The EMBO Journal vol.4 no.10 pp.2665–2668, 1985

Isolation and structural organization of the *Neurospora crassa* copper metallothionein gene

Karl Münger, Ursula A.Germann and Konrad Lerch

Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Communicated by W.Schaffner

The *Neurospora crassa* copper metallothionein gene was cloned and its complete nucleotide sequence is reported. Enriched metallothionein mRNA was used as a template for cDNA synthesis, primed by a metallothionein-specific, synthetic undecanucleotide. The sequence of the cDNA obtained allowed the synthesis of a unique 21-mer which was used to screen a genomic DNA library of *N. crassa*. In agreement with the published amino acid sequence, the gene codes for a polypeptide 26 amino acid residues in length. The coding region is interrupted by a small intron (94 nucleotides). The gene structure is compared with those of mammalian metallothioneins. In both cases, the coding regions are split by introns, the intron-exon boundaries, however, are in different positions. The *Neurospora* copper metallothionein gene is, to our knowledge, the smallest gene interrupted by an intron isolated so far.

Key words: ascomycetes/genomic cloning/intron-exon structure/metal-binding proteins

Introduction

Metallothioneins (MTs) are a class of low mol. wt., heavy metal-binding proteins, containing high amounts of cysteine. They have been isolated from various organisms including vertebrates, invertebrates and fungi (Kägi and Nordberg, 1979). In mammals MTs have a typical chain length of 61 amino acids and they are synthesized in response to heavy metal ions and glucocorticoids (Kägi *et al.*, 1984).

We have previously shown that the ascomycete *Neurospora crassa* contains a small copper-binding protein, 25 amino acid residues in length (Lerch, 1980). It is induced by copper only and does not bind any other metal ions *in vivo*. The protein exhibits a pronounced amino acid sequence homology to the amino-terminal part of mammalian MTs. In particular, the sequence positions of the seven cysteinyl residues are conserved (Lerch, 1980). The ability of *Neurospora* MT to bind heavy metals other than copper has recently been demonstrated *in vitro* (Beltramini *et al.*, 1984). The biological functions of *Neurospora* metallothionein remain unclear although it has been assigned to play a role in copper storage and detoxification. It was, moreover, demonstrated *in vitro*, that it acts as an efficient copper donor to the apo-forms of copper-containing proteins such as *Neurospora* tyrosinase and *Carcinus* hemocyanin (Lerch and Beltramini, 1983).

The structures of mouse and human MT genes have been investigated in great detail (Glanville *et al.*, 1981; Karin and Richards, 1982). The genes are interrupted by two introns, the positions of which are conserved. Both for the induction by heavy metals and glucocorticoids, transcriptional control has been demonstrated (Durnam and Palmiter, 1981; Hagar and Palmiter, 1981). Their effects are mediated, however, through the activation of different promoters (Karin *et al.*, 1984a). Both promoters have been characterized in the cases of the human MT-II (Karin *et al.*, 1984a, 1984b) and the mouse MT-I (Stuart *et al.*, 1984; Carter *et al.*, 1984) genes.

More recently, a gene coding for a cysteine-rich, copper-inducible protein has been cloned from the yeast *Saccharomyces cerevisiae*, and its primary structure elucidated (Butt *et al.*, 1984a, 1984b; Karin *et al.*, 1984c). It shows, however, only a remote sequence homology to mammalian MTs.

Both in view of the close structural similarity of *Neurospora* copper MT to mammalian MTs and the observed differences in the mode of induction, we set out to clone the *Neurospora* MT gene. Because *N. crassa* represents a simple, genetically well-established eukaryotic organism, this system seems to be highly suitable to investigate the molecular mechanisms of MT induction by copper and the overall biological function(s) of MTs.

Here, we report the cloning and the nucleotide sequence of the *N. crassa* copper metallothionein gene. Its structure is compared with those of yeast, human and mouse MTs.

