Volume 11 Number 22 1983 Nucleic Acids Research

# Transcriptional activity of the human pseudogene $\psi\alpha$ globin compared with $\alpha$ globin, its functional gene counterpart

E.Whitelaw and N.J.Proudfoot

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received 27 September 1983; Accepted 28 October 1983

ABSTRACT

Transcriptional analysis of the human pseudogene $\psi\alpha$ globin has revealed the following features: (1) The promoter with a 23 bp deletion between the CCAAT and ATA boxes is functional both in vitro and in vivo, 3 fold and 10 fold less efficient, respectively, than $\alpha$. (2) Both the $\psi\alpha$ and $\alpha$ globin gene promoters are active in the absence of transcriptional enhancers, either a gene-encoded or viral enhancer. (3) The mutated poly(A) addition signal in $\psi\alpha$ (AATGAA) appears to be completely nonfunctional. This result provides an explanation for the absence of $\psi\alpha$ transcripts in human erythroid cells.

## INTRODUCTION

A considerable number of pseudogenes have been identified in eukaryotes with examples demonstrated in virtually all well characterised gene families; for example, in $\alpha$ and $\beta$ globins (1), immunoglobulin (2) and tubulin (3). A generally acceptable definition for a pseudogene would be a DNA sequence that bears significant sequence homology (detectable by cross-hybridisation) to a functional gene but has sequence differences within the coding region that would result in either a prematurely terminated polypeptide or a polypeptide that bears little or no homology to the functional gene product. This definition places emphasis on the translational competence of the pseudogene. However, many pseudogenes, in addition to translational termination and frameshift mutations have sequence changes in regions strongly implicated in transcriptional processes.

Human $\psi\alpha$ globin gene can be considered a good example of the above generalities. This pseudogene is about 80% homologous to $\alpha$ and is defined as a pseudogene by a short deletion in exon 2 that alters the translational reading frame of the gene resulting in a premature in-phase termination codon. Clearly the gene cannot encode an $\alpha$-like polypeptide chain (4). Interestingly, $\psi\alpha$ also has a number of transcriptional control mutations as

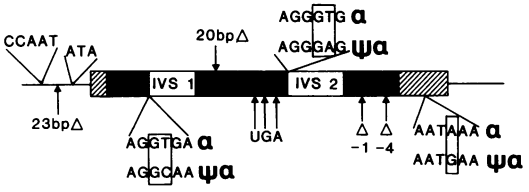



Figure 1. ψα gene diagram indicating sequence features that define ψα as a pseudogene. Hatched box indicates non-coding sequence; filled in box indicates exon; open box indicates intron; line indicates gene flanking sequence; Δ denotes deletion.

depicted in Figure 1.

Although the promoter sequences, the CCAAT and ATA boxes, which have been demonstrated to play critical roles in initiation of transcription (5,6) appear intact in ψα, a 23 bp deletion between them brings these sequences unusually close together, 17 bp rather than a more normal 40 bp apart. It has been demonstrated in globin genes that the spatial arrangement of the CCAAT and ATA boxes with respect to themselves and the mRNA cap site is highly conserved (7). We would therefore predict that this deletion might at least partially inactivate the ψα gene promoter. As indicated in Figure 1, neither introns of ψα are likely to be correctly spliced. Both intron donor sites are mutated so that the invariant GT sequence becomes GC in intron 1 and GA in intron 2. However, it is likely that alternative donor sites may be utilized (cryptic donors) as has been shown to occur in β-thalassaemic genes that lack normal splice sites (8). Finally, ψα has a variant poly(A) addition site sequence AATGAA rather than AATAAA. The sequence AATAAA has been found 15-20 bp 5' of the poly(A) tail of nearly all mammalian mRNAs (9) and has been directly implicated in mRNA 3' end formation in SV40 (10).

In these studies we have investigated the transcriptional activity of ψα as compared to α. Since we considered it unlikely that this pseudogene would function at detectable levels in its entirety, we analysed the promoter and poly(A) addition site of ψα separately by replacing either the 5' or 3' ends of the gene with an equivalent part of the human α gene. The results we obtain indicate that the ψα gene promoter is still functional, both *in vivo* and *in vitro*, but at reduced efficiency. We further demonstrate that both the ψα and α promoters are active without an enhancer element (11). However, the poly(A) addition site is virtually 100% non-functional and is therefore the likely cause of this pseudogene's inactivity *in vivo*.

METHODS

1. Template preparation

In vitro transcription experiments were carried out on DNA fragments isolated from a pBR322 subclone (pHB) containing the whole $\psi\alpha$ gene plus the 5' half of the $\alpha 2$ gene (12). pHB was digested with PvuII plus HindIII and the 1.6 Kb size fraction was purified by agarose gel fractionation, electroelution, and DEAE cellulose chromatography (13). This 1.6 Kb DNA contains 2 DNA fragments: one is $\psi\alpha$ 5' half and the other $\alpha 2$ 5' half. Both DNAs are therefore in identical yield and state of purity. DNA fragments containing either the 5' half or the entire $\psi\alpha$ and $\alpha 2$ genes (for $\psi\alpha$ a 1.6 Kb PvuII fragment and a 3.2 Kb SmaI-HpaI fragment; for $\alpha 2$ a 1.6 Kb PvuII-HindIII fragment and a 2.1 Kb PvuII fragment) were isolated from subclones containing these two genes and purified as before.

In vivo transcription experiments were carried out on various modifications of the human $\alpha$, $\psi\alpha$ and $\beta$ globin genes inserted in pSVOd (11, figure 2a). Figure 2b shows the two hybrid $\psi\alpha/\alpha$ and $\alpha/\psi\alpha$ constructs. $\psi\alpha/\alpha$ contains a 5' HinfI fragment of $\psi\alpha$ including the promoter region and part of the 5' non-coding sequence replacing the 250 bp SmaI fragment from $\alpha 1$ which includes the promoter through to the 5' end of intron 1. $\alpha/\psi\alpha$ contains the 3' portion of $\psi\alpha$ (BstEII-PvuII) including exon 3, the 3' non-coding region and 3' flanking sequence in place of the equivalent region of $\alpha 1$.

