Volume 13 Number 21 1985 Nucleic Acids Research

Comparative nucleotide sequence analysis of two types of larval β-globin mRNAs of *Xenopus laevis*

Walter Knöchel, Wolfgang Meyerhof, Jürg Stalder[1] and Rudolf Weber[1]

Institut für Molekularbiologie und Biochemie der Freien Universität Berlin, Arnimallee 22, D-1000 Berlin 33, FRG, and [1]Zoologisches Institut der Universität Bern, Abteilung für Zell- und Entwicklungsbiologie, Baltzerstrasse 4, CH-3012 Bern, Switzerland

Received 30 July 1985; Revised and Accepted 1 October 1985

ABSTRACT

The complete nucleotide sequence of the cDNA insert of the clone pXGL25 derived from the larval β_{II}-globin mRNA of Xenopus laevis has been determined. The sequence of 593 nucleotides represents part of the 5'nontranslated region, the coding region for 146 amino acids and the entire 3'nontranslated region. It diverges from the related larval β_I-sequence by 24.9% in the coding region. Alignment of the 5' and 3'nontranslated regions of the two related larval β-sequences to maximum matching resulted in 31.2% and 46.7% divergence, respectively. Divergence between the corresponding adult and larval sequences considerably exceeds that of related larval sequences, suggesting that larval genes may have arisen by gene duplication prior to genome duplication. In contrast to mammalian β-globin mRNAs, replacement and silent base substitutions are equally abundant, thus indicating less functional constraint on the larval Xenopus laevis β-globin chains. The larval β_I- and β_{II}-globins diverge by 30.8% and show most variation in the α_1/β_2-chain interaction sites.

INTRODUCTION

The genes encoding the protein subunits of vertebrate hemoglobins provide attractive models for investigating developmental regulation and evolution of eukaryotic genes (1). In the anuran amphibian Xenopus laevis a single switch in hemoglobin phenotypes occurs at metamorphosis and is characterized by the complete replacement of the larval by adult globin chains (2, 3). cDNA clones derived from larval and adult erythroblasts revealed eight abundant sequences four being preferentially expressed in larval and four others in the adult stage, each group comprising two pairs of closely related α- and β-globin sequences respectively (4). Further studies disclosed a unique chromosomal organization of the Xenopus laevis globin genes. They are arranged in two clusters of related genes, each containing an adult (A) α-gene linked to an adult β-gene (5), which are flanked by two very closely related larval (L) α- and also by two larval β-genes in the order: 5'-α_a^L-α_a^L-α^A-β^A-β_a^L-β_b^L-3' (6). The existence of a second β^L-gene in each cluster is as yet based only on ge-

netic evidence, and needs confirmation. Therefore, it is not certain if the known β_I^L- and β_{II}^L-genes belong to the same or different pairs of related genes. However, the fact that larval a- and b-genes crosshybridize under stringent conditions indicates that they must be very closely related (6).

The two clusters of the Xenopus laevis globin gene family were taken as further evidence for the hypothesis that Xenopus laevis might have arisen from an ancestor by genome duplication or tetraploidization some 30 million years ago (7).

To elucidate the evolutionary history of the Xenopus laevis globin genes, and to assess the functional characteristics of the stage-specific globin chains, sequence analysis of cloned cDNAs was undertaken by several laboratories. The primary structure of the α_I^A- (8, 9) as well as the related β_I^A- (11) and β_{II}^A-globin chains (12) has been derived from the respective mRNAs. Information on the larval globin sequences is less complete comprising only data on the α_I^L- (9, 10) and β_I^L-mRNAs (13).

In this paper we present the complete nucleotide sequence of the Xenopus laevis β_{II}^L-globin mRNA and the corresponding amino acid sequence. These sequences are then compared to previously published data on the related β_I^L-globin mRNA (13, 14) and to all the other known globin mRNA sequences of Xenopus laevis. Moreover, the amino acid sequences of the β-globin chains are analyzed for specific functional characteristics in primary structure.

