Vol. 117, No. 2, 1983 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
December 16, 1983 Pages 601-609

NUCLEOTIDE SEQUENCE OF A cDNA CLONE FOR HUMAN ALDOLASE B

Claude BESMOND[1], Jean-Claude DREYFUS[1], Claudine GREGORI[1], Monique FRAIN[2], Mario M. ZAKIN[3], José SALA TREPAT[2] and Axel KAHN[1]

1. Institut de Pathologie Moléculaire, INSERM U 129, CHU COCHIN, 24 rue du Faubourg St Jacques, 75674 PARIS CEDEX 14 - FRANCE
2. Laboratoire d'Enzymologie du CNRS, 91110 GIF SUR YVETTE - FRANCE
3. Unité de Biochimie Cellulaire, Institut Pasteur, 28 rue du Dr Roux, 75724 PARIS CEDEX 15 - FRANCE

Received August 29, 1983

SUMMARY. Two specific clones for human aldolase B were isolated from a human liver cDNA library using a rat aldolase B cDNA probe. The clones were identified by positive hybridization-selection and one of them was sequenced. The 127 C-terminal residues of the human protein were deduced from this nucleotide sequence analysis. They showed 92 % homology with the corresponding previously published amino-acid sequence of rat liver aldolase B.

The enzyme aldolase (EC 4.1.2.13) exists as three genetically distinct isozymes (1) : muscle-type (aldolase A), liver type (aldolase B) and brain-type (aldolase C). Aldolase B is the only isozyme able to metabolize fructose ; its deficiency is responsible for an inborn error of metabolism, Hereditary Fructose Intolerance (2-4).

In young fetuses, the liver mainly synthesizes both aldolases A and C, development being associated with a progressive disappearance of these isozymes correlated with a parallel increase of aldolase B (5).

Hepatocarcinogenesis is associated with the repression of aldolase B synthesis and with the reappearance of the "fetal isozymes" (5-7). It appears therefore that aldolase B is a strong marker of hepatocytic differentiation, whose expression is modified in cancer, and which is also involved in a non exceptional metabolic disease. The availability of a specific DNA probe is essential for studying the mechanisms of Fructose Intolerance at the gene level and for analyzing the modifications of aldolase B synthesis during cancer and development. In a preceding work (8) we described the iden-

0006-291X/83 $1.50

tification of rat aldolase B cDNA clones. The rat aldolase B cDNA insert was used as a probe for isolating specific clones from a human liver cDNA library. In this paper we report the partial sequence analysis of human aldolase B cDNA as compared to the sequence of rat aldolase B cDNA recently published by Tsutsumi et al (9). In addition, the human probe is used for characterizing aldolase B mRNA and specific genomic DNA fragments.

MATERIALS AND METHODS

Tissue sampling, RNA purification, cell free synthesis and detection of neosynthesized aldolase B by immunoaffinity microchromatography (10-12), positive hybridization-selection (8), construction of the human liver ds cDNA library (13), Northern and Southern blot analyses (13) have been described in previous papers.

The human ds cDNA library was transferred to Whatman 540 filters (Masiakowsky, P., personal communication) and hybridized with purified, nick translated rat aldolase B cDNA insert. The recombinant plasmid A-4C9 was amplified and purified (14-16), then cut by Pst1 restriction enzyme. After electrophoresis in 1.5 % (w/v) agarose gel, the insert was transferred to NA-45 filter (DEAE-membrane Schleicher and Schuell) by electroblotting (17-18). The NA-45 membrane strip was stained with ethidium bromide and the insert band was cut out, washed and eluted (19-20). The eluate was diluted 10 times to adjust NaCl concentration to 0.1 M and DNA was precipitated by 10 mM spermine hydrochloride (21) and nick-translated (about 2 X 10^8 cpm/µg DNA). The Whatman 540 filters carrying the human cDNA clones were hybridized as described (8) for 18 h in the presence of 10,000 cpm per ml of labeled probe. In order to reduce the non specific hybridization, 200 µg/ml of sonicated, boiled total nucleic acids from E. Coli transformed by non recombinant pBR-322 were also added to the hybridization mixture.

The nucleotide sequences were determined according to the chemical method of Maxam and Gilbert (22).