Figure 2c shows the various $\alpha/\beta$ constructs. Essentially, four different $\alpha 1$ gene fragments, the $\alpha$ promoter region plus exon 1 (fragment A), the rest of the $\alpha$ gene (fragment B), intron 1 and exon 2 (fragment C) and the 3' end of the $\alpha$ gene including exon 3 and the 3' non-coding region (fragment D) were inserted into the BamHI site (3' end of exon 2) of the $\beta$-gene in both orientations.

2. Transcriptional analysis

a. In vitro transcription The two 1.6 Kb DNA fragments (together and separately) containing the $\psi\alpha$ and $\alpha 2$ 5' halves as well as the separate full length gene fragments of $\psi\alpha$ and $\alpha$ were added to HeLa cell, in vitro transcription extracts, together with nucleoside triphosphates, one $\alpha$-$^{32}$P labelled (14). Labelled RNA run off transcripts were fractionated on methylmercury agarose gels (15). The gels were dried down and radioautographed.

b. Transient expression The different pSVOd globin gene constructs (Figure 2) were transformed into monkey Cos7 cells grown in culture. This cell line is a defective SV40 transformed CV1 monkey cell line that expresses

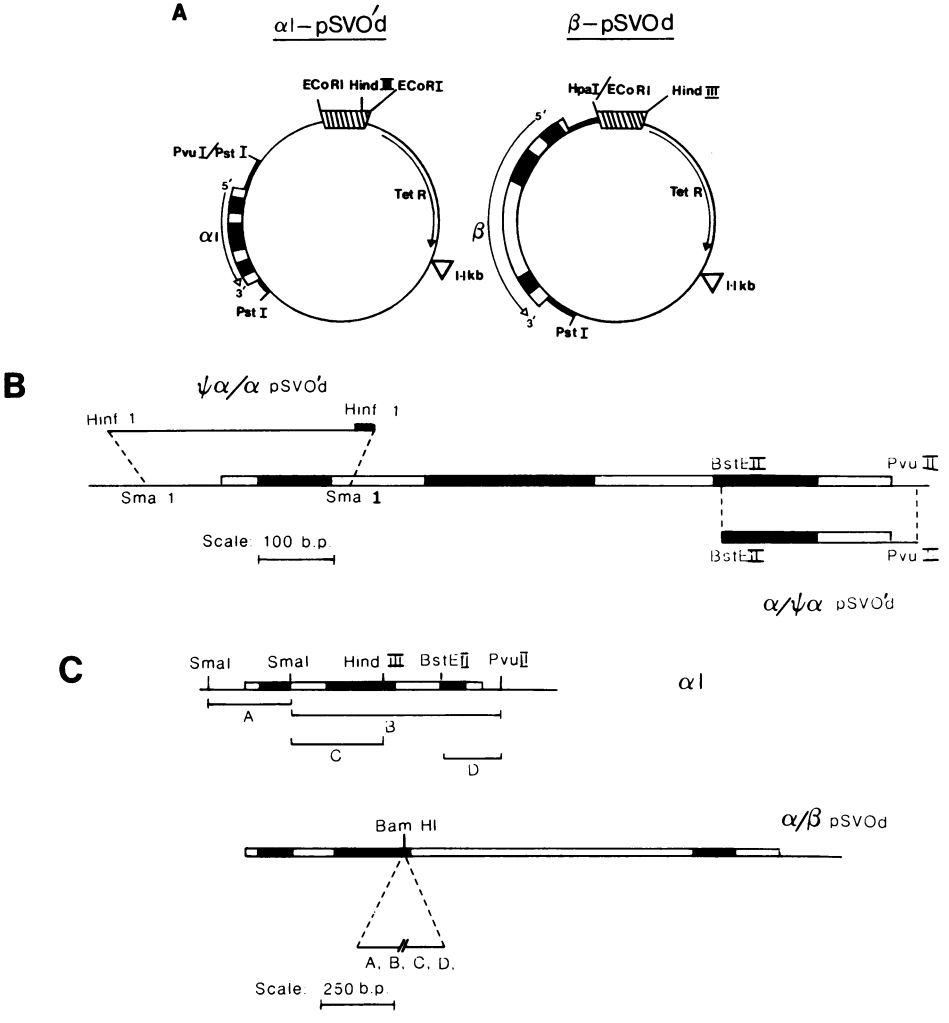



Figure 2.
A. Maps of αl-pSVO'd and β-pSVOd. Thin line indicates pBR322 sequence; thick line indicates α or β gene flanking sequence; open box indicates gene non coding sequence or intron; filled in box indicates exons; hatched box indicates SV40 origin sequence. For pSVO'd this is the EcoRII G origin fragment of SV40 inserted into pBR322 with EcoRI linkers. For pSVOd the EcoRII G fragment is cut down to the HindIII site indicated in pSVOd. This origin fragment contains less of the SV40 early gene but still has an intact origin (11). Δ indicates deletion of 1.1 Kb.
B.C. Restriction maps of ψα/α , α/ψα and α/β pSVOd.

sufficient levels of SV40 T antigen to allow replication of plasmids containing SV40 origin sequences, such as pSVOd (16). DNAs were precipitated with calcium phosphate and added to subconfluent Cos7 cells. After 12 hours the medium was changed and the cells were incubated a further 48 hours. Cells were then harvested and nuclear or cytoplasmic RNAs were purified as previously described (8). Essentially cells were gently lysed and nuclei spun through a sucrose cushion. Nuclear and cytoplasmic fractions were then incubated with proteinase K, phenol extracted and ethanol precipitated. Finally, nuclear RNA was DNase 1 treated to remove chromosomal and plasmid DNAs.