MATERIALS AND METHODS

Isolation of globin mRNA and cDNA cloning

9S poly(A)-containing RNA was isolated from immature red blood cells of anemic Xenopus larvae at stage 53-54 (15) and transcribed into cDNA. The cDNA was inserted into the Pst I site of pBR 322 using G/C-tailing. Construction and characterization of the recombinant plasmids have been reported previously (4, 12). The clone pVD12, representing the β_I^L-sequence was originally described in a study on the β_I^L-globin-gene (14). The clone pXGL25 was previously described by Widmer et al. (4) as larval β_1-sequence, but in view of the genomic arrangement of the corresponding gene was later designated as β_{II}^L-sequence (6).

Nucleotide sequence determination

For nucleotide sequence determination the cDNA inserts of the clones pVD12 and pXGL25 were digested with appropriate restriction enzymes as shown in

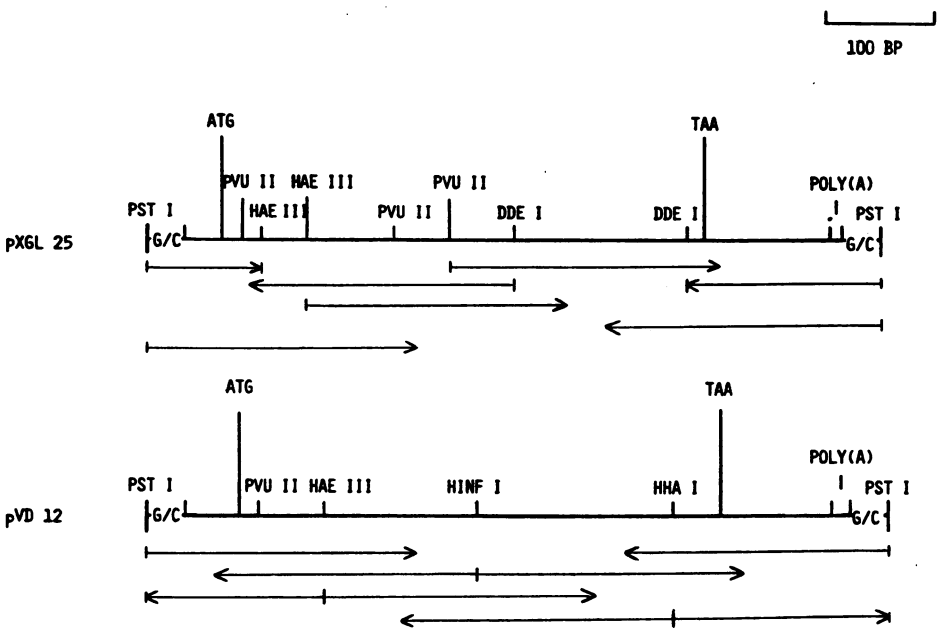

Figure 1: Restriction maps and restriction fragments of clones pXGL25 (β^L_{II}) and pVD12 (β^L_I) used for sequencing.
Arrows show direction and extent of nucleotide readings.

Fig. 1. Restriction fragments were separated by gel electrophoresis and either 5'-end labelled by polynucleotide kinase with (γ-^{32}P)ATP or 3'-end labelled by terminal transferase with (α-^{32}P)ddATP using commercial kits (Amersham, Buchler) and the protocols of the suppliers. After cleavage by a second restriction enzyme the fragments were separated by gel electrophoresis and submitted to the chemical sequencing reactions of Maxam and Gilbert (16). By two electrophoretic runs of each reacted fragment, it was possible to read sequences up to 250 nucleotides.

