The EMBO Journal vol.7 no.11 pp.3337 - 3344, 1988

Nuclear scaffold attachment sites in the human globin gene complexes

A.P.Jarman and D.R.Higgs

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

Communicated by D.J.Weatherall

In an analysis of a 90-kb region around the human β-globin gene complex we have identified at least eight sites of attachment to the nuclear scaffold (SARs). While these have many potential functions, there appears to be a particular association with sequences important in the regulation of the complex. Two SARs are close to the known enhancer-like elements of the β-globin gene. SARs flanking the complex co-habit with the boundaries of the putative β-like globin gene regulatory domain. In contrast, we have detected no SARs within a 140-kb region of the human α-globin gene complex. If SARs play a role in the regulation of gene expression then this structural difference would imply a difference in the regulation of the two complexes.

Key words: nuclear matrix/scaffold-associated regions/globin genes/developmental regulation

Introduction

Eukaryotic chromatin appears to be organized into large loops or domains, the bases of which are attached to the nuclear matrix or scaffold (Berezney and Coffey, 1974; Benyajati and Worcel, 1976; Cook and Brazell, 1975). As well as a mode of structural packaging for chromosomes, it seems likely that these loops have functional significance. Chromosomal interactions with the matrix may be important in mitosis, replication and gene regulation, consequently the nature of the regions of DNA attachment to the matrix are of considerable interest and several methods for their identification have been reported. One particular method has identified candidates for a specific class of matrix associations in the vicinities of a number of genes. This method relies on the preparation of nuclear matrices using lithium-3',5'-di-iodosalicylate (LIS) as the extracting agent in place of the usual high salt buffers; the structures so defined are termed nuclear scaffolds (Mirkovitch *et al.*, 1984).

The regions of attachment thus identified in *Drosophila* (scaffold-associated region, SAR) have been characterized as short (~500 bp), A/T-rich segments of DNA closely flanking those genes analysed. Certain sequence elements are shared by the SARs, the most important being the increased occurrence of consensus interaction sequences for topoisomerase II (Gasser and Laemmli, 1986a). That this is more than coincidence is suggested by the fact that this enzyme is an abundant component of both the nuclear matrix (Berrios *et al.*, 1985) and chromosomal scaffold (Lewis and Laemmli, 1982; Earnshaw *et al.*, 1985). A number of developmentally *cis*-regulated genes are bounded by SARs. Frequently these coincide with genetically identified *cis*-regulatory elements (termed co-habitation, Gasser and Laemmli, 1986b). The presence of bounding SARs invites the notion that these define the limits of the independent regulatory domain for the genes in question. Strikingly, this appears to be the case for two *Drosophila* genes, in which only transgenes including the flanking SARs are consistently immune from position effects (Gasser and Laemmli, 1986b). In the vertebrate genes that have been analysed by this technique (Cockerill and Garrard, 1986; Kas and Chasin, 1987; Phi-Van and Stratling, 1988), the SARs detected have very similar structural features to those of *Drosophila*; furthermore, one is close to a *cis*-regulatory enhancer and others are at the boundaries of an active chromatin domain.

Here, we present an analysis of the human α- and β-globin gene complexes to determine the presence and functional implications of SARs detectable by this method. A number of SARs are detected in the β-globin gene complex, two of which have been precisely localized close to the β gene. They conform to the previously observed structure of SARs and, significantly, they exhibit co-habitation, each being close to a known *cis*-regulatory sequence for the β-globin gene. The remaining SARs within the complex may identify sites of possible regulatory significance that should be investigated. It has been proposed that regions flanking the whole complex that include 'super' hypersensitive sites for DNase I may represent the domain boundaries of the complex (Tuan *et al.*, 1985), and recently it has been shown that these regions are necessary for allowing complete positionally independent high-level expression of β-globin (Grosveld *et al.*, 1987). We show that SARs are present in close association with both these areas. In contrast, no SARs have been detected within 140 kb of the α-globin gene complex. If SARs are playing a role in gene expression then this apparent difference between two coordinately expressed complexes raises important questions concerning modes of gene regulation, perhaps pointing to fundamentally different mechanisms of expression control in the two complexes.