Results

Primer extension and screening of a genomic DNA library

For specific cDNA synthesis, a primer encoding four amino acids (residues 11 – 14) in the sequence of *Neurospora* copper MT (Lerch, 1980) was prepared (Figures 1a and b). This primer was selected over the other seven possibilities based on the published codon usage of the trp-1 (Schechtman and Yanofsky, 1983) and histones H-3 and H-4 (Woudt *et al.*, 1983) genes from *N. crassa*. High mol. wt. cDNA obtained from primer extension reactions was subjected to chemical DNA sequence analysis

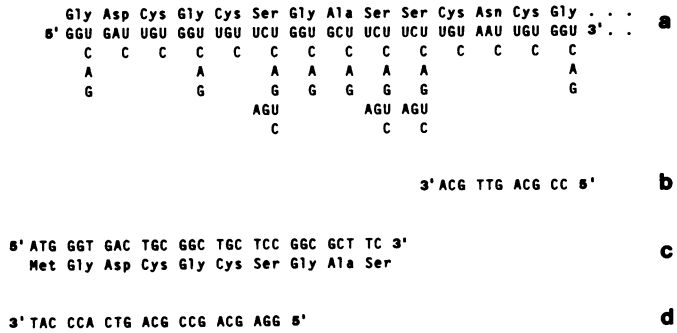

Fig. 1. Cloning strategy of the *N. crassa* copper MT gene. Based on the known amino acid sequence (Lerch, 1980) and the possible mRNA sequences **(a)**, a DNA primer was synthesized **(b)** and used for specific cDNA synthesis with enriched *Neurospora* MT mRNA (Münger *et al.*, 1983). High mol. wt. primer extension products were subjected to sequence analysis (Maxam and Gilbert, 1980) **(c)**. The amino acid sequence derived was in complete agreement with the amino terminus of *Neurospora* copper MT. This allowed the synthesis of a unique 21-mer **(d)** which was used to screen a genomic *Neurospora* DNA library (Schechtman and Yanofsky, 1983).

