependent upon exogenous cytidine for survival (Kelsall and Meuth, 1988). We then attempted to complement this auxotrophy by DNA mediated gene transfer. When the DNA transfers met with little success, we introduced purified metaphase chromosomes isolated from CTP synthetase proficient HeLa cells into our deficient strains since previous observations showed that chromosome mediated gene transfer (CMGT) was more effective in complementing purine auxotrophs (Yamauchi *et al.*, 1989). This strategy has proved considerably more successful as, after several rounds of chromosome transfer and selection, genomic fragments bearing the structural gene were cloned. Here we describe the isolation and characterization of genomic and cDNA sequences encoding the human CTP synthetase.

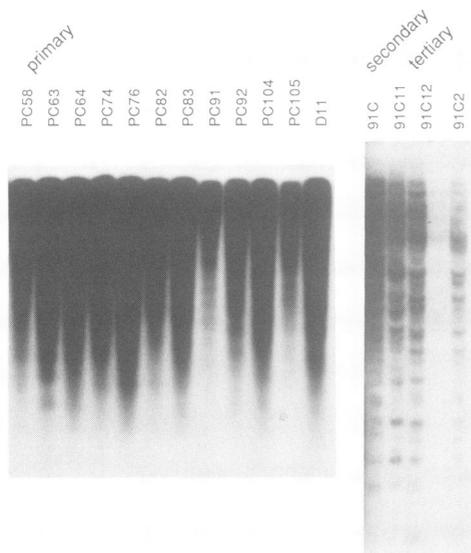
## Results

### Complementation of CTP synthetase deficient mutants by CMGT

The Chinese hamster (CHO) CTP synthetase deficient cytidine auxotroph CR<sup>-2</sup> was treated with DNA or mitotic chromosomes co-precipitated with calcium phosphate. After recovery and expression, the cells were plated in selective medium (lacking cytidine) to allow formation of colonies acquiring the human CTP synthetase gene. The frequency of such colonies in the first round of transfer was  $\sim 10^{-6}$ . When purified DNA was used instead of isolated chromosomes, the frequency of colony formation was about  $10^{-7}$ . In control cultures (no added DNA or chromosomes) no colonies were observed (frequency  $< 10^{-7}$ ).

To determine if the colonies survived as a result of the uptake and expression of the human CTP synthetase gene, DNAs purified from several of the isolates were digested with the restriction endonuclease *EcoRI*. Digests were fractionated by electrophoresis on agarose gels and transferred to nitrocellulose filters. These filters were then hybridized with labelled human *Alu* repeat BLUR 8. Figure 1 shows that DNA purified from the colonies recovered after CMGT contained a smear of fragments hybridizing with the BLUR 8 probe. Colonies derived from DNA transfers, on the other hand, showed no human *Alu* containing DNA (data not shown).