3. RNA mapping

a. Primer extension Both in vitro transcripts and in vivo, transient expression transcripts were analysed by primer extension analysis (13). For $\psi\alpha$ in vitro transcripts, a 3' end labelled, single-strand probe from exon 1 was purified as indicated in Figure 3c. For $\psi\alpha/\alpha$ in vivo transcripts, an AvaII fragment from $\alpha$1 intron 1 was 3' end labelled and strand separated as shown in Figure 4c. Finally for $\alpha/\beta$ in vivo transcripts an $\alpha$ exon 1 HinfI-HaeIII DNA fragment and a $\beta$ exon 1 intron1 HinfI DNA fragment was purified, 3' end labelled and strand separated as shown in Figure 6. These end labelled primers were annealed to in vitro or in vivo RNAs and extended using reverse transcriptase (17). The products of these reactions were fractionated by acrylamide gel electrophoresis.

b. S1 mapping experiments The 5' end of $\psi\alpha/\alpha$ transcripts was mapped using a kinase labelled agarose gel purified HindIII fragment containing the 5' half of $\psi\alpha/\alpha$ through to the HindIII site in exon 2 of $\psi\alpha/\alpha$ pSVOd (see Figure 4b). The 5' ends of the different $\alpha/\beta$ constructs were mapped using a kinase labelled strand separated HaeIII/HinfI fragment containing the promoter and 5' end of the $\beta$ gene. The 3' ends of $\psi\alpha/\alpha$, $\alpha/\psi\alpha$ or $\alpha$ pSVOd was mapped using BstEII digests of these different plasmids. In each case the enzyme cuts the different plasmids once, creating a linear molecule. These DNAs were then filled in using $^{32}$P dNTP and the Klenow subfragment of E.coli DNA polymerase 1 and the 3' end labelled DNAs were purified by phenol extraction and G100 Sephadex gel filtration (18). These different $^{32}$P end labelled DNA probes were annealed to the various transient expression RNAs. Annealing reactions were carried out under R loop conditions (80% formamide at 52°C). Annealed nucleic acids were then digested with S1 nuclease and the digestion products fractionated on polyacrylamide gels (19).

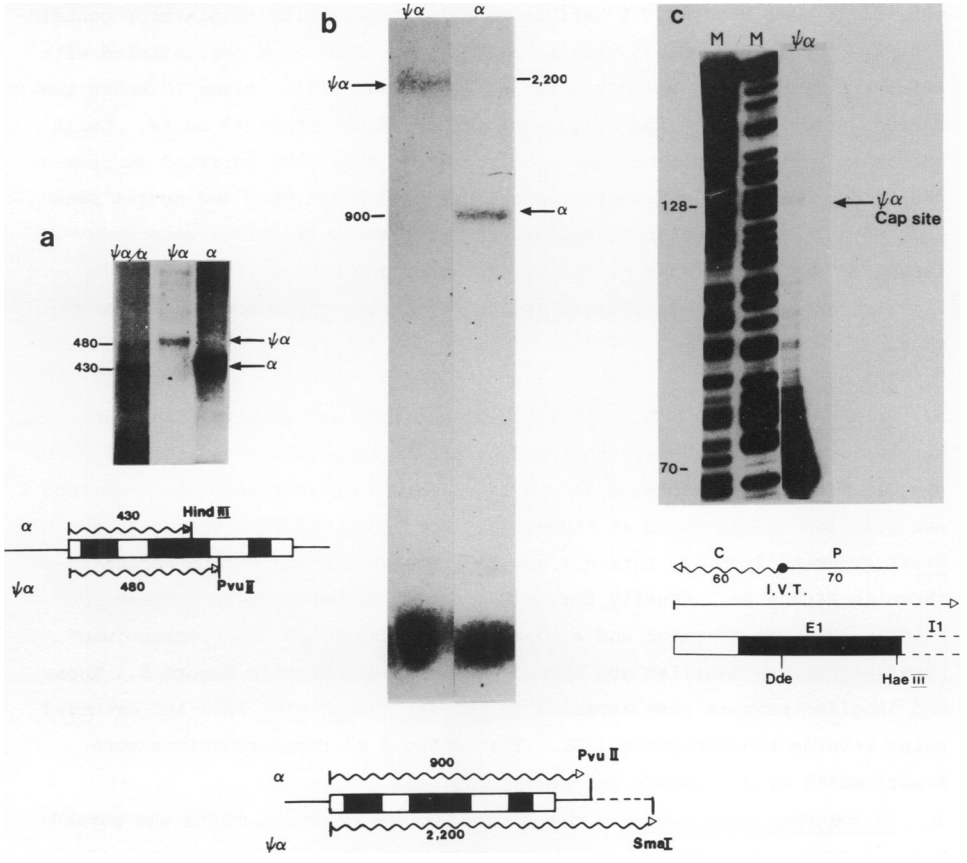



Figure 3.
a.b. Run off in vitro transcripts of α and ψα as indicated below radioautograph and in Methods.
c. Primer extension 5' end analysis of ψα in vitro transcripts. The strong band at 70 bp is excess primer whilst the faint band at 128 bp is the primer extension product. C denotes cDNA transcript; P denotes primer; I.V.T. denotes in vitro transcript. Lanes M are size markers (sequence ladders).