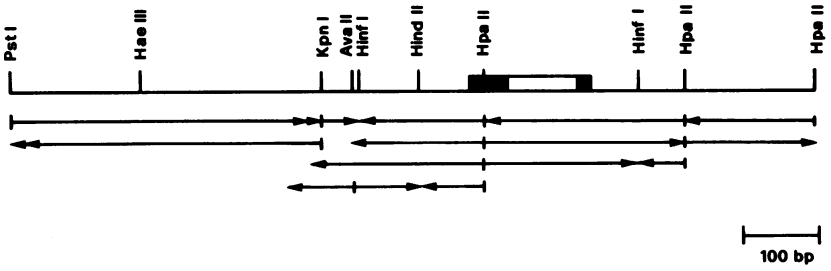

Fig. 2. Restriction map of the *Neurospora* copper MT gene. Only the sites used to determine the sequence are shown. ■, Exons of the MT gene, arrows indicate the direction and extent of sequencing. For chemical sequencing (Maxam and Gilbert, 1980) (→), restriction fragments were labelled at their 5′ ends and either subjected to strand separation (Maxam and Gilbert, 1980) or cut with another restriction enzyme. Chain termination sequencing (Sanger *et al.*, 1979) (⇀) was performed using the wild-type phages mp18 and mp19 (Norrander *et al.*, 1983).

```
                                    -450                                -400
CTGCAGGTTCTGAGACATCACATGAACATTGCAACCATTCACGCTTTTGGACATCAACCATCTCTATCGTCATCGCCTTACATACTGTACATCCTATGTA

                                    -350                                -300
CACTACACTGTAACGGTAGTGTACGATGAGGGTATTTCGCCGCGTGGGCTGGGCGGTAGGATGCCAAGCAGGTGGCCCAGAACGCCGCCGCACAAAAGAC

                                    -250                                -200
GCAGCCACCACGGCAGCGAACACCCGCCTCCACTCTTGATGCAGCCGCTACGGCACAAATATCGCAAACCGTGCGGGCGGACGTCCCCGATAGGCTCGGA

                                    -150                                -100
TCGTCGCTAACATGCCATGCCATGCCATGCCATGCCGTGTCGTGGAATGGAAGGTAACAGACAGTCACAGTCACAGTTCCGAACGGTGCGTAGCAGTGTA
                                                                        +1
                                    -50                                 ▼
GTGTACGTGACAGCAGGAACAAGGTACCGTACGTAGGCATCATTCAGGGTATATAAAGCCACGGGACCCTGACTCCTCCTCAAATTCCTTTACTTGTCAT

                                    +50                                 +100
CAACCGACGAACTAAATCATCACTGCATTGAAACACTCACACAAAGACAACATCAGTCTTTCGTCAACCTACACAACCACAACAACAACTCAATTCCTCT

                                    +150
CAAAAGAACAACCAGTTCTATCAAA  ATG GGT GAC TGC GGC TGC TCC GGC GCT TCT TCC TGC AAC TGC GGC TCT GGC T  GT

                           Met Gly Asp Cys Gly Cys Ser Gly Ala Ser Ser Cys Asn Cys Gly Ser Gly C

        +200                                +250
AAGTTTTTTTTACTCCCCCCTCCCCCGTCGCCTGTCCGATCCCATCTTCAACTACTGGTACTGGCATGGTACTAACCTTTGAAACTTTCTAG  GC TCT

                                                                        ys Ser

                    +300                                +350
TGC TCC AAC TGC GGC AGC AAG TAA  ATTTACCTCTCCGCACTGCTCAATGCATCCACGTACCCGAGCGACCATGAGCTTTCTTTGAATCACT

Cys Ser Asn Cys Gly Ser Lys ***

                    +400                                +450
AGTCGCAGCGACAGCGCGTAAGGTTGGGAGACCTTTGGCTGCCTGACATACCGCCGGGGGGTTTTGTTCCTATTTCGCCATGACGCCCATTGGCACTTTC

                    +500                                +550
CCCCTTTGGAGAGATATGGATTAGATGAAAACAAAACAAGTCTTTTCCCTCCTTCACAGATAATATTACTCTTGAGATTATGTTTTGATTTGACTGTGAT

                    +600
GACAACGCCCGAACTCAATGCTTATGATCCTCTACGCC
```

Fig. 3. Primary structure of the *N. crassa* copper MT gene. The promoter sequence (Breathnach and Chambon, 1981) is boxed, transcription initiation sites are indicated by ▼. The sequence in the intron, possibly involved in splicing (Langford *et al.*, 1984) is indicated by a dashed box, the putative transcription termination sequence (Benoist *et al.*, 1980) is underlined (see text).

(Figure 1c). The amino acid sequence deduced is in complete agreement with the published primary structure of *Neurospora* copper MT (Figure 1c). This allowed the synthesis of a unique 21-mer (Figure 1d) which was used as a hybridization probe to screen a genomic *Neurospora* DNA library.

Structure and organization of the gene

The strategy of DNA sequencing, together with a restriction map of the 1105 nucleotides-long *Pst*I-*Hpa*II fragment are presented in Figure 2. The sequence of the *Kpn*I-*Hpa*II fragment was elucidated by means of chemical sequencing (Maxam and Gilbert, 1980) of the three *Hpa*II sub-fragments as outlined in Figure 2. The sequence of the *Pst*I-*Kpn*I fragment was determined by chain termination sequencing (Sanger *et al.*, 1979) of both strands, using the M13 wild-type phages mp18 and mp19, respectively. Every sequence was determined at least twice. The complete nucleotide sequence of the *Pst*I-*Hpa*II fragment containing the structural gene as well as 5′- and 3′-flanking regions is presented in Figure 3. Nuclease S1 mapping to determine the start of transcription of the *Neurospora* MT gene yielded two nuclease-resistant bands, beginning at the *Hpa*II site at position +147 and ending at C −1 and A +1 (data not shown). At positions C +20 and A +21 two weak S1-resistant bands were also detected. Control experiments, omitting either MT mRNA or enzyme, yielded no detectable signals in this region.