To reduce the number of human DNA fragments,

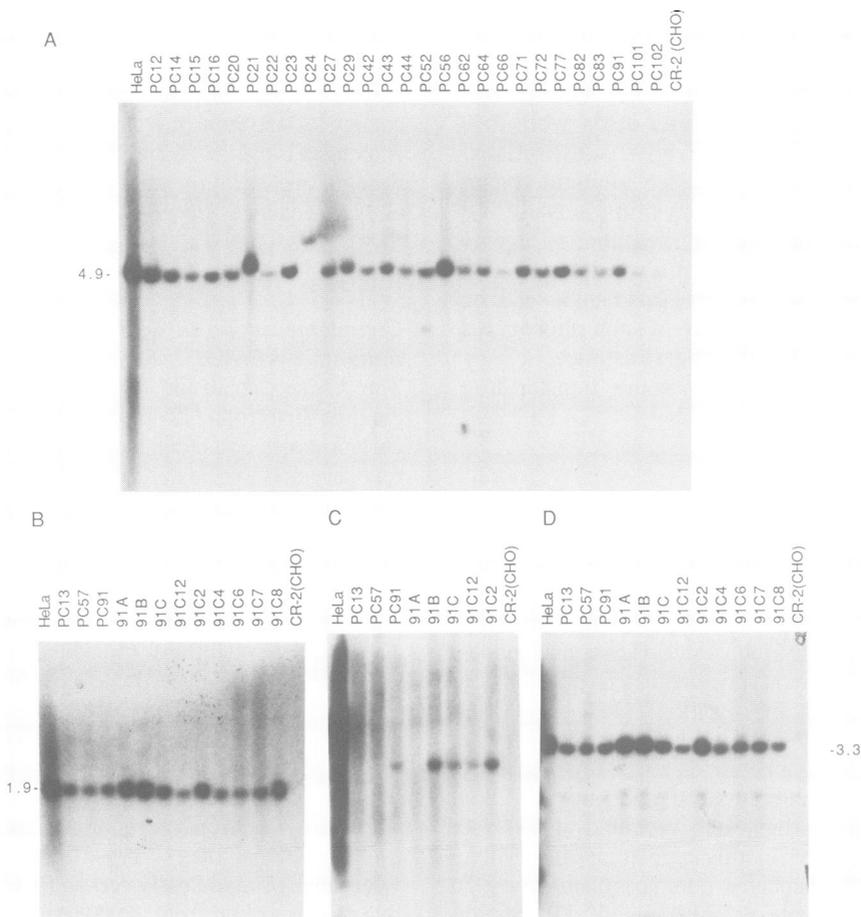


**Fig. 1.** Southern blot analysis of DNA purified from primary, secondary and tertiary transfectants of the cytidine auxotroph CR<sup>-2</sup>. *EcoRI* digests were blotted and probed with labelled BLUR8 human *Alu* sequence. Primary, secondary and tertiary transfectants are indicated.

chromosomes isolated from one of the transfectants (apparently containing the least number of *Alu*<sup>+</sup> fragments, PC91) were used to transfect CR<sup>-2</sup> for a second round of transfection and chromosomes purified from a secondary transfectant (PC91C) were subsequently used in a third round. The frequencies of colonies growing in the absence of cytidine in the latter two rounds were somewhat lower than the first ( $1-2 \times 10^{-7}$ ), but DNA purified from these survivors clearly contained human *Alu* sequences (Figure 1). It also appeared that the colonies obtained from these further rounds of CMGT had progressively fewer human *Alu* containing fragments.

**Isolation of human *Alu* containing fragments linked to CTP synthetase**

As common bands were evident among the tertiary transformants, DNA obtained from one of the isolates (91C12) was digested with *EcoRI*, size fractionated and ligated to *EcoRI* cleaved DNA from the  $\lambda$  vector NM1149. The recombinant phage libraries derived were screened for human *Alu* bearing fragments: 120 independent isolates were eventually picked and cloned inserts from these recombinants were used to screen a panel of primary, secondary and tertiary transfectants. Presumably any fragments retained by all the strains would be closely linked with the CTP



**Fig. 2.** Screens of human *Alu*<sup>+</sup> fragments isolated from a tertiary transfectant (91C12) against a panel of DNAs purified from primary, secondary and tertiary transfectants. (A) A 4.9 kb human fragment present in DNA digests of all primary transfectant and human (HeLa) DNAs but not CHO recipient cell (CR<sup>-2</sup>) DNA. (B) Southern blot showing the presence of a 1.9 kb human fragment in primary, secondary and tertiary transfectants. (C) An example of one of the many human *Alu*<sup>+</sup> fragments, cloned from the tertiary transfectant, not represented in all primary transfectants. (D) Hybridization of the conserved 3.3 kb fragment (obtained by its linkage to the two common fragments in a genomic  $\lambda$  clone) to primary, secondary and tertiary transfectants. Note the faint hamster specific fragments detected.

synthetase structural gene. Of the 120 fragments screened in this manner only two were common to the entire panel (Figure 2A–C).