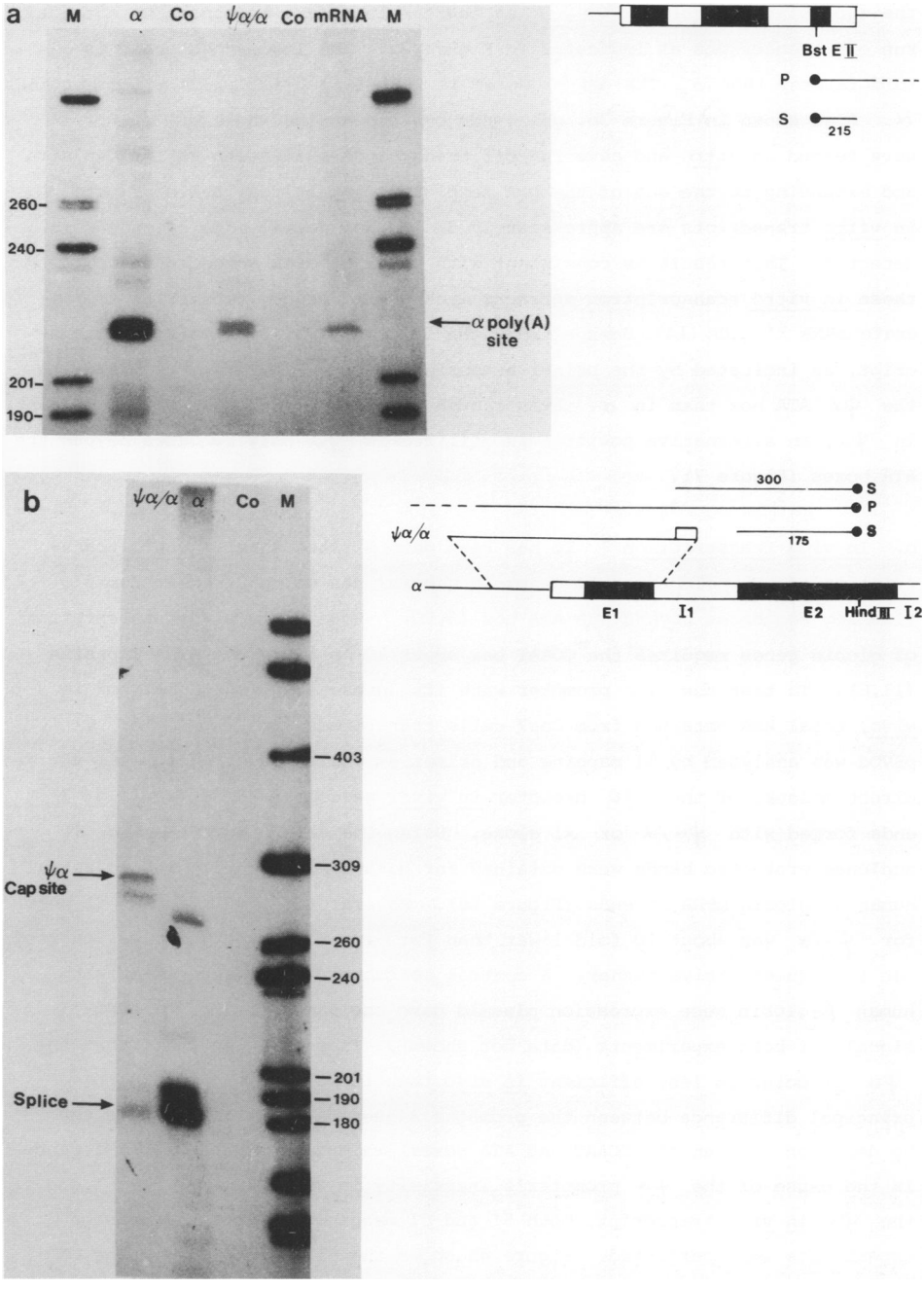
## RESULTS

### Analysis of the ψα globin gene promoter

a. In vitro transcription To assess the activity of the ψα promoter, we carried out transcription experiments both in vitro and in vivo. Previous preliminary data indicates that ψα can be transcribed in vitro (13). However, we decided to confirm this result and to obtain a more quantitative measure of the ψα promoter's efficiency as compared to α. Furthermore, we wished to define the precise 5' end of the ψα in vitro transcript. Using

the equimolar mixture of ψα , α as the template (see Methods), we obtained run off transcripts as indicated in Figure 3a. The larger ψα band is 3-fold fainter than α. The ψα promoter is therefore 3-fold less efficient than α. Also shown in Figure 3b, gel bands containing the whole of ψα or α were tested *in vitro* and gave run off transcripts initiating at the Cap site and extending to the end of the DNA templates. Again the levels of α to ψα *in vitro* transcripts are approximately 3 : 1. No normal mRNA 3' ends were detected. This result is consistent with other studies which indicate that these *in vitro* transcription extracts lack the necessary activities to generate mRNA 3' ends (14). Unexpectedly the 5' end of the ψα *in vitro* transcript, as indicated by the primer extension band in Figure 3c, is closer to the ψα ATA box than in α. Even though the α Cap site sequence is present in ψα, an alternative position is utilized in ψα only 20 bases beyond the ATA boxes (Figure 7).

b. *In vivo transcription* It has been demonstrated that *in vitro* transcription of RNA polymerase II genes in general has no requirement for 5' flanking sequences beyond the ATA box (5,20). However *in vivo* transcription of globin genes requires the CCAAT box about 70 bp 5' to the mRNA Cap site (11,6). To test the ψα promoter with its unusually close CCAAT box *in vivo*, total RNA obtained from Cos7 cells transformed with ψα/α or α1 pSVOd was analysed by S1 mapping and primer extension experiments. As a direct measure of the ψα promoter *in vivo*, we compared the levels of 3' ends formed with ψα/α or α1 alone. Using the same α BstEII probe, S1 nuclease protected bands were obtained for both DNAs corresponding to normal human α-globin mRNA 3' ends (Figure 4a). However the level of mRNA 3' ends for ψα/α was about 10 fold lower than for α. This experiment was carried out in a quantitative manner. A control cotransformation experiment with human β-globin gene expression plasmid gave the same level of β mRNA signal for both experiments (data not shown). These data indicate that the ψα promoter is less efficient *in vivo* than it is *in vitro*. Since the principal difference between the promoter sequences of ψα and α is the 23 bp deletion between the CCAAT and ATA boxes, we predict that this difference is the cause of the ψα promoter's inactivity *in vivo*. To map the 5' end of the ψα *in vivo* transcript, both 5' and S1 mapping and primer extension experiments were performed. Figure 4b shows the S1 mapping data. The DNA probe 5' end labelled at the HindIII site in exon 2 of ψα/α gives a strong signal corresponding to the 3' junction of intron 1 when hybridised to α1


a
M
α
Co
ψα/α
Co
mRNA
M
260–
240–
201–
190–
α poly(A) site
Bst E II
P
S
215
b
ψα/α
α
Co
M
403
309
260
240
201
190
180
ψα Cap site
Splice
300
S
P
ψα/α
175
S
α
E1
I1
E2
Hind III
I2