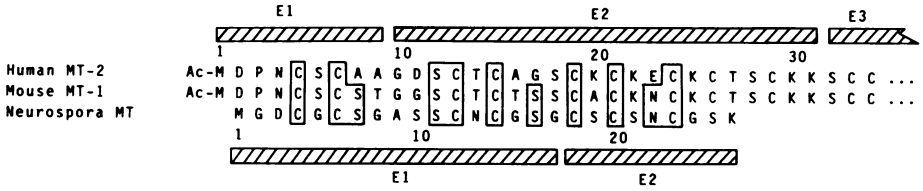

Fig. 4. Comparison of the human MT-2 (Karin and Richards, 1982), mouse MT-1 (Glanville *et al.*, 1981) and *Neurospora* copper MT genes. Amino acid residues are abbreviated using the one-letter code: A, Ala; C, Cys; D, Asp; E, Glu; G, Gly; K, Lys; M, Met; N, Asn; P, Pro; S, Ser; T, Thr; V, Val. Identical amino acid residues are boxed, exons are indicated by hatched bars.

Table I. Comparison of codon usage in the *N. crassa* copper MT gene with *trp-1* (Schechtman and Yanofsky, 1983) and histone H-3 and H-4 (Woudt *et al.*, 1983) genes

		MT	*trp-1*	H-3/H-4
Cys	TGT	–	–	–
	TGC	7	10	1
Ser	TCT	3	14	4
	TCC	3	20	7
	TCA	–	1	–
	TCG	–	8	1
	AGT	–	4	–
	AGC	1	12	2
Gly	GGT	1	18	13
	GGC	5	34	9
	GGA	–	7	2
	GGG	–	3	–

Discussion

Computer analysis of the nucleotide sequence obtained, demonstrated that the reading frame for *Neurospora* copper MT spans a total of 172 nucleotides. The coding region is interrupted by a short, 94-bp intron (Figure 3). A polypeptide of 26 amino acids is encoded, the sequence of which is, with the exception of the amino-terminal methionine, in complete agreement with the published amino acid sequence (Lerch, 1980). Recent h.p.l.c. analysis of *Neurospora* copper MT revealed the existence of isoforms with the N-terminal methionine still present (K. Lerch, unpublished data). The gene exhibits 53% homology to the 5′ part of the mouse MT-I gene when the nucleotide sequences are aligned as shown in Figure 4.

The coding block is significantly G+C rich (64%) compared with the flanking regions and the intron (49%). The codon usage for the most abundant amino acids cysteine, glycine and serine is presented in Table I. All seven cysteines are encoded by TGC. This strong preference is in agreement with other *Neurospora* genes sequenced so far (Schechtman and Yanofsky, 1983; Woudt *et al.*, 1983). The same finding has also been reported for the mouse MT-I (Glanville *et al.*, 1981) and the human MT-II (Karin and Richards, 1982) gene. Similarly, the codon usage for glycine and serine are in agreement with other *Neurospora* genes (Schechtman and Yanofsky, 1983; Woudt *et al.*, 1983; Alton *et al.*, 1982). All the data taken together show that the *Neurospora* MT gene exhibits a strong preference for codons ending in G or C (81%).

The position of the transcription initiation site at A +1 established by nuclease S1 mapping, is in agreement with the consensus sequence PyCAPy (where A denotes position +1 and Py a pyrimidine nucleoside) derived from a comparison of different mRNA start sites (Corden *et al.*, 1980). The sequence around A +21 also agrees with the consensus sequence mentioned above. At this position, two weak nuclease S1-resistant bands were also detected. Therefore we suggest, that a major transcription initiation site is located at A +1 and a minor one at A +21.

A typical TATA box (Breathnach and Chambon, 1981) is centered at position −45. A second element which is present in many eukaryotic promoters in the −100 region (the CAAT box) (Breathnach and Chambon, 1981) is absent in the *Neurospora* MT gene.