RESULTS AND DISCUSSION

The nucleotide sequences of the clones pVD12 and pXGL25 representing the β^L_I- and β^L_{II}-globin mRNA respectively, together with the derived amino acid sequences are shown in Fig. 2. Both cDNA inserts contain part of the 5'nontranslated region, viz., including the initiation codon ATG, 48 nucleotides in the β^L_I-sequence (pVD12), and 32 nucleotides in the β^L_{II}-sequence (pXGL25), the entire coding region for 146 amino acids, and the 3'nontranslated region comprising, the termination codon TAA included, 109 nucleotides in the β^L_I- and

```
A:                                   TACTTCTGCCATTTGAAGTCTTAGCAGCT ACTCCC  ATCTAC AGCCACC ATG
B:                                                ..T..TC..C.....TTA...A..A...A... ...

                                                  10                                      20
A:    GTG CAC CTG TCA GCT GAT GAG AAA TCC GCC ATT AAC GCG GTA TGG TCA AAG GTC AAC ATT
B:    ..C ..T TG. A.. ... ..A ... ..G G.. ... ..C .CT T.T ... ... CAG ... ... ..T G.A
A*:   Val His Leu Ser Ala Asp Glu Lys Ser Ala Ile Asn Ala Val Trp Ser Lys Val Asn Ile
B*:    .   .  Trp Thr  .  Glu  .   .  Ala  .   .  Thr Ser  .   .  Gln  .   .   .  Val

                                                  30                                      40
A:    GAA AAT GAT GGC CAT GAT GCC CTG ACC AGG CTG CTG GTT GTG TTT CCC TGG ACC CAG AGG
B:    ... C.. ... ... ... ... ... ... GG. ... ... ... A.. ... .AC ... ... ... ... ..A
A*:   Glu Asn Asp Gly His Asp Ala Leu Thr Arg Leu Leu Val Val Phe Pro Trp Thr Gln Arg
B*:    .  His  .   .   .   .   .   .  Gly  .   .   .  Ile  .  Tyr  .   .   .   .   .

                                                  50                                      60
A:    TAT TTC AGC AGC TTT GGA AAC CTA TCC AAT GTG GCT GCC ATC TCT GGA AAT GCC AAG GTT
B:    ..C ... ..T .A. ... ... ... ..C ... ... TCA ... ..T G.T G.. ... ... ... ... ...
A*:   Tyr Phe Ser Ser Phe Gly Asn Leu Ser Asn Val Ala Ala Ile Ser Gly Asn Ala Lys Val
B*:    .   .   .  Asn  .   .   .   .   .   .  Ser  .   .  Val Ala  .   .   .   .   .

                                                  70                                      80
A:    CGT GCC CAT GGC AAG AAG GTT TTG TCT GCC GTC GAT GAA TCC ATC CAC CAT TTA GAT GAT
B:    .AA ... ... ... ... ... ... C.T ..A ..T ..T .GC A.T G.. ..T AG. ... A.T ..C AG.
A*:   Arg Ala His Gly Lys Lys Val Leu Ser Ala Val Asp Glu Ser Ile His His Leu Asp Asp
B*:   Gln  .   .   .   .   .   .   .   .   .   .  Gly Asn Ala  .  Ser  .  Ile  .  Ser

                                                  90                                     100
A:    ATC AAG AAC TTC CTC TCT GTA CTG AGC ACG AAG CAC GCT GAG GAA CTC CAC GTG GAC CCT
B:    G.G ... TC. .CT ... CAA CA. ..C ..T .A. .TC ..T ..C ACT ... ..G TTT ... ... ...
A*:   Ile Lys Asn Phe Leu Ser Val Leu Ser Thr Lys His Ala Glu Glu Leu His Val Asp Pro
B*:   Val  .  Ser Ser  .  Gln Gln  .   .  Lys Ile  .   .  Thr  .   .  Phe  .   .   .

                                                 110                                     120
A:    GAA AAC TTC AAG CGT CTT GCT GAT GTG CTG GTG ATT GTC TTG GCT GGC AAA CTG GGA GCT
B:    ..G ... ..T ... ... T.. .G. .GA ..T ... ..C ... ... ... .G. .C. ... ... ... A..
A*:   Glu Asn Phe Lys Arg Leu Ala Asp Val Leu Val Ile Val Leu Ala Gly Lys Leu Gly Ala
B*:    .   .   .   .   .  Phe Gly Gly  .   .   .   .   .   .  Gly Ala  .   .   .  Thr

                                                 130                                     140
A:    GCC TTC ACT CCT CAA GTC CAG GCT GCC TGG GAG AAA TTC AGC GCT GGG CTG GTG GCT GCT
B:    ... ... ... ... A.. ..T ... ... ..T ... ... ... ... .TT ..A .TT T.. ..T .A. .G.
A*:   Ala Phe Thr Pro Gln Val Gln Ala Ala Trp Glu Lys Phe Ser Ala Gly Leu Val Ala Ala
B*:    .   .   .   .  Lys  .   .   .   .   .   .   .   .  Ile  .  Val  .   .  Asp Gly

A:    CTT AGC CAT GGC TAC TTC TAA A GAATTTCATTCCAT GCAAAGCC  CAAT   ATC TGCTCGCC TGCA
B:    ... ... ..G ... ..T AA. ... .T.TT...TTG.TTT.G......T.AG..C.GCC..AG...ATA.AA..AT
A*:   Leu Ser His Gly Tyr Phe
B*:    .   .  Gln  .   .  Asn

A:    ACAAAGACAACTCTCACAAT   AAGTGCAT TGCTT   GTTGTC AATAAA  GATTTCTTGTCAAACTGCT POLY(A)
B:    ..C.....C.TGGAT.....GTC.TC..T..A.....CAACAAC.GA......AA.CAC....A.G POLY(A)
```