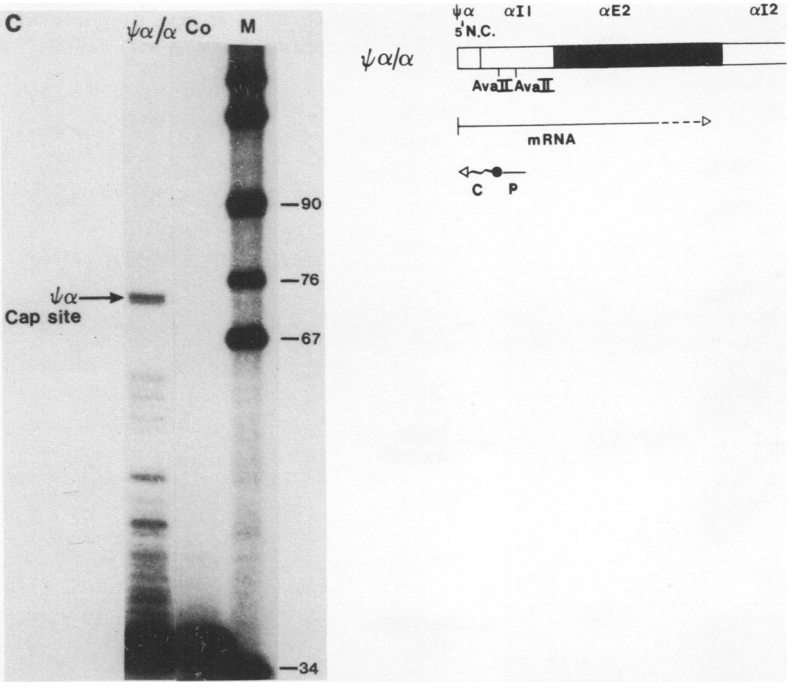



Figure 4.
a. 3' end S1 nuclease analysis of ψα/α versus α pSVOd. M denotes restriction fragment markers, α denotes α pSVOd. Co denotes minus RNA control, ψα/α denotes ψα/α pSVOd and mRNA denotes α-globin mRNA control. P denotes probe and S signal.
b. 5' end S1 nuclease analysis of ψα/α versus α pSVOd.
c. Primer extension 5' end analysis of ψα/α .

pSVOd transformed Cos7 RNA. However, ψα/α lacks the intron 1 donor site (see Figure 2b). Intron 1 of ψα/α is therefore only partially spliced as indicated by a faint band corresponding to intron 1 3' end (Figure 4b). Presumably an alternate cryptic donor site is utilised at low efficiency in ψα/α . Two larger bands are evident in the ψα/α experiment. The smaller band may be an over digestion product while the larger band corresponds to the 5' end of the ψα/α transcript. This DNA was too large to size precisely, so we used the alternate technique of primer extension to obtain a precise position for the ψα Cap site *in vivo*. Using a short single strand AvaII fragment in intron 1 of ψα/α , an extension product was obtained that precisely maps the ψα/α mRNA 5' end (Figure 4c). This places the 5' end of the ψα *in vivo* transcript at the α Cap site position, a different position to the ψα *in vitro* transcript (Figure 7).

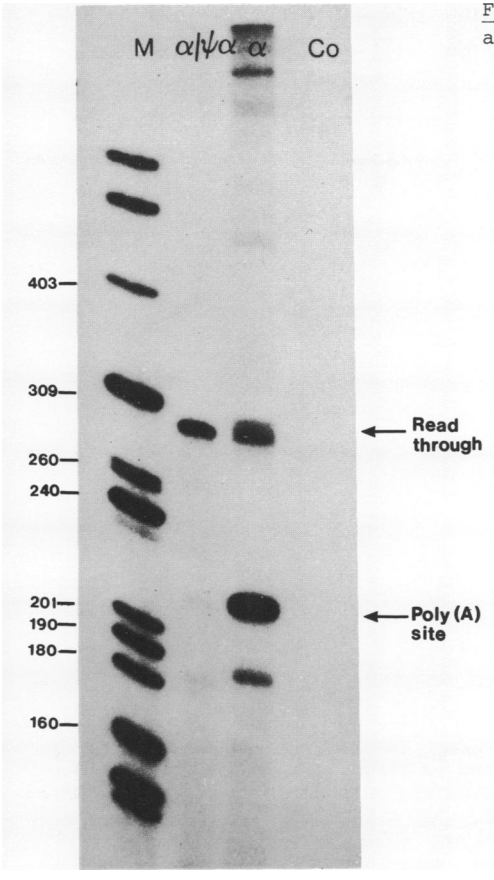
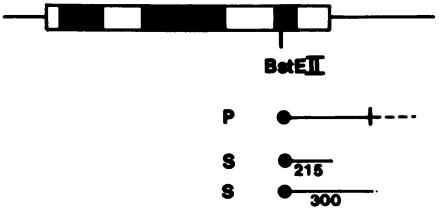



Figure 5. 3' end S1 nuclease analysis of α/ψα versus α pSVOd.




Analysis of ψα poly(A) addition site

To test the ability of ψα to generate mRNA 3' ends, we carried out in vivo transcription experiments, since in vitro transcription extracts lack the necessary activities to create mRNA 3' ends (Figure 3b). Furthermore, we

fused the more efficient α promoter onto the ψα 3' portion (α/ψα pSVOd) to obtain higher signals in the RNA mapping experiments (see Figure 2). Figure 5 demonstrates that whereas α gives the expected 3' end signal corresponding to α globin mRNA, α/ψα gives no detectable band ≃ 20 bases beyond the abherrant poly(A) addition site, but rather a strong read through band caused by mismatch between the probe and α/ψα transcript about 100 bases further 3'. These data demonstrate that the ψα poly(A) addition site is virtually non-functional. Transcripts read through this site and end at unmapped positions within the pSVOd vector sequence presumably at cryptic poly(A) addition sites. It seems very likely that the variant AATAAA sequence AATGAA is the cause of this effect. Figure 5 also demonstrates a read through band for α1 pSVOd at about 10% the level of the poly(A) addition site band. This indicates that the α poly(A) addition site is 90% efficient.

Analysis of α/β hybrid globin genes

We have demonstrated that the ψα globin gene promoter functions approximately 10 fold less well than α in Cos7 cells. Furthermore, like α, the ψα gene is functional in the absence of a viral transcriptional enhancer sequence. This result is in contrast to many other eukaryotic genes such as rabbit or human β-globin (21) or sea urchin histone genes (22) which do require viral enhancer sequences. However, certain eukaryotic genes, in particular immunoglobulin genes have recently been shown to contain their own enhancer sequences (23-25). We therefore wished to establish whether or not ψα and α possessed enhancer sequences within the gene as has been shown for immunoglobulin genes. To address this question for α-globin, we inserted sections of the α-gene into the human β gene in pSVOd. Since pSVOd lacks the SV40 enhancer sequence (11), the β-gene does not work at significant levels in this vector. If the α globin gene contains an enhancer sequence it seemed plausible that certain parts of α globin, when placed close to the β promoter, might allow β to transcribe without an SV40 enhancer. In other words, α might enhance β.