The metal regulation sequences of the mouse MT-I and human MT-II genes were identified by an extensive series of mutation studies (Karin *et al.*, 1984b; Carter *et al.*, 1984). A 12-bp DNA motif, repeated several times in these promoter regions, conferred metal regulation to thymidine kinase when inserted into the promoter of this gene (Stuart *et al.*, 1984). A computer search in the 5′ region of the TATA box in the *Neurospora* MT gene did not reveal sequences with extensive homology to the metal regulation elements of mammalian MT genes. The absence of homologous elements, however, may reflect the different mode of induction reported for *Neurospora* copper MT. No sequence homology in the 5′ regions could be detected when *Neurospora* and yeast copper MT genes were compared. This result is not too surprising since the primary structures of the gene products also show no significant correspondence. Taken together, these data suggest that *Neurospora* and yeast copper MTs did not evolve from a common ancestor.

As in many yeast mRNAs (Zaret and Sherman, 1982), the typical poly(A) addition signal AATAAA (Proudfoot and Brownlee, 1976) is absent in the *Neurospora* MT gene. A sequence centered at position +515 (TTTTCCCTCC), however, has eight out of 10 bases in common with the model sequence TTTTCACTGC for the 3′-terminal region of eukaryotic genes (Benoist *et al.*, 1980).

The coding region of the *Neurospora* MT gene is interrupted by an intron. In agreement with the rule proposed by Breathnach *et al.* (1978) it begins with the dinucleotide GT and ends with AG. Fifteen bases upstream the 3′ intron exon junction, the sequence TACTAACC was detected. It is identical in seven of eight bases with the TACTAACA box, a common feature (Langford and Gallwitz, 1983; Pikielny *et al.*, 1983), and essential splicing signal sequence in yeast introns (Langford *et al.*, 1984). The sequence found is complementary in five positions to the 5′ intron exon junction T:GTAAGTT. Similar sequences were reported in a number of other split *Neurospora* (Woudt *et al.*, 1983; Arends and Sebald, 1984) and fungal (Dons *et al.*, 1984) genes. The splicing mechanism of yeast pre-mRNA, involving lariat structures (Domdey *et al.*, 1984; Rodriguez *et al.*, 1984) may therefore also be functional in splicing of certain *Neurospora* and other fungal pre-mRNAs.

The *Neurospora* MT gene is, to our knowledge, the smallest gene interspersed by an intron isolated so far. Whether or not the intron-exon structure reflects functional domains in copper binding is being investigated by n.m.r. spectroscopy. In contrast

to the *Neurospora* MT gene, those from mammals are made up of three exons (Glanville *et al.*, 1981; Karin and Richards, 1982) (Figure 4). For mammalian MTs, it has been shown by ^{113}Cd n.m.r. spectroscopy (Otvos and Armitage, 1980), that the seven metals are bound in two distinct domains. The four-metal domain A involves 11 cysteines of the C-terminal part (corresponding to exon 3); the three-metal domain B comprises nine cysteines of the amino-terminal part (corresponding to exon 1 + 2) of the molecule (Boulanger *et al.*, 1983). In view of the striking amino acid sequence homology between *Neurospora* copper MT and the amino-terminal part of mammalian MTs, it is of interest to compare the structures of corresponding parts of the genes. As shown in Figure 4, the DNA sequences are interrupted by an intron in both cases. However, the intron-exon boundaries are located at different sequence positions. In *Neurospora* copper MT, the intron splits the sequence at amino acid residues 17/18; in human and mouse MTs at amino acid residues 9/10. Following the idea that fungal, invertebrate and vertebrate MTs evolved from a common ancestor (divergent evolution) (Kägi *et al.*, 1984; Lerch, 1980) the course of evolution may be more complex than proposed by Glanville *et al.* (1981) based on primary and secondary structures of the mouse MT-I gene and its mRNA, respectively. The possibility of convergent evolution of fungal, invertebrate and vertebrate MTs may, however, not be ruled out, until more structures of MT genes become available.

Materials and methods

Synthesis of oligodeoxynucleotides

Oligodeoxynucleotides were synthesized on a manual solid-phase DNA synthesizer (BACHEMGENTEC) according to the phosphotriester method (Miyoshi *et al.*, 1980a, 1980b). The fully deprotected oligomers were purifed by gel-filtration on Sephadex G-25 (Pharmacia), followed by electrophoresis on 8 M urea/20% polyacrylamide gels and chromatography on NACS-52 (BRL).