$A = \beta_I^L$ (pVD12) $\quad A^* = \beta_I^L$ amino acids

$B = \beta_{II}^L$ (pXGL25) $\quad B^* = \beta_{II}^L$ amino acids

Figure 2: Nucleotide sequences of the cDNA clones pVP12 = β_I^L (A) and pXGL25 = β_{II}^L (B) as well as the derived amino acid sequences (A*, B*). The 5' and 3'nontranslated regions are aligned for maximum homology. Identical nucleotides and amino acids are indicated by•

120 nucleotides in the β^{L}_{II}-sequence. Since both cDNA sequences carry poly(A) stretches of 18 (pVD12) and 10 (pXGL25) adenine residues, respectively, we conclude that in both cases the entire 3'nontranslated regions and part of the poly(A) tails have been cloned. Both sequences contain the putative polyadenylation signal AATAAA (17, 18) at a distance from the polyadenylation site of 19 nucleotides in the β^{L}_{I}-sequence and 13 nucleotides in the β^{L}_{II}-sequence. An even larger difference in length of the 3'nontranslated region was reported for the α^{L}_{I}- and α^{L}_{II}-globin mRNAs of Xenopus laevis (19).

The nucleotide sequence of the cDNA in pVD12 is identical to the sequence of the β^{L}_{I}-gene reported earlier (14), and like the gene sequence differs in the same three positions from the β^{L}_{I}-sequence described by Banville et al. (13). We assume that these discrepancies reflect genetic polymorphism.

Divergence of the nucleotide sequences

After alignment to maximum homology the 5'nontranslated regions (without ATG) diverge by 31.2% (= 10 changes/32 nucleotides) and by 46.7% (=56/120) in the 3'nontranslated regions (including TAA). In the coding regions (without ATG) the divergence is only 24.9% (= 109/438). The overall divergence amounts to 29.5% which represents twice the divergence calculated from the melting temperature of homo- and heteroduplexes of the corresponding cDNA clones (4). The coding regions differ by 109 nucleotides, of which 46 (= 42.2%) represent silent substitutions. This is in striking contrast to the human and rabbit β-globin mRNA sequences, in which as much as 67% of the substitutions are silent (20). The greater proportion of replacement substitutions in the larval Xenopus laevis β-globin mRNAs suggests that the corresponding proteins are under minor functional constraint than those of mammals. Comparison of the coding regions of the related larval α-globin mRNAs (Table 1) revealed a very similar divergence of 24.1% again exceeding the value obtained from melting curves (4) by a factor of two. It is clear, therefore, that melting of hybrid molecules does not accurately reflect sequence relatedness. To obtain a more reliable estimate of the relatedness of the Xenopus laevis globin-mRNAs, we determined the divergence in the coding regions between pairs of related α- and β-sequences.