To determine the proximity of these two fragments, one of them (1.9 kb in size) was used to probe genomic libraries prepared from partially digested human DNA ligated into the  $\lambda$ DASH vector. One isolate having a 16 kb human DNA insert was obtained in the screen. To our surprise, the insert also contained the second (4.9 kb) 'linked' fragment only ~6 kb away from the first (Figure 3). Furthermore, a unique 3.3 kb fragment subcloned from the 16 kb insert gave faint hamster specific signals when it was used to probe blots containing digests of human, hamster and transfectant DNA, indicating that these fragments carried sequences conserved in both hamsters and humans (Figure 2D).

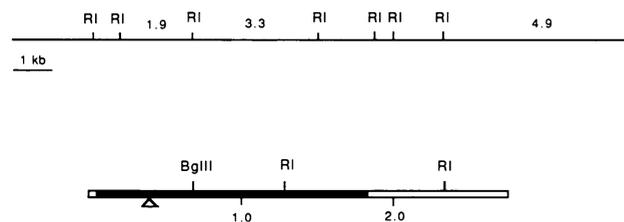
Given the conserved sequences present on this fragment, RNA prepared from human cells was fractionated and probed on Northern blots with the labelled 3.3 kb fragment. Figure 4 shows a clear band of ~3 kb hybridizing with this probe. In contrast RNA prepared from the CTP synthetase deficient hamster strain showed no band.

#### Isolation of human cDNAs homologous to the conserved sequences

The conserved 3.3 kb fragment was used to screen human cDNA libraries. Of several screened only one (prepared from human testis DNA) yielded positives. DNA purified from all the positive clones had common 1.1 and 0.5 kb *Eco*RI fragments, but one (clone 14) had a larger, 2.3 kb, insert (Figure 3). All of the clones hybridized with the original 4.9 kb fragment as well.

#### Sequence analysis of the cDNA clone and isolation of the complete coding sequence

To determine whether the 2.3 kb cDNA encoded the human CTP synthetase, the fragment was sequenced and the open reading frames found were used to search protein databases. This analysis revealed a long open reading frame which had a strong homology with only one protein, that of the *Escherichia coli* CTP synthetase (Weng *et al.*, 1986). The match was extensive, >40% identical amino acids and >70% similarity when conservative substitutions were considered, and it included the entire open reading frame.



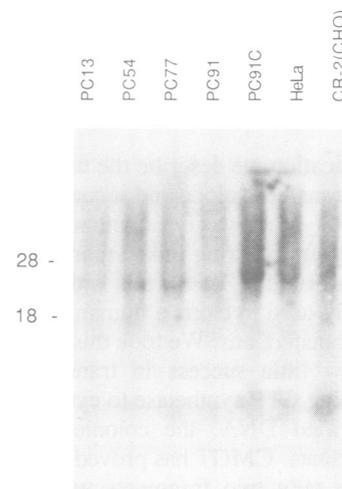
**Fig. 3.** Maps of cloned genomic and cDNA fragments. **Top** figure shows the 16 kb  $\lambda$ DASH insert isolated by probing human libraries with the 1.9 kb fragment common to all transfectants (Figure 2B). This fragment also contained the common 4.9 kb fragment (Figure 2A). The conserved 3.3 kb fragment was used to probe Northern blots and cDNA libraries. **Bottom** figure shows the map of the cDNA. The portion to the right of the triangle represents that cloned from the human testis cDNA library using the conserved 3.3 kb fragment as probe, the remainder was obtained by RACE PCR. The filled portion represents the 591 amino acid open reading frame. Although the alignments between the genomic and cDNA sequences have not been precisely determined, the 3.3 kb fragment hybridizes to the *Bgl*III–*Eco*RI cDNA fragment.

However, it only covered about two thirds of the *E. coli* gene (the carboxyl portion up to amino acid 146) indicating that the amino end was still to be cloned.