RESULTS

Isolation and identification of human aldolase B cDNA clones. Only 4 human ds cDNA clones out of the 7200 tested were found to hybridize with the rat probe (Fig. 1). These clones were amplified and the plasmids were purified and analyzed by positive hybridization-selection (Fig. 2). The plasmids derived from the 7OH10 and 14F4 human clones both hybridized with a messenger RNA encoding human aldolase B. Restriction maps of the inserts are shown in fig. 3. The existence of two close BstN I and Taq I restriction sites on both inserts allows one to determine the relative position of these inserts. About 890 bases of the human aldolase B mRNA sequence were covered by the overlapping 7OH10 and 14F4 clones. The other clones did not hybridize with any translatable mRNA probably because of their small size (less than

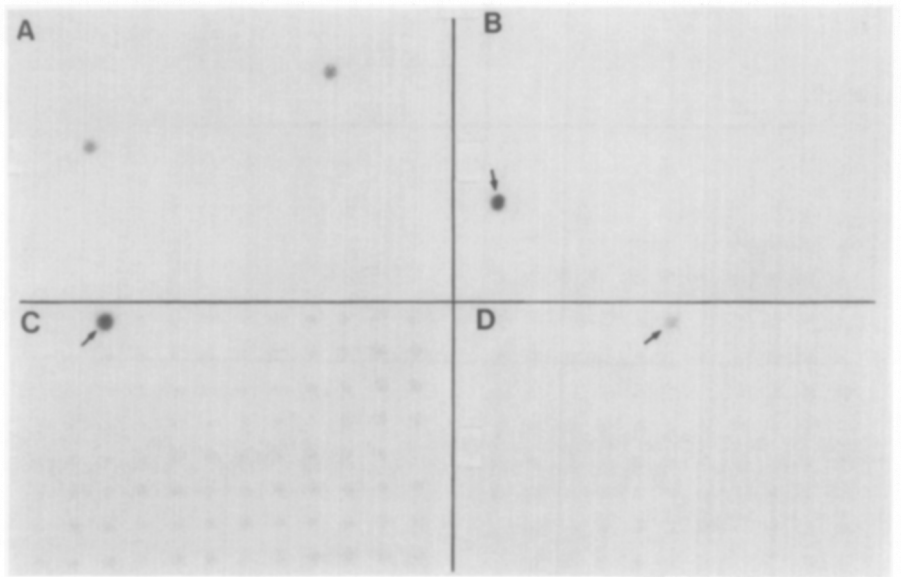

Figure 1 : Screening of the human cDNA library with rat aldolase B labeled cDNA. A : control hybridization : the filter containing the two clones of rat aldolase B was hybridized in the same mixture than those from the human library. The two rat aldolase B cDNA clones were specifically detected ; B, C and D : detection of human aldolase B cDNA clones, respectively 66FI, 70H10 and 14F4.

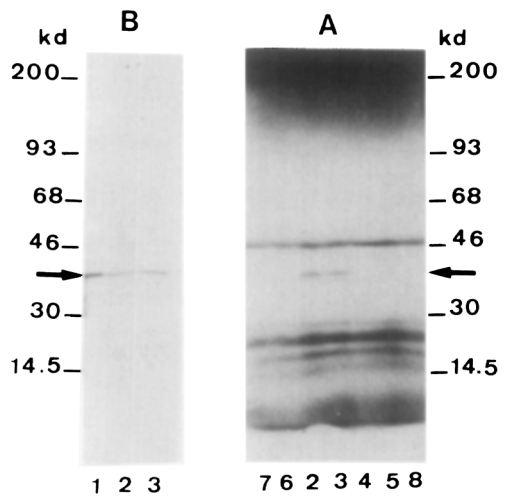

Figure 2 : Positive hybridization-selection analysis. Plasmid DNA was bound to nitrocellulose filters then hybridized with poly A-containing human liver mRNA as indicated in "Materials and Methods". Eluted mRNA was translated in 25 µl of a reticulocyte cell free system, 5 µl of the translation mixture being dissociated and analyzed by SDS polyacrylamide gel electrophoresis (A) ; translation mixture showing specific pattern was then passed through a micro-column containing antihuman aldolase B antibodies (B). 1 : ^{14}C labeled human aldolase B ; 2 to 7 : translation products obtained under direction of RNA eluted from filters bound to plasmid 70H10 (2), plasmid 14F4 (3), plasmid 66F1 (5), plasmid 7A6 (6), non recombinant PBR 322 (7) ; 8 : translation blank, without addition of exogenous RNA.