As shown in Figure 2c, various portions of α were inserted into the β-gene in both orientations. However, in no case was β significantly expressed. Figure 6a shows the level of β promoter activity in the various α/β pSVOd constructs using 5' S1 analysis. As a positive control for β-gene promoter activity (lane βE), an SV40-pBR322 plasmid containing a large portion of SV40 including the enhancer, origin and T antigen gene was transformed into HeLa cells (6). This expression system allows replication off the

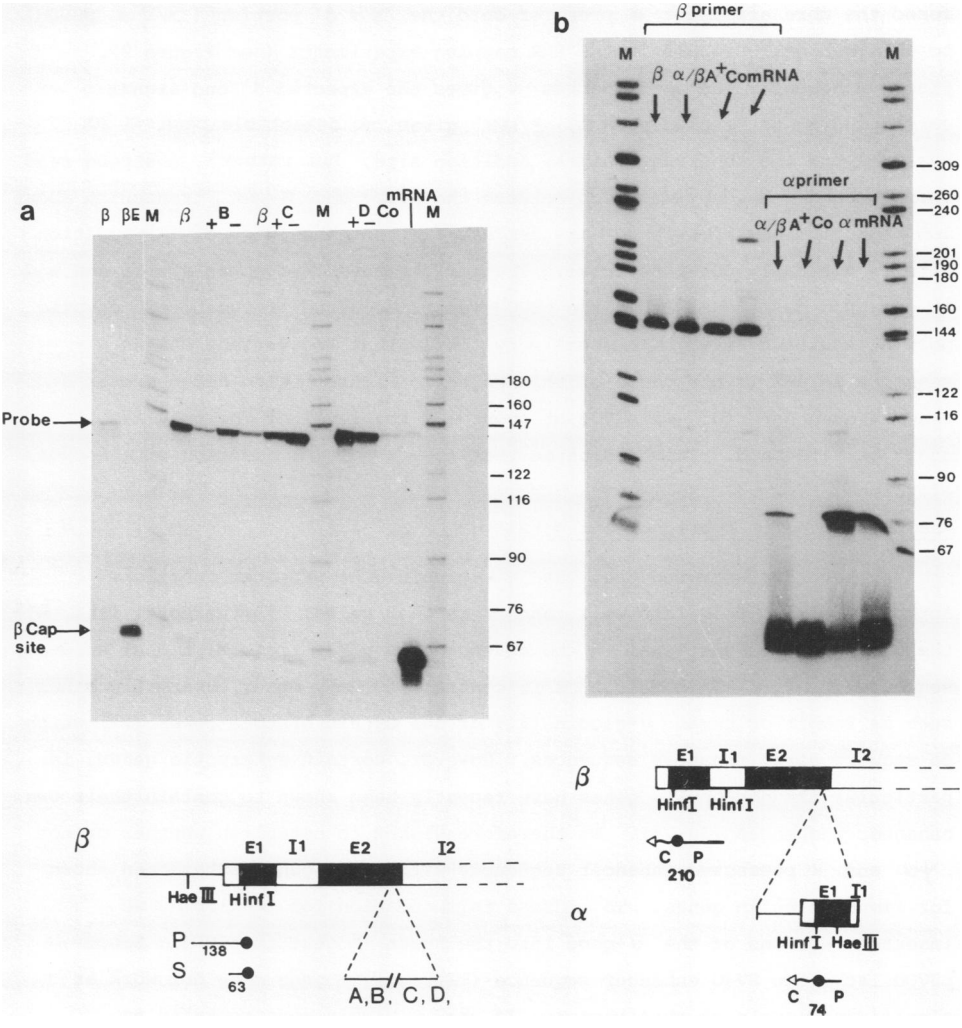



Figure 6.
a. 5' end S1 nuclease analysis of the various α/β pSVOd constructs. β denotes β pSVOd, βE denotes β transcription with SV40 enhancer (see text). +- refer to the different orientations of the α gene fragments, A, B, C and D in β pSVOd, + denotes same orientation as the β gene, - denotes opposite orientation.
b. Primer extension 5' end analysis of α/β $A^+$ pSVOd versus β pSVOd. Both the α and β promoters in this construct were analysed.

plasmid in an analogous way to pSVOd plasmids in Cos7 cells (11). As indicated, none of the α/β constructs allows significant transcription of the β promoter as compared to the control. It therefore would appear that α1

does not possess intrinsic enhancer activity. Figure 6b shows the promoter analysis (using primer extension) of $\alpha$1A+/$\beta$ pSVOd. This $\alpha/\beta$ construct contains the promoter and exon 1 of $\alpha$1 inserted into the BamHl site of the $\beta$-gene (Figure 2c). Whereas the $\beta$-promoter is still inactive in this construct, the $\alpha$ promoter, although within the $\beta$ gene, displays significant activity. Presumably the $\alpha$ promoter's low level of activity as compared to the intact $\alpha$ gene in $\alpha$1 pSVOd reflects its unphysiological situation in the middle of another gene.

From the data presented in Figure 6, we conclude that the $\alpha$ gene does not contain enhancer-like activity since it cannot activate the $\beta$ promoter but rather the $\alpha$ promoter region is capable of transcription in the absence of an enhancer. A similar argument can be made for $\psi\alpha$. As shown in Figure 4, $\psi\alpha/\alpha$ pSVOd initiates transcription, albeit at low levels. This construct contains the promoter sequence of $\psi\alpha$ through to the middle of the 5' non-coding sequence. Since there are no enhancer elements in this plasmid, we conclude that, like $\alpha$, the $\psi\alpha$ promoter functions independently of transcriptional enhancers. It should be mentioned that we cannot discount the possibility that $\psi\alpha$ might possess an enhancer sequence in the 5' flanking region. The fact that $\psi\alpha$ and $\alpha$ apparently function without an enhancer argues that there may be features in common between the $\alpha$ and $\psi\alpha$ promoters that allow for enhancer independent transcription.