As shown by Table 1, the results are surprisingly concordant for each type of sequence. In agreement with earlier data from melting curves (4) we find less divergence between the related adult than between the related larval sequences of both α- and β-types. However, in both α- and β-sequences the diver-

Table 1: Divergence between the coding regions of Xenopus laevis globin mRNAs

Sequences	Divergence		References
	Substitutions/ sequence	%	
$\alpha^A_I - \alpha^A_{II}$	32/423	7.6	(9) (12)
$\beta^A_I - \beta^A_{II}$	40/372	10.8	(11) (12)
$\alpha^L_I - \alpha^L_{II}$	86/357	24.1	(10)* G. Micheli (unpublished)
$\beta^I_I - \beta^L_{II}$	109/438	24.9	Fig. 2
$\alpha^A_I - \alpha^L_I$	184/423	41.4	(9) (10)*
$\alpha^A_{II} - \alpha^L_{II}$	137/359	38.2	(12) G. Micheli (unpublished)
$\beta^A_I - \beta^L_I$	195/435	44.8	(11) Fig. 2
$\beta^A_{II} - \beta^L_{II}$	164/375	43.7	(12) Fig. 2

*) Banville et al. (10) discovered discrepancies between the nucleotide sequence of α^L_I-globin mRNA, reported by Andres et al. (9), and their own data derived from the same cDNA clone pXGL 19. Reexamination of the original sequencing protocols now revealed that the sequence reported by Andres et al. (9) refers to clone pXGL 21 which represents a very closely related α^L_I-globin mRNA sequence. Moreover, the nucleotide sequence originally derived from pXGL 19 by Andres et al. (unpublished) is identical to that reported by Banville et al. (10) for the same clone.

gence is much greater between the larval than between the adult sequences. This finding is in obvious conflict with the genome duplication hypothesis, which actually implies simultaneous duplication of all genes. Since the tetraploidization hypothesis of the Xenopus laevis genome is further supported by the occurrence of pairs of related vitellogenin (21) and albumin genes (22), we must assume that the difference in divergence between adult and larval globin sequences might reflect differential functional constraint on the gene products.

We find a very similar divergence between larval and adult sequences of the α- or the β-type, but it is greater than the divergence between both types of related larval sequences. This suggests that the larval genes may have arisen by gene duplication prior to tetraploidization, by which pairs of related genes have emerged.

Table 2: Codon usage in larval β-globin mRNAs of Xenopus laevis

		U	β^L_I	β^L_{II}		C	β^L_I	β^L_{II}		A	β^L_I	β^L_{II}		G	β^L_I	β^L_{II}	
U	Phe	UUU	2	4	Ser	UCU	3	2	Tyr	UAU	1	1	Cys	UGU	0	0	U
		UUC	6	3		UCC	3	2		UAC	1	2		UGC	0	0	C
	Leu	UUA	1	0		UCA	2	2	Term	UAA	1	1	Term	UGA	0	0	A
		UUG	2	2		UCG	0	0		UAG	0	0	Trp	UGG	3	4	G
C	Leu	CUU	2	2	Pro	CUU	2	2	His	CAU	4	6	Arg	CGU	2	1	U
		CUC	2	3		CCC	1	1		CAC	4	0		CGC	0	0	C
		CUA	1	0		CCA	0	0	Gln	CAA	1	3		CGA	0	0	A
		CUG	8	6		CCG	0	0		CAG	2	4		CGG	0	0	G
A	Ile	AAU	3	5	Thr	ACU	1	4	Asn	AAU	3	4	Ser	AGU	0	3	U
		AUC	3	2		ACC	2	1		AAC	5	4		AGC	5	2	C
		AUA	0	0		ACA	0	1	Lys	AAA	3	3	Arg	AGA	0	1	A
	Met	AUG	1	1		ACG	1	0		AAG	7	8		AGG	2	1	G
G	Val	GUU	3	8	Ala	CGU	10	7	Asp	GAU	7	3	Gly	GGU	0	3	U
		GUC	4	4		GCC	9	9		GAC	1	2		GGC	4	5	C
		GUA	2	2		GCA	0	1	Glu	GAA	4	3		GGA	3	4	A
		GUG	7	3		GCG	0	0		GAG	3	3		GGG	1	0	G

Strongly favoured codons common to both sequences are underlined

Primary structure of the larval β-globin chains

Larval and adult amphibian hemoglobins differ in their functional characteristics, such as binding of oxygen and allosteric molecules, as well as in the stability of their tertiary and quarternary structure (23). Having established the primary structure of the two related larval β-globins, it was of interest to search for molecular characteristics that might account for the functional peculiarities of these proteins.