Results

Attachment sites occur close to the β-globin genes, but not to the α-globin genes

In the basic method of SAR detection, isolated nuclei are treated with LIS, an agent that extracts all histones and many non-histone proteins. This leaves a stable, residual structure resembling the nuclear matrix I (Lebkowski and Laemmli, 1982) to which decondensed, naked DNA is still bound (although similar to the nuclear matrix, this structure has been named the nuclear scaffold to indicate the different extraction procedure). After exhaustive washing to remove

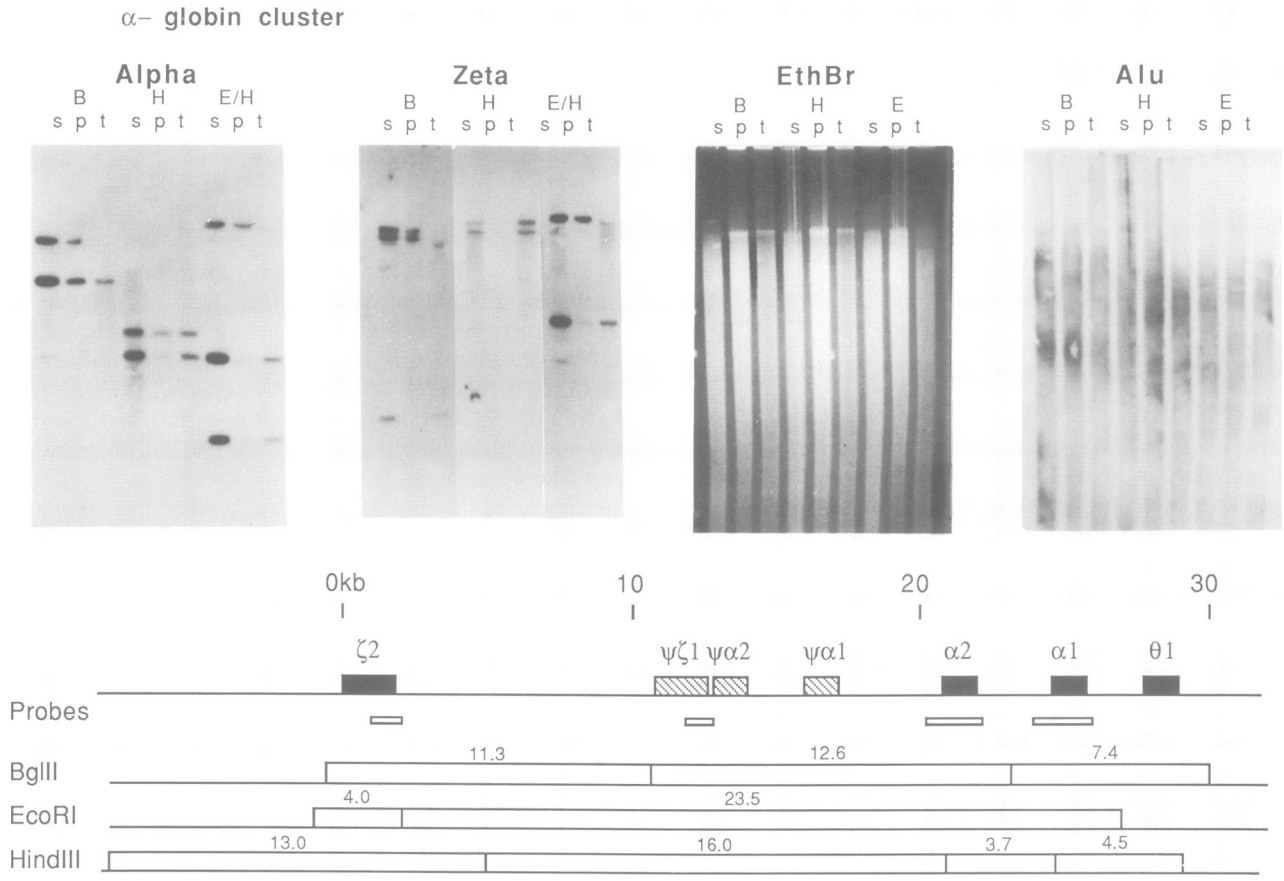

Fig. 1. Screening the vicinity of the α-like globin genes for SARs. Supernatant (S) and pellet (P) DNA fractions from restriction enzyme-digested LIS matrices, and total human DNA were probed with sequences from the ζ- and α-globin genes. The autoradiographs of three such digests (B, *Bgl*II; H, *Hin*dIII; E, *Eco*RI, or *Eco*RI/*Hin*dIII) show all fragments to be released (depleted from the pellet fraction). The fragments detected are represented by the open boxes on the map beneath. The position of the probe sequences is indicated by small boxes below the genes. Ethidium bromide staining before blotting shows that approximately equal weights of DNA have been loaded in supernatant, pellet and total DNA tracks as in the example shown. Hybridization of this with a probe to the Alu dispersed repetitive element confirms this. The ζ-globin gene probe is a 1.8-kb *Sac*I fragment from pBRζ, the α-globin gene probe is a 1.5-kb *Pst*I fragment from pDH7. Map coordinates are shown with the origin being the ζ2-globin mRNA cap site. Some of the high mol. wt fragment in the total human DNA tracks are poorly visible on the photograph due to a blotting artefact on the original autoradiograph.

the LIS, restriction enzymes are added. Complete digestion of the DNA is achieved in a few hours, by which time ~75% of the DNA is solubilized (the S fraction). The remaining fragments are firmly bound to the scaffold which can be collected by centrifugation (the P fraction). DNAs are isolated from the two fractions and then electrophoresed in equal amounts (by weight) on an agarose gel, and transferred to a nylon membrane. Probing with cloned DNA fragments of interest will then reveal a differential enrichment of particular fragments in one or other fraction. Thus, assuming a non-random binding, most fragments will be expected to predominate in the S fraction, but any including an SAR will be enriched in the P fraction.