Isolation of RNA and oligonucleotide-primed cDNA synthesis

mRNA was isolated from copper-exposed mycelium of *N. crassa* wild-type strain (Fungal Genetic Stock Center #321, Arcata CA) by the method of Chirgwin *et al.* (1979), followed by poly(A) selection on oligo(dT)-cellulose (Aviv and Leder, 1972). After size fractionation on isokinetic sucrose density gradients (Noll, 1967) and *in vitro* translation of individual fractions, the size of *N. crassa* MT mRNA was estimated to be 7.6S (Münger *et al.*, 1983). Taking this mRNA as a template, cDNA was synthesized essentially as described by Nathans and Hogness (1983). After size fractionation on Sephacryl S-300 (Pharmacia), high mol. wt. primer extension products were subjected to chemical sequence analysis (Maxam and Gilbert, 1980).

Screening of a genomic library

A library of genomic *Neurospora* DNA was generously provided by M. Schechtman, Syracuse University (Schechtman and Yanofsky, 1983). The library was screened according to Hanahan and Meselson (1980) using a synthetic oligodeoxynucleotide, ^{32}P-labelled at its 5′ end, as a hybridization probe.

Restriction mapping

Restriction enzymes obtained from Boehringer Mannheim, Amersham, or Anglian Biotechnology were used according to the protocols supplied. Digests were analyzed on 6% polyacrylamide or 1% agarose gels. DNA in agarose gels was denatured and blotted to nitrocellulose sheets (Schleicher and Schuell) (Southern, 1975).

Nuclease S1 mapping

Nuclease S1 mapping was performed according to Berk and Sharp (1977) as modified by Weaver and Weissmann (1979).

DNA sequencing

The sequence of the 1.1-kb *Pst*I-*Hpa*II fragment was determined by a combination of chemical (Maxam and Gilbert, 1980) and chain termination (Sanger *et al.*, 1979) methods using the M13 wild-type phages mp18 and mp19 (Norrander *et al.*, 1983). DNA sequences were analyzed with a computer program (Queen and Korn, 1980).

Acknowledgements

We thank Dr. M. Schechtman (Syracuse University) for providing a genomic DNA library from *N. crassa* and B. Suter for the replicative forms of mp18 and mp19. This work was supported by Swiss National Science Foundation Grant No. 3.285-0.82 and by the Kanton of Zürich.