Fig. 1 shows that the β^L_I- and the β^L_{II}-globins diverge by 45 in 146 amino acids (= 30.8%). Amino acid substitutions are most abundant between positions 70-94, which corresponds to the second half of exon 2 of all known β-globin genes. The divergence to the adult globin amounts to 54% (13) for β_I-globin and 47% for β_{II}-globin, respectively.

Table 2 reveals non-random selection of synonymous codons and sequence specific preferences in codon usage. In each β^L-sequence eight codons are preferentially used, but only four of them are favoured in both sequences. On the other hand, 17 codons are avoided in each sequence, the majority of which con-

Table 3: Divergence (%) in functional regions of vertebrate β-like globins related to the larval β-globins of Xenopus laevis

Region	Globin	Xenopus laevis				Rana catesbeiana	Chicken			Human		
		β^L_I	β^L_{II}	β^A_I	β^A_{II}	β^L	ε	ρ	β^A	ε	γ^G	β
	Ref.	Fig. 2	Fig. 2	(11)	(12)	(26)	(27)	(27)	(28)	(29)	(30)	(31)
Heme contact *	β^L_I	-	12.5	25	25	12.5	25	25	25	25	25	18.8
(16 residues)	β^L_{II}	12.5	-	18.8	18.8	18.8	31.3	31.3	31.3	31.3	31.3	25
α1/β1 contact *	β^L_I	-	37.5	56.3	50	50	43.8	43.8	37.5	31.3	31.3	37.5
(16 residues)	β^L_{II}	37.5	-	43.8	56.3	31.3	43.8	47.5	31.3	37.5	43.8	37.5
α1/β2 contact *	β^L_I	-	23.1	30.1	30.1	7.7	23.1	23.1	23.1	23.1	23.1	23.1
(13 residues)	β^L_{II}	23.1	-	46.2	30.1	23.1	38.5	38.5	38.5	38.5	38.5	30.1

* according to Fermi (25)

tains the dinucleotide CG or is ending in A, but of these only 12 codons are avoided in both the β^L_I- and the β^L_{II}-sequences. The marked difference in avoidance of codons among related sequences, which are simultaneously expressed within the same cell, i.e. the larval erythroblast, suggests that functional constraints on the globin chains might be more effective in determining codon usage than abundance of isoacceptor tRNAs, as has been suggested for mammalian cells (24).

To get information on the functional significance of divergences between β-like globin chains of vertebrates, we compared the residues assumed to be involved in heme contact as well as in α/β-chain interactions, as derived from human deoxyhaemoglobin (25). The results which refer to embryonic and adult β-like globins of amphibians and higher vertebrates are summarized in Table 3 and expressed as % divergence from the larval Xenopus laevis β-globins.

On the whole we note less divergence in the heme contact regions than in those involved in chain interactions. However, only 10 out of 16 residues are conserved, which is in obvious contrast to the α-like globin chains, where the heme contact sites are almost invariant (9). This data is consistent with an observation of Richardson et al. (32), indicating less conservation in β-globins. As for the interchain contacts, divergence is greater in the α_1/β_1- than in the α_1/β_2-contact sites, yet - except for Xenopus laevis - there are apparently no stage-specific differences in divergence, as has been reported for

α-like globins (9). Out of the 16 α_1/β_1-contact sites only two are conserved, whereas in the α_1/β_2-interaction sites as many as 6 out of 13 residues are invariant. So far neither stage nor species-specific patterns in amino acid substitution are apparent except for the replacement of β93 Ser and β146 His in the adult by Ala and Phe or Asn, respectively, in the larval β-globin which might be responsible for the absence of the alkaline Bohr effect and the Root effect in the larval hemoglobin (23).

ACKNOWLEDGEMENTS

The authors are grateful to Dr. G. Micheli for communication of unpublished data and to Mrs. S. Friedli for typing the manuscript.

This work was supported in part by the "Swiss National Science Foundation" (Project 3.324.82).