To obtain upstream portions of the CTP synthetase message, the RACE polymerase chain reaction (RACE PCR) technique (Frohman *et al.*, 1988) was employed. An oligonucleotide complementary to the sequence just downstream of the *Bgl*III site of our cDNA (Figure 3) was used as primer for reverse transcription of the CTP synthetase message in RNA prepared from HeLa cells. Transcripts were then A tailed and a second (nested) oligonucleotide complementary to the cDNA together with a tail specific oligonucleotide (also bearing a *Sal*I site for subsequent cloning) were used to amplify the products generated from reverse transcription. A portion of the products of the amplification were fractionated by electrophoresis on agarose gels and blotted onto nitrocellulose for probing with an appropriate cDNA fragment (upstream of the *Bgl*III site). This revealed that the extension was successful and a further 527 bp upstream was sequenced from *Bgl*III–*Sal*I digested products subcloned into M13. From this determination a potential ATG start codon (consistent with the criteria of Kozak, 1986) was found in a position which would correspond to the start of the *E. coli* enzyme and an open reading frame of 1773 nucleotides or 591 amino acids was evident (Figure 5). A polyadenylation signal was present 905 bp downstream from the stop codon but there was no poly(A)tail. This sequence has been deposited into the EMBL database, accession number X52142.

#### Sequence features of the CTP synthetase cDNA

The 591 amino acid open reading frame would represent a protein of 66 700 kd, virtually identical to the mol. wt of 66 000 determined by protein purification protocols (Kizaki *et al.*, 1985; Thomas *et al.*, 1988). Alignment of this sequence with that of the *E. coli* CTP synthetase reveals a striking similarity over the entire length with the human enzyme being only 46 amino acids larger, mainly at the carboxyl end. 248 amino acids (48%) were identical with another 139 representing conservative substitutions (for an



**Fig. 4.** Northern blot analysis of RNA obtained from primary transfectants, HeLa and CR<sup>-</sup>2 hamster cells using the 3.3 kb fragment as probe. A ~3 kb RNA species is detected in human and transfectant cells. Positions of 18 and 28S ribosomal subunits are indicated.



phenotype and provide a means of rapidly assaying its development.

## Materials and methods

### Cells

The cytidine auxotrophic strain CR<sup>-2</sup>, originally isolated from the Toronto Pro<sup>-</sup> CHO cell line (Kelsall and Meuth, 1988), was grown in  $\alpha$  minimal essential medium ( $\alpha$ -MEM, GIBCO) supplemented with 7.5% fetal calf serum, 10  $\mu$ M thymidine, and 20  $\mu$ M cytidine. When CTP synthetase proficient transfectants were selected, 5% dialysed fetal calf serum was used for the cytidine deficient medium.

### Chromosome and DNA mediated gene transfer

Metaphase chromosomes were prepared essentially as described by Lewis *et al.* (1980). Colcemid was added to a 1 l exponentially growing culture of HeLa cells to a final concentration of 50 ng/ml. After 18 h cells were harvested and resuspended in 100 ml of ice cold 75 mM KCl for 30 min. After recentrifugation, the cells were then resuspended in 2 ml ice cold 15 mM HEPES(Na) pH 7.0, 3 mM CaCl<sub>2</sub>, and 1% NP-40 for disruption by a Dounce homogenizer (10 strokes, on ice). The homogenate was centrifuged at 200 g to remove debris and the resulting supernatant was then centrifuged at 1500 g for 20 min at 4°C. The pellet obtained was washed once with ice cold homogenization buffer without detergent. This chromosome preparation was then resuspended in 10 ml of gene transfer buffer (25 mM HEPES(Na) pH 7.1, 140 mM NaCl, and 0.75 mM Na<sub>2</sub>HPO<sub>4</sub>) and 0.5 ml of 2.5 M CaCl<sub>2</sub> was added. The resultant co-precipitate was immediately added to recipient (CR<sup>-2</sup>) cells (2–3  $\times$  10<sup>6</sup>/10 cm dish) in growth medium supplemented with antibiotics (0.05 mg/ml gentamycin and 1  $\mu$ g/ml amphotericin B, Sigma), 15 mM HEPES(Na) pH 7.1, and 0.1% polyethylene glycol (MW1500).