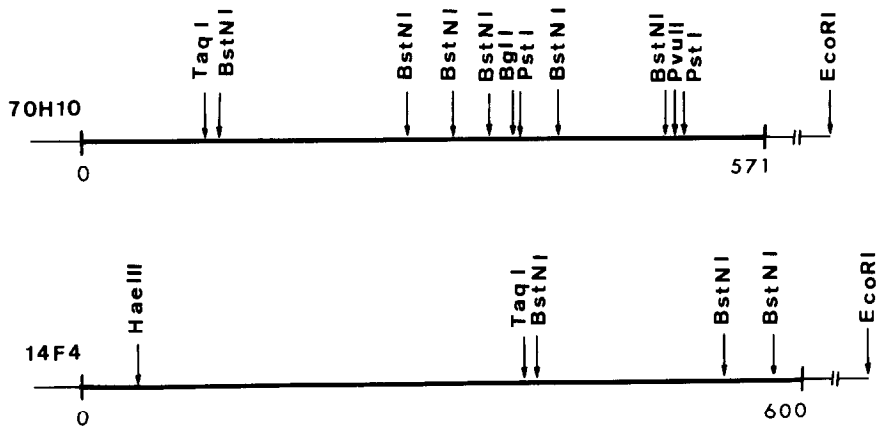

Figure 3 : Restriction endonuclease maps of human aldolase B cDNA inserts (70H10 and 14F4). The maps were oriented from 3' to 5' end relative to aldolase B mRNA (from left to right). BstN I and Taq I restriction sites allow their relative positioning.

200 bp ; not shown) as compared to the 70H10 and the 14F4 clones (571 and 600 bp respectively).

Northern blot analysis of human and rat aldolase B mRNA with the human 14F4 and 70H10 probes. Fig. 4 shows that the human 14F4 and 70H10 probes both hybridized with a single mRNA species, whose length approximated 1.8 kb. The amount of aldolase B mRNA in adult human liver far exceeded that of a four-month-old fetus liver. No aldolase B mRNA was detected in human adult skeletal muscle. The human probes cross-hybridized with rat aldolase B messenger RNA, whose size was identical to that of its human counterpart. Cross-hybridization was higher with the 70H10 than with the 14F4 probe. It suggested that the former insert contained a longer coding sequence than the latter, since the coding sequences are more conserved during evolution than the 3' non coding sequences (23). Therefore, the nucleotide sequence of the 70H10 insert rather than that of the 14F4 probe was determined.

Nucleotide sequence of the 70H10 insert. Fig. 5 shows the nucleotide sequence and the deduced amino-acid sequence of the 70H10 insert. The nucleotide fragment encodes the last 127 amino-acid residues of the enzyme and contains a 187 bp stretch after the translation termination codon UAG at the 3' end. The deduced protein sequence agrees well with the known sequence of

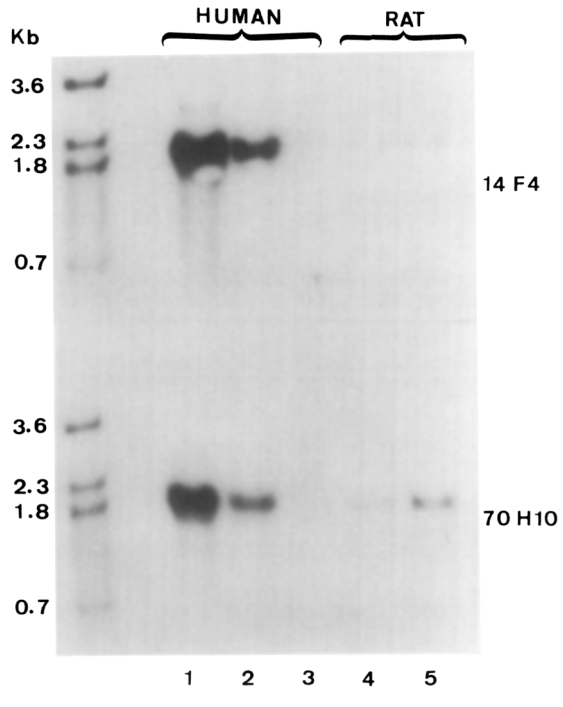

Figure 4 : Size determination of human aldolase B mRNA. 2 µg of poly(A)$^+$ RNA from various tissues of man and rat were subjected to agarose-formaldehyde gel electrophoresis. The RNA was transferred to the nitrocellulose filters and the RNA bands were detected by autoradiography after hybridization with ^{32}P-labeled human aldolase B cDNAs (70H10 and 14F4 respectively). The RNA samples were prepared from : 1 : adult human liver ; 2 : fetal human liver (4 months) ; 3 : skeletal human muscle ; 4 and 5 : rat liver (fasted and carbohydrate fed, respectively).

the C-terminal dodecapeptide of rabbit aldolase B (24), except for the 4th codon, preceding the UAG stop signal, which corresponds to a cysteine in man and to a serine in rabbit. The 70H10 cDNA clone encodes the last 5 residues of the active site, which are identical to that of other active sites in different species (25,26).