## DISCUSSION

### The $\psi\alpha$ promoter

Figure 7 compares the DNA sequences of the $\psi\alpha$ and $\alpha$ promoters. As discussed above, the most striking difference between these two promoters is the 23 bp deletion between the CCAAT and ATA boxes in $\psi\alpha$. Another feature of the $\psi\alpha$ promoter is that the DNA sequence between the ATA box and Cap site is almost completely diverged from $\alpha$. In particular it should be noted that two additional potential Cap sites have been created in the $\psi\alpha$ sequence (denoted by * in Figure 7). *In vitro* transcription is generally found to be insensitive to promoter sequence elements 5' to the ATA box (5,20). Thus the efficiency of initiation of $\psi\alpha$ *in vitro* transcription is only partly reduced as compared to $\alpha$, about threefold down. However, the $\psi\alpha$ *in vitro* transcript initiates 5' to the $\alpha$ position, only 20 bp away from the ATA box at one of the two alternative Cap site sequences mentioned above, even though this brings the *in vitro* Cap site unusually close to the ATA box. *In vivo* transcriptional initiation reveals a quite different picture. The level

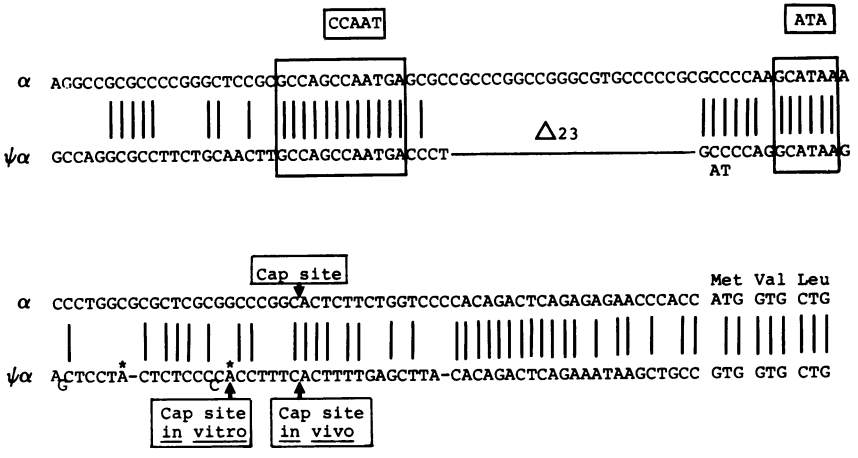



Figure 7. Sequence comparison of the α and ψα promoters. Boxed sequences indicate regions of extended sequence homology around the CCAAT and ATA boxes of ψα and α.

of initiation is at least 10 fold down as compared to α. Thus the deletion between CCAAT and ATA appears to be critical. Furthermore the site of initiation corresponds to the α Cap site. The previously demonstrated role of the ATA box in positioning the Cap site of the mRNA (6) appears to function correctly for the ψα gene in vivo but not in vitro. In other words, the sequence of the Cap site appears to be more critical than the position of the ATA box for in vitro transcription.

Another interesting feature of both the ψα and α gene promoters as described above is the fact that both function in the absence of transcriptional enhancers either their own enhancer (23) or a viral enhancer (21). We have demonstrated that α globin is incapable of enhancing β-globin even when different portions of α are inserted into β, close to the β-globin gene promoter. This result is in agreement with other studies (26,27) in which separate α and β genes were placed in the same plasmid, transformed into tissue culture cells and α and β promoter activity monitored. In both studies, α but not β transcription was observed. However, it could be argued in these studies that the putative α enhancer was too far away from β to enhance the β promoter. Such an argument is backed up by the finding that a mouse immunoglobulin lambda gene promoter requires the SV40 enhancer close to it, within several hundred bp, for correct initiation of transcription (28). Our studies confirm that α does not enhance β, since fragments of the gene have been placed in close proximity to the β-promoter

(see Figure 2c) and still fail to enhance it.

The hypothesis that $\alpha$ and $\psi\alpha$ globin genes might contain their own enhancer sequences is therefore discounted. Rather there must be intrinsic features of these two promoters that allow for enhancer independent transcription. Indeed it is plausible to suggest that one reason for the $\psi\alpha$ promoter's reduced efficiency as compared to $\alpha$ is that it has lost some of the $\alpha$ promoter's sequence features necessary for enhancer independent transcription.

The $\psi\alpha$ poly(A) addition site

We have demonstrated that the $\psi\alpha$ poly(A) addition site is non-functional. Thus over 99% of $\psi\alpha$ transcripts read through the normal end of the mRNA. Although there are a number of different base changes between $\psi\alpha$ and $\alpha$ in their 3' non-coding sequences, we presume that the critical mutation is within the highly conserved poly(A) addition signal AATAAA which becomes AATGAA in $\psi\alpha$. Functional variants of AATAAA have been demonstrated in a few eukaryotic genes, in particular AGTAAA (29,30) and ATTAAA (31-34). However, two non-functional variants of AATAAA have recently been described. The sequence AATAAG has been identified in an $\alpha$ thalassaemia gene and shown to be the cause of this gene's malfunction (35). Similarly the AATAAA sequence of the adenovirus E1A gene has been mutated to AACAAA by site directed mutagenesis. Again this mutant gene fails to generate normal 3' ends with consequent transcriptional read through into the adjacent gene E1B (36). It seems highly probable that in $\psi\alpha$, the cause of the gene's observed transcriptional read through is the mutated poly(A) addition signal AATGAA. Apparently variations in the AATAAA sequence at the second position are permissible whilst variations in the third, fourth and sixth positions are not. The precise sequence rules governing the activity of the poly(A) addition signal are still not fully understood and will only be resolved when a complete set of AATAAA variants are constructed and tested.

$\psi\alpha$ Transcription in humans?