References

Alton,N.K., Buxton,F., Patel,V., Giles,N.H. and Vapnek,D. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1955-1959.
Arends,H. and Sebald,W. (1984) *EMBO J.*, **3**, 377-382.
Aviv,J. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.
Beltramini,M., Lerch,K. and Vašák,M. (1984) *Biochemistry (Wash.)*, **23**, 3422-3427.
Benoist,C., O'Hare,K., Breathnach,R. and Chambon,P. (1980) *Nucleic Acids Res.*, **8**, 127-142.
Berk,A.J. and Sharp,P.A. (1977) *Cell*, **2**, 721-732.
Boulanger,Y., Goodman,C.M., Forte,C.P., Fesik,S.W. and Armitage,I.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1501-1505.
Breathnach,R., Benoist,C., O'Hare,K., Gannon,F. and Chambon,P. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 4853-4857.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349-383.
Butt,T.R., Sternberg,E., Herd,J. and Crooke,S.T. (1984a) *Gene*, **27**, 23-33.
Butt,T.R., Sternberg,E.J., Gorman,J.A., Clark,D., Hamer,D.H., Rosenberg,M. and Crooke,S.T. (1984b) *Proc. Natl. Acad. Sci. USA*, **81**, 3332-3336.
Carter,A.D., Felber,B.K., Walling,M.-J., Jubier,M.-F., Schmidt,C.J. and Hamer,D.H. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7392-7396.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry (Wash.)*, **18**, 5294-5299.
Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C. and Chambon,P. (1980) *Science (Wash.)*, **209**, 1406-1414.
Domdey,H., Apostol,B., Lin,R.-J., Newman,A., Brody,E. and Abelson,J. (1984) *Cell*, **39**, 611-621.
Dons,J.J.M., Mulder,G.H., Rouwendal,G.J.A., Springer,J., Bremer,W. and Wessels,J.G.H. (1984) *EMBO J.*, **4**, 2101-2106.
Durnam,D.M. and Palmiter,R.D. (1981) *J. Biol. Chem.*, **256**, 5712-5716.
Glanville,N., Durnam,D.N. and Palmiter,R.D. (1981) *Nature*, **292**, 267-269.
Hager,L.J. and Palmiter,R.D. (1981) *Nature*, **291**, 340-342.
Hanahan,D. and Meselson,M. (1980) *Gene*, **10**, 63-67.
Kägi,J.H.R. and Nordberg,M. (1979) *Metallothionein*, published by Birkhäuser, Basel, Switzerland.
Kägi,J.H.R., Vašák,M., Lerch,K., Gilg,D.E.O., Hunziker,P., Bernhard,W.R. and Good,M. (1984) *Environ. Health Perspect.*, **54**, 93-103.
Karin,M. and Richards,R.I. (1982) *Nature*, **299**, 797-802.
Karin,M., Haslinger,A., Holtgreve,H., Cathala,G., Slater,E. and Baxter,J.D. (1984a) *Cell*, **36**, 371-379.
Karin,M., Haslinger,A., Holtgreve,H., Richards,R.I., Krauter,P., Westphal,H.M. and Beato,M. (1984b) *Nature*, **308**, 513-519.
Karin,M., Najarian,R., Haslinger,A., Valenzuela,P., Welch,J. and Fogel,S. (1984c) *Proc. Natl. Acad. Sci. USA*, **81**, 337-341.
Langford,C.J. and Gallwitz,D. (1983) *Cell*, **33**, 519-527.
Langford,C.J., Klinz,F.-J., Donath,C. and Gallwitz,D. (1984) *Cell*, **36**, 645-653.
Lerch,K. (1980) *Nature*, **284**, 368-370.
Lerch,K. and Beltramini,M. (1983) *Chem. Scripta*, **21**, 109-115.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Miyoshi,K., Miyaka,T., Hozumi,T. and Itakura,K. (1980a) *Nucleic Acids Res.*, **8**, 5473-5490.
Miyoshi,K., Huang,T. and Itakura,K. (1980b) *Nucleic Acids Res.*, **8**, 5491-5505.
Münger,K., Germann,U.A. and Lerch,K. (1983) *Experientia*, **39**, 651.
Nathans,J. and Hogness,D.S. (1983) *Cell*, **34**, 807-814.
Noll,H. (1967) *Nature*, **215**, 360-363.
Norrander,J., Kempe,T. and Messing,J. (1983) *Gene*, **26**, 101-106.
Otvos,J.D. and Armitage,I.M. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7094-7098.
Pikielny,C.W., Teem,J.L. and Rosbash,M. (1983) *Cell*, **34**, 395-403.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211-214.
Queen,C.L. and Korn,J.L. (1980) *Methods Enzymol.*, **65**, 595-609.
Rodriguez,J.R., Pikielny,C.W. and Rosbash,M. (1984) *Cell*, **39**, 603-610.
Sanger,F., Nicklen,S. and Coulson,A.R. (1979) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5468.
Schechtman,M.G. and Yanofsky,C. (1983) *J. Mol. Appl. Genet.*, **2**, 83-99.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
Stuart,G.W., Searle,P.F., Chen,H.Y., Brinster,R.L. and Palmiter,R.D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7318-7322.
Weaver,R.F. and Weissmann,C. (1979) *Nucleic Acids Res.*, **7**, 1175-1193.
Woudt,L.P., Pastink,A., Kempers-Veenstra,A.E., Janson,A.E.M., Mager,W.H. and Planta,R.J. (1983) *Nucleic Acids Res.*, **11**, 5347-5360.
Zaret,K.S. and Sherman,F. (1982) *Cell*, **28**, 563-573.

Received on 8 July 1985