REFERENCES

1. Jeffreys, A.J. (1981) In: Genetic Engineering (R. Williamson, ed.), 2, 1-48.
2. Just, J.J., Schwager, J., Weber, R., Fey, H. and Pfister, H. (1980) Wilhelm Roux's Archives 188, 75-80.
3. Hosbach, H.A., Widmer, H.J., Andres, A.-C. and Weber, R. (1982) In: Embryonic Development, Part A: Genetic Aspects (M.M. Burger and R. Weber, eds.). Alan R. Liss, New York.
4. Widmer, H.J., Andres, A.-C., Niessing, J., Hosbach, H.A. and Weber, R. (1981) Dev. Biol. 88, 325-332.
5. Jeffreys, A.J., Wilson, V., Wood, D., Simons, J.P., Kay, R.M. and Williams, J.G. (1980) Cell 21, 555-564.
6. Hosbach, H.A., Wyler, T. and Weber, R. (1983) Cell 32, 45-53.
7. Bisbee, C.A., Baker, M.A., Wilson, A.C., Hadji-Azimi, I. and Fischberg, M. (1977) Science 195, 785-787.
8. Kay, R.M., Harris, R., Patient, R.K., and Williams, J.G. (1983) Nucleic Acids Res. 11, 1537-1542.
9. Andres, A.-C., Hosbach, H.A. and Weber, R. (1984) Biochim. Biophys. Acta 781, 294-301.
10. Banville, D. and Williams, J.G. (1985) Nucleic Acids Res. 13, 5407-5421.
11. Williams, J.G., Kay, R.M. and Patient, R.K. (1980) Nucleic Acids Res. 8, 4247-4258.
12. Knöchel, W., Meyerhof, W., Hummel, S. and Grundmann, U. (1983) Nucleic Acids Res. 11, 1543-1553.
13. Banville, D., Kay, R.M., Harris, R. and Williams, J.G. (1983) J. Biol. Chem. 258, 7924-7927.
14. Meyerhof, W., Klinger-Mitropoulos, S., Stalder, J., Weber, R. and Knöchel, W. (1984) Nucleic Acids Res. 12, 7705-7719.
15. Nieuwkoop, P.D. and Faber, J. (1956) Normal table of Xenopus laevis (Daudin) North-Holland, Amsterdam.
16. Maxam, A.M. and Gilbert, W. (1980) Meth. Enzym. 65, 499-560.
17. Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211-214.

18. Konkel, D.A., Tilghman, S.M. and Leder, P. (1978) Cell 15, 1125-1132.
19. Widmer, H.J., Hosbach, H.A. and Weber, R. (1983) Dev. Biol. 99, 50-60.
20. Kafatos, F.C., Efstratiadis, A., Forget, B.C. and Weissmann, S.M. (1977) Proc. Nat. Acad. Sci. USA 74, 5618-5622.
21. Wahli, W., Dawid, I., Ryffel, G.U. and Weber, R. (1981) Science 212, 298-304.
22. Westley, B., Wyler, T., Ryffel, G.U. and Weber, R. (1982) Nucleic Acids Res. 9, 3557-3574.
23. Perutz, M.F. (1983) Mol. Biol. Evol. 1, 1-28.
24. Grantham, R., Gauthier, C., Young, M., Mercier, R. and Pave, A. (1980) Nucleic Acids Res. 8, r49-r62.
25. Fermi, G. (1975) J. Mol. Biol. 97, 237-256.
26. Watt, K.W.K., Maruyama, T. and Riggs, A. (1980) J. Biol. Chem. 255, 3294-3301.
27. Dodgson, J.B., Stadt, S.J., Choi, O.R., Dolan, M., Fischer, H.D. and Engel, J.D. (1983) J. Biol. Chem. 258, 12685-12692.
28. Dolan, M., Dodgson, J.B. and Engel, J.D. (1983) J. Biol. Chem. 258, 3983-3990.
29. Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 621-626.
30. Slightom, J.L., Blechl, A.E. and Smithies, O. (1980) Cell 21, 627-638.
31. Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) Cell 21, 647-651.
32. Richardson, C., Capello, J., Cochran, M.D., Armentrout, R.W. and Brown, R.D. (1980) Dev. Biol. 78, 161-172.