The 3' non coding extension of the insert does not reach the poly A tail. Nevertheless, a putative polyadenylation signal (AATAAA) is found 37 bases upstream the 3' end of the fragment.

Southern blot analysis. One single restriction fragment was detected with both 70H10 and 14F4 probes using Pst I, EcoR I, BamH I and Bgl I restriction enzymes. The length of the restriction fragments ranged from 5.4 kbp for Bgl I to 13.8 kbp for BamH I. Pvu II gives 3 fragments with both

```
His Ala Cys Thr Lys Lys Tyr Thr Pro Glu Gln Val Asp Met Ala Thr Val Thr Ala Leu
CAT GCC TGC ACC AAG AAG TAT ACT CCA GAA CAA GTA GAT ATG GCC ACC GTA ACA GCT CTC
                        .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                              60

His Arg Thr Val Pro Ala Ala Val Pro Gly Ile Cys Phe Leu Ser Gly Gly Met Ser Glu
CAC CGT ACT GTT CCT GCA GCT GTT CCT GGC ATC TGC TTT TTG TCT GGT GGC ATG AGT GAA
   .  .           .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                             120

Glu Asp Ala Thr Leu Asn Leu Asn Ala Ile Asn Leu Cys Pro Leu Pro Lys Pro Trp Lys
GAG GAT GCC ACT CTC AAC CTC AAT GCT ATC AAC CTT TGC CCT CTA CCA AAG CCC TGG AAA
         .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                             180

Leu Ser Phe Ser Tyr Gly Arg Ala Leu Gln Ala Ser Ala Leu Ala Ala Trp Gly Gly Lys
CTA AGT TTC TCT TAT GGA CGG GCC CTG CAG GCC AGT GCA CTG GCT GCC TGG GGT GGC AAG
      .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                             240

Ala Ala Asn Lys Glu Ala Thr Gln Glu Ala Phe Met Lys Arg Ala Met Ala Asn Cys Gln
GCT GCA AAC AAG GAG GCA ACC CAG GAG GCT TTT ATG AAG CGG GCC ATG GCT AAC TGC CAG
               .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                             300

Ala Ala Lys Gly Gln Tyr Val His Thr Gly Ser Ser Gly Ala Ala Ser Thr Gln Ser Leu
GCG GCC AAA GGA CAG TAT GTT CAC ACG GGT TCT TCT GGG GCT GCT TCC ACC CAG TCG CTC
  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
                                                                             360

Phe Thr Ala Cys Tyr Thr Tyr ***
TTC ACA GCC TGC TAT ACC TAC TAG GGTCCAATGCCCGCCAGCCTAGCTCCAGTGCTTCTAGTAGGAGGG
            .
                                                                             429

CTGAAAGGGAGCAACTTTTCCTCCAATCCTGGAAATTCGACACAATTAGATTTGAACTGCTGGAAATACAACACATGT
                                                                             507

TAAATCTTAAGTACAAGGGGGAAAAAATAAATCAGTTATTGAAACATAAAAATGAATACCAAGG
                                                                             571
```

Figure 5 : Partial nucleotide sequence of human aldolase B mRNA and the corresponding amino-acid sequence. The sequence was determined from cDNA insert of 7OH10. The hexanucleotide AATAAA is underlined. Dots under the nucleotide sequence indicate the different nucleotides between man and rat. Lines on the amino-acid sequence indicate the amino-acids differing from those of rat aldolase B.

probes, 2 of them (2.7 and 2.4 kbp-long respectively) being common and one specific for each probe (2.1 kbp for 7OH10 and 5.25 kbp for 14F4). The 7OH10 probe allowed to detect three Bgl II fragments (15, 5.2 and 3.1 kbp), only the last two hybridizing with 14F4, and two Hind III fragments (2.6 and 1.9 kbp), the 2.6 kbp fragment being the only one to hybridize with 14F4 (Fig. 6).

The same type of pattern was found with leukocyte DNA derived from three unrelated healthy volonteers (not shown).