Since we have established that the $\psi\alpha$ gene functions, albeit at low levels, following transformation in Cos7 cells, we decided to search for $\psi\alpha$ transcripts in human erythroid cells. Embryonic, foetal and adult erythroid total cell RNA was hybridised to a $\psi\alpha$ 3' probe and S1 treated. No bands corresponding to $\psi\alpha$ 3' ends (neither the read through mismatch band or poly(A) site band) were detectable even after long exposures (data not shown). We conclude that $\psi\alpha$ gene transcripts are not present at detectable levels in human erythroid cells at any developmental stage.

We have demonstrated that the $\psi\alpha$ promoter is active in Cos7 cell transient expression systems. However it should be emphasised that such systems are certainly non-physiological so that even though the $\psi\alpha$ promoter may function in transient expression, it may be completely switched off in humans. Nevertheless the poly(A) addition sequence is virtually non-functional. Transcripts that read through the normal mRNA 3' end are likely to be unstable. In the study of the $\alpha$-thalassaemia gene with an AATAAA mutation (35), it was found that read through transcripts were detected in RNA from HeLa cells transformed with the $\alpha$-thalassaemia gene, but were not detected in the patients' reticulocytes. The authors account for this difference by arguing that the $\alpha$-thalassaemia "read through" mRNA is relatively unstable, so that it is absent from the relatively old globin mRNA population found in reticulocytes. Therefore, even though the $\psi\alpha$ promoter may be active, we would predict that the poly(A) site mutation is the principal cause of the absence of $\psi\alpha$ transcripts in humans.

It has previously been suggested (1) that pseudogenes, following their initial translational inactivation, undergo a strong selection process to turn off gene transcription. Such a process would be advantageous to the cell as a whole, since the presence of non-functional mRNA might impede translation of other functional mRNAs. This suggestion is clearly confirmed for $\psi\alpha$ since we have demonstrated that both the initiation of transcription and mRNA 3' end formation in this gene are ineffective processes.

ACKNOWLEDGEMENTS

We gratefully acknowledge Tom Maniatis in whose laboratory some of the in vitro transcription experiments described here were carried out. We also acknowledge helpful discussions with Richard Triesman and Tom Maniatis and the availability of their unpublished work. This work was funded by a British Medical Research Council research project grant (no. G8108316CB).

REFERENCES

1. Proudfoot, N.J., Nature 286, 840-841 (1980)
2. Bentley, D.L. and Rabbitts, T.H., Nature 288, 730-733 (1980)
3. Lee, M.G-S., Lewis, S.A., Wilde, C.D. and Cowan, N.J., Cell 33, 477-487 (1983)
4. Proudfoot, N.J. and Maniatis, T., Cell 21, 537-544 (1980)
5. Grosveld, G.C., Shewmaker, C.K., Jat, P. and Flavell, R.A., Cell 25, 215-226 (1981)
6. Grosveld, G.C., de Boer, E., Shewmaker, C.K. and Flavell, R.A., Nature 295, 120-126 (1982)

7. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C, Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J., Cell 21, 653-668 (1980)
8. Triesman, R., Proudfoot, N.J., Shander, M. and Maniatis, T., Cell 29, 903-911 (1982)
9. Proudfoot, N.J. and Brownlee, G.G., Nature 263, 211-214 (1976)
10. Fitzgerald, M. and Shenk, T., Cell 24, 251-260 (1981)
11. Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T., Cell 27, 279-288 (1981)
12. Lauer, J., Shen, C-K.J. and Maniatis, T., Cell 20, 119-130 (1980)
13. Proudfoot, N.J., Shander, M.H.M., Manley, J.L., Gefter, M.L. and Maniatis, T., Science 209, 1329-1336 (1980)
14. Manley, J.L., Fire, A., Cano, A., Sharp, P.A. and Gefter, M.L., Proc. Nat.Acad.Sci.USA 77, 3855-3859 (1980)
15. Bailey, J.M. and Davidson, N., Anal.Biochem. 70, 75- (1976)
16. Gluzman, Y., Cell 23, 175-182 (1981)
17. Devine, J.M., Tsang, A.S. and Williams, J.G., Cell 28, 793-800 (1982)
18. Maniatis, T., Fritsch, E.F. and Sambrook, J., Molecular Cloning, a laboratory manual. Cold Spring Harbor Press (1982)
19. Weaver, R.F. and Weissman, C., Nucl.Acids Res. 6, 1175-1193 (1979)
20. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P., Science 209, 1406-1414 (1980)
21. Banerji, J., Rusconi, S. and Schaffner, W., Cell 27, 299-308 (1981)
22. Bendig, M.M. and Hentschel, C.C., Nucl.Acids Res. 11, 2337-2346 (1983)
23. Banerji, J., Olson, L. and Schaffner, W., Cell 33, 729-740 (1983)
24. Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S., Cell 33, 717-728 (1983)
25. Queen, C. and Baltimore, D., Cell 33, 741-748 (1983)
26. Treisman, R., Green, M.R. and Maniatis, T., Proc.Natl.Acad.Sci.USA, in press
27. Humphries, R.K., Ley, T., Turner, P., Moulton, A.D. and Nienhuis, A.W., Cell 30, 173-183 (1982)
28. Picard, D. and Schaffner, W., Proc.Natl.Acad.Sci.USA 80, 417-421 (1983)
29. Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. and Goeddel, D.V., Nature 302, 33-34 (1983)
30. Donehower, L., Huang, A. and Hager, G., J.Virol. 37, 226-238 (1981)
31. Hagenbuchle, O., Krikeles, M.S., Sprague, K.V., J.Biol.Chem. 254, 7157-7169 (1980)
32. Jung, A., Sippel, A.E., Grez, M. and Schutz, G., Proc.Natl.Acad.Sci.USA 77, 5759-5763 (1980)
33. Fornwald, J., Kuncio, G., Peng, I. and Ordahl, C.P., Nucl.Acids Res. 10, 3861-3876 (1982)
34. Zakut, R., Shani, M., Givol, D., Neuman, S., Yaffe, D. and Nudel, U., Nature 298, 857-859 (1982)
35. Higgs, D.R., Goodbourn, S.E.Y., Lamb, J., Clegg, J.B., Weatherall, D.J. and Proudfoot, N.J., Nature, in press
36. Montell, C., Fisher, E.F., Caruthers, M.H. and Berk, A.J., Nature, in press