For DNA mediated gene transfer, DNA was resuspended in the same gene transfer buffer and added to cells as described above.

Recipient cells were exposed to chromosome or DNA–calcium phosphate co-precipitate for 8–16 h, treated with 10% DMSO in the last 30 min, and then allowed to recover in nonselective growth medium for 24 h. Treated cells were trypsinized, washed in PBS twice and plated at a density of 2–5  $\times$  10<sup>5</sup>/10 cm dish in the absence of cytidine. Colonies appearing after 2–4 weeks were picked for further analysis.

### Genomic and cDNA cloning

*Eco*RI digested DNA purified from tertiary transfectants was fractionated by electrophoresis on agarose gels followed by electroelution of DNA of the desired size as described previously (Nalbantoglu *et al.*, 1986). This size fractionated DNA was then ligated with *Eco*RI cut DNA of the  $\lambda$  insertion vector NM1149. These reactions were then packaged and plated on the strain NM514 *recA hfl* giving efficiencies of  $\sim 1 \times 10^6$  recombinants/ $\mu$ g vector. These libraries were screened with the human *Alu* repeat BLUR8 (Amersham) and the positives were picked for further screening against our panel of transfectants.

*Sau*3A partial digests of HeLa cell DNA were ligated with *Bam*HI digested DNA purified from the  $\lambda$ Dash vector (Stratagene). Libraries of recombinant phage were obtained by plating the packaged DNA on *E. coli* strain P2392 and screened using a 1.9 kb fragment present in all our transfectants (isolated in the screens described above).

A human  $\lambda$ gt11 cDNA library, prepared with RNA isolated from human testis (Clontech), was screened with the conserved 3.3 kb unique human DNA fragment common to all our transfectants. Positives obtained were then used for DNA sequence analysis.

### RACE PCR

Total RNA prepared from HeLa cells for RACE PCR (rapid amplification of cDNA end polymerase chain reaction; Frohman *et al.*, 1988) was reverse transcribed using an oligonucleotide primer (nucleotides 784–>765 or 5'-CAACATGGCAGAACATTGAT-3') and super RT (Anglian Biotec Ltd, UK). The product was A tailed by nucleotidyl transferase (BRL) and the opposite DNA strand was synthesized using a (dT)17 adaptor and the Taq polymerase (Perkin Elmer Cetus). The resultant double stranded DNA was then amplified using the adaptor and a nested primer (nucleotides 758–>739 or 5'-TCCTTCACTGATGTGTCAAG-3') by PCR. The product of this reaction was digested with *Sal*I and *Bgl*II and cloned into the M13 vectors mp18 and 19. The clone with the longest insert (700 bp) was sequenced.

### DNA sequencing

Fragments purified from the  $\lambda$ gt11 isolates were subcloned into the M13 vectors mp18 or 19. Sequencing was performed by the dideoxy chain

termination method using [ $\alpha$ -<sup>35</sup>S]dATP. Polymerization reactions using Sequenase (USB) were primed by either the 17mer universal primer or appropriate internal primers. Sequencing reactions were electrophoresed on 5% polyacrylamide denaturing gels.

### DNA sequence analysis

Sequences were analysed using Intelligenetics software. Protein databases were screened using the program Prosrch (Coulson *et al.*, 1987) on an AMT DAP (Distributed Array Processor) 610.

## Acknowledgements

We are grateful to Iain Goldsmith for the supply of oligonucleotides and Michelle Ginsburg for her help with computer analysis of DNA and amino acid sequences. Richard Iggo is thanked for helpful suggestions regarding RACE PCR. Rick Wood and Peter Goodfellow are thanked for useful comments concerning the manuscript.