DISCUSSION

When we started studying the human aldolase B gene organization, we considered two alternative strategies : either screening the human cDNA library directly with a human probe (namely, an enriched ss cDNA probe prepared from partially purified human aldolase B mRNA), or cloning rat aldolase B cDNA first so as to use it as a probe for screening the human library.

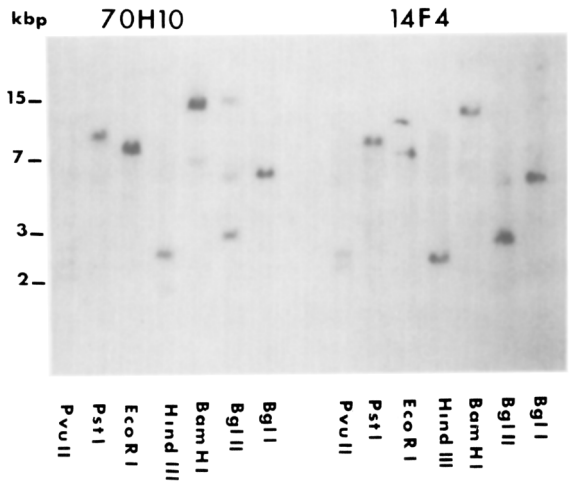

Figure 6 : Southern blot analysis of human leukocyte DNA with the 70H10 and 14F4 inserts, respectively, as probes. 10 µg DNA were digested to completion with the various restriction enzymes indicated under the figure, then electrophoresed in a 0.7 % (w/v) agarose gel for 14 hours at 1.5 volts/cm. After transfer to nitrocellulose, the blot was hybridized in the nick translated insert, as indicated in Materials and Methods (10^6 cpm/ml, 18 hours hybridization). Washing was performed under stringent conditions (0.1 % SSC + 0.5 % SDS at 65° C). Exposure time : 36 hours at -80° C with intensifying screens.

Since aldolase B mRNA was found to be inducible by a carbohydrate-rich diet (8), we preferred the second approach. The procedure adopted for cloning rat liver aldolase mRNA is described elsewhere (8). It consisted in a differential screening of the rat liver cDNA library with ss cDNA probes reverse transcribed from either starved or carbohydrate fed rat liver mRNAs.The number of aldolase B cDNA clones found in the human library was low (4 out 7200, i.e. 0.06 %). Some clones may have been missed because of their very short insertion or because they corresponded to upstream coding sequences which do not cross-hybridize with our rat cDNA probe (1050 base long). Moreover, the liver used for making the human cDNA library was derived from a fasted patient.

The two longest human aldolase B cDNA clones were overlapping. They covered a 820 base long portion of the mRNA sequence. We determined the nucleotide sequence of 70H10 insert which contains the longest coding sequence, while most of the 14F4 insert is composed of the 3' non coding region (about 400 non coding and 200 coding bases, as deduced from the alignment of the inserts).

The deduced protein sequence of the 127 C-Terminal residues was similar in man and rat (9). They only differed by 9 residues as shown in Fig. 5. At the nucleotide level, a 87 % homology was found in the last 381 coding bases between man and rat aldolase B mRNA, only 48 codons being different between the two species. Nine codons differing from the rat counterpart by either the first base or the second were found to encode different amino-acid residues. At position 66 and 201, two codons encoded the same residue, i.e. Arg, although they differed by two bases (the first and the third base of the codon).

By contrast, the 3' non coding regions are very different in man and rat as shown for a number of mammalian mRNAs (23). This is the reason why the 14F4 probe (composed of a majority of 3' non coding bases) exhibited a lower cross-hybridization with rat mRNA than the 70H10 insert. This non coding region is 387 bases long in rat. If we assume that the same is true in man (which is highly probable in view of the identical length of the rat and human RNAs), this means that the 14F4 clone contains all or most of the 3' coding extension.

One striking difference between the human and rat aldolase B mRNA 3' non coding regions consists in the location of the polyadenylation site. One canonical polyadenylation site (AATAAA) in man is located far from the 3' end of the mRNA whereas this site in rat is located as usual about 15 bases upstream the poly-A tail. We can speculate that a similar polyadenylation site exists in human mRNA, the proximal site being not used.

The hybridization of the cDNA probes with genomic DNA restriction fragments indicates that the gene for aldolase B may be present as a single copy. Indeed, one or two restriction fragments only were detected with most of the enzymes used. When several fragments could be observed, some of them hybridized specifically (or preferentially) with one of the two probes. This suggests that these multiple fragments result from the cleavage of different internal sites rather than from the splitting of different genes. This is particularly obvious with Pvu II which is known to cleave within the cDNA sequence.

Finally, the pattern of the restriction fragments was found to be similar regardless of the washing conditions, (stringent or non stringent), a feature which argues against the presence of partially cross-hybridizing sequences.

In conclusion, we have obtained the partial sequence of a cDNA clone corresponding to human aldolase B. This probe is now used for characterizing genomic clones, with the aim of analyzing the genetic defect in Hereditary Fructose Intolerance.

REFERENCES

1. Rutter, W.J., Rajkumar, T., Penhoet, E., Kochman, M. and Valentine, R. (1968) Ann. New York Acad. Sci. 151 : 102-117.
2. Hers, H.G. and Joassin, G. (1961) Enzymol. Biol. Clin. 1 : 4-14.
3. Schapira, F., Nordmann, Y. and Dreyfus, J.C. (1968) Rev. Fr. Et. Clin. Biol. 13 : 267-269.
4. Schapira, F., Grégori, C. and Hatzfeld, A. (1977) Clin. Chim. Acta 78 : 1-7.
5. Schapira, F., Hatzfeld, A. and Weber, A. (1975). In "Isozymes", C.L. Markert (ed), Academic Press (New York), III, pp 987-1003.
6. Schapira, F. (1973) Isozymes and cancer, in Klein-G, Weinhouse, S. (eds) : Advances in Cancer Research, vol 18, Academic Press (New York), pp 77-153.
7. Schapira, F., Dreyfus, J.C. and Schapira, G. (1963) Nature 220 : 995-997.
8. Simon, M.P., Besmond, C., Cottreau, D., Weber, A., Chaumet-Riffaud, P., Dreyfus, J.C., Sala Trépat, J., Marie, J. and Kahn, A. J. Biol. Chem. In press.
9. Tsutsumi, K., Mukai, T., Hidaka, S., Miyahara, H., Tsutsumi, R., Tanaka, T., Hori, K. and Ishikawa, K. (1983) J. Biol. Chem. 248 : 6537-6542.
10. Besmond, C., Benarous, R. and Kahn, A. (1981) Biochem. Biophys. Res. Commun. 103 : 587-594.
11. Grégori, C., Besmond, C., Kahn, A. and Dreyfus, J.C. (1982) Biochem. Biophys. Res. Commun. 104 : 369-375.
12. Kahn, A., Cottreau, D., Daegelen, D. and Dreyfus, J.C. (1981) Eur. J. Biochem. 116 : 7-12.
13. Uzan, G., Park, I., Besmond, C., Maessen, G., Frain, M., Sala Trépat, J., Zakin, M.M. and Kahn, A. (1983) Submitted.
14. Norgard, M.W., Emigholz, K. and Monahan, J.J. (1979) J. Bacteriol. 138 : 270-272.
15. Wahl, G.M., Padgott, R.A. and Stark, G.R. (1979) J. Biol. Chem. 254 : 8679-8699.
16. Birboim, H.C. and Poly, J. (1979) Nucleic Acids Res. 7 : 1513-1523.
17. Stellwag, E.J. and Dahlberg, A.E. (1980) Nucleic Acids Res. 8 : 253-264.
18. Bittner, M., Kupferer, P. and Morris, C.F. (1980) Anal. Biochem. 102 : 459-471.
19. Winberg, G. and Hammarrkjold, M.L. (1980) Nucleic Acids Res. 8 : 253-264.
20. Dretzen, G., Bellard, M., Sassone-Corsi, P., and Chambon, P. (1981) Anal. Biochem. 112 : 295-298.
21. Hoopes, B.C. and Mc Clure, W.R. (1981) Nucleic Acids Res. 9 : 5493-5504.
22. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology 65 : 499-560.
23. Cleveland, D.W., Lopata, M.A., Mac Donald, R.I., Cowan, N.J., Rutter, W.J. and Kirschner, M. (1980) Cell 20, 95-105.
24. Hannapel, E., Mac Gregor, J.S. and Davoust, S.G. (1981) Feder. Proc. 40 : 1679-1683.
25. Lai, C.Y. (1975) Arch. Biochem. Biophys. 166 : 358-368.
26. Ting, S.M., Lai, C.Y. and Horecker, B.L. (1971) Arch. Biochem. Biophys. 144 : 476-484.