## References

- Aronow, B., Watts, T., Lassetter, J., Washteyn, W. and Ullman, B. (1984) *J. Biol. Chem.*, **259**, 9035–9043.
- Caras, I., Levinson, B.B., Fabry, M., Williams, S.R. and Martin, D.W., Jr. (1985) *J. Biol. Chem.*, **260**, 7015–7022.
- Chu, E.H.Y., McLaren, J.D., Li-I.-C. and Lamb, B. (1984) *Biochem. Genet.*, **22**, 701–715.
- Coulson, A.F.W., Collins, J.F. and Lyall, A. (1987) *Computer J.*, **30**, 420–424.
- de Saint Vincent, B.R. and Buttin, G. (1980) *Biochim. Biophys. Acta*, **610**, 352–359.
- Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8998–9002.
- Kaufman, E.R. (1986) *Mutat. Res.*, **161**, 19–27.
- Kelsall, A. and Meuth, M. (1988) *Somat. Cell Mol. Genet.*, **14**, 149–154.
- Kizaki, H., Ohsaka, F. and Sakurada, T. (1985) *Biochim. Biophys. Acta*, **829**, 34–43.
- Kozak, M. (1986) *Cell*, **44**, 283–292.
- Lewis, W.H., Srinivasan, P.R., Stokoe, N. and Siminovitch, L. (1980) *Somat. Cell Genet.*, **6**, 333–347.
- McPartland, R.P. and Weinfeld, H. (1979) *J. Biol. Chem.*, **254**, 11394–11398.
- Meuth, M. (1981) *Somat. Cell Genet.*, **7**, 89–102.
- Meuth, M. (1989) *Exp. Cell Res.*, **181**, 305–316.
- Meuth, M., Goncalves, O. and Thom, P. (1982) *Somat. Cell Genet.*, **8**, 423–432.
- Meuth, M., Goncalves, O. and Trudel, M. (1985) In de Serres, F.J. (ed.), *Genetic Consequences of Nucleotide Pool Imbalance*. Plenum Press, New York, pp. 297–312.
- Meuth, M., L'Heureux, H., Huard, N. and Trudel, M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 6505–6509.
- Meuth, M., Trudel, M. and Siminovitch, L. (1979) *Somat. Cell Genet.*, **5**, 303–318.
- Miller, C.L. and Ruddle, F.H. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 3346–3350.
- Nalbantoglu, J., Hartley, D., Phear, G., Tear, G. and Meuth, M. (1986) *EMBO J.*, **5**, 1199–1204.
- Olsen, L.C., Aasland, R., Wittwer, C.U., Krokan, H.E. and Helland, D.E. (1989) *EMBO J.*, **8**, 3121–3125.
- Porteous, D.J., Morten, J.E.N., Cranston, G., Fletcher, J.M., Mitchell, A., van Heyningen, V., Fantes, J.A., Boyd, P.A. and Hastie, N.D. (1986) *Mol. Cell. Biol.*, **6**, 2223–2232.
- Pritchard, C.A. and Goodfellow, P.N. (1987) *Genes Dev.*, **1**, 172–178.
- Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T. (1985) *Nucleic Acids Res.*, **13**, 2035–2043.
- Thelander, L. and Berg, P. (1986) *Mol. Cell. Biol.*, **6**, 3433–3442.
- Thomas, P.E., Lamb, B.J. and Chu, E.H.Y. (1988) *Biochim. Biophys. Acta*, **953**, 334–344.
- Weng, M., Makaroff, C.A. and Zalkin, H. (1986) *J. Biol. Chem.*, **261**, 5568–5574.
- Yamauchi, M., Ayusawa, D., Shimizu, K., Seno, T. and Matsuhashi, M. (1989) *Somat. Cell. Mol. Genet.*, **15**, 39–48.

Received on February 9, 1990; revised on April 12, 1990