Our initial aim was to identify potential SARs within the immediate vicinity of the globin genes using this assay (blot-hybridization assay). For this, nuclear scaffolds prepared from K562 cells (a human haemopoietic cell line) were digested with *Bgl*II, *Hin*dIII and *Eco*RI, and Southern blots made. In general, when using a variety of probes, we observe that most fragments are released to a greater or lesser degree (~0–50% remaining in the P fraction) and are defined here as not containing SARs. Of these, the proportion remaining bound is, to a certain extent, fragment dependent, and also seems to be length dependent such that large fragments (particularly >10 kb) are invariably partially retained in the pellet fraction. This is probably due to non-specific interactions between incompletely extracted proteins. Nevertheless, a proportion of fragments are very clearly partitioned into the pellet fraction (~80–100%), and are thus operationally defined as containing the SARs of interest.

All fragments detected using α-like globin gene probes show 50% or less partitioning in the pellet fraction, and thus we conclude that no SARs exist in the near vicinity of the α-globin complex (Figure 1). In contrast, a number of fragments from the β-globin complex are very specifically bound (Figure 2): the fragments detected by the β-gene probe itself in all three digests are strongly enriched in the pellet fraction (>90%). Furthermore, both β-specific *Eco*RI fragments are pellet enriched (upstream 5.5-kb and downstream 3.8-kb fragments). Similarly, δ- and ψβ-specific fragments are scaffold associated (although the smaller *Bgl*II ψβ fragment is bound more weakly, ~50%, see later) (data not shown). All γ-globin fragments are, in contrast, released (~50% $^{G}\gamma$, ~95% $^{A}\gamma$), but a further SAR is detected upstream of the ε-globin gene. Here, a 5′-extended 8.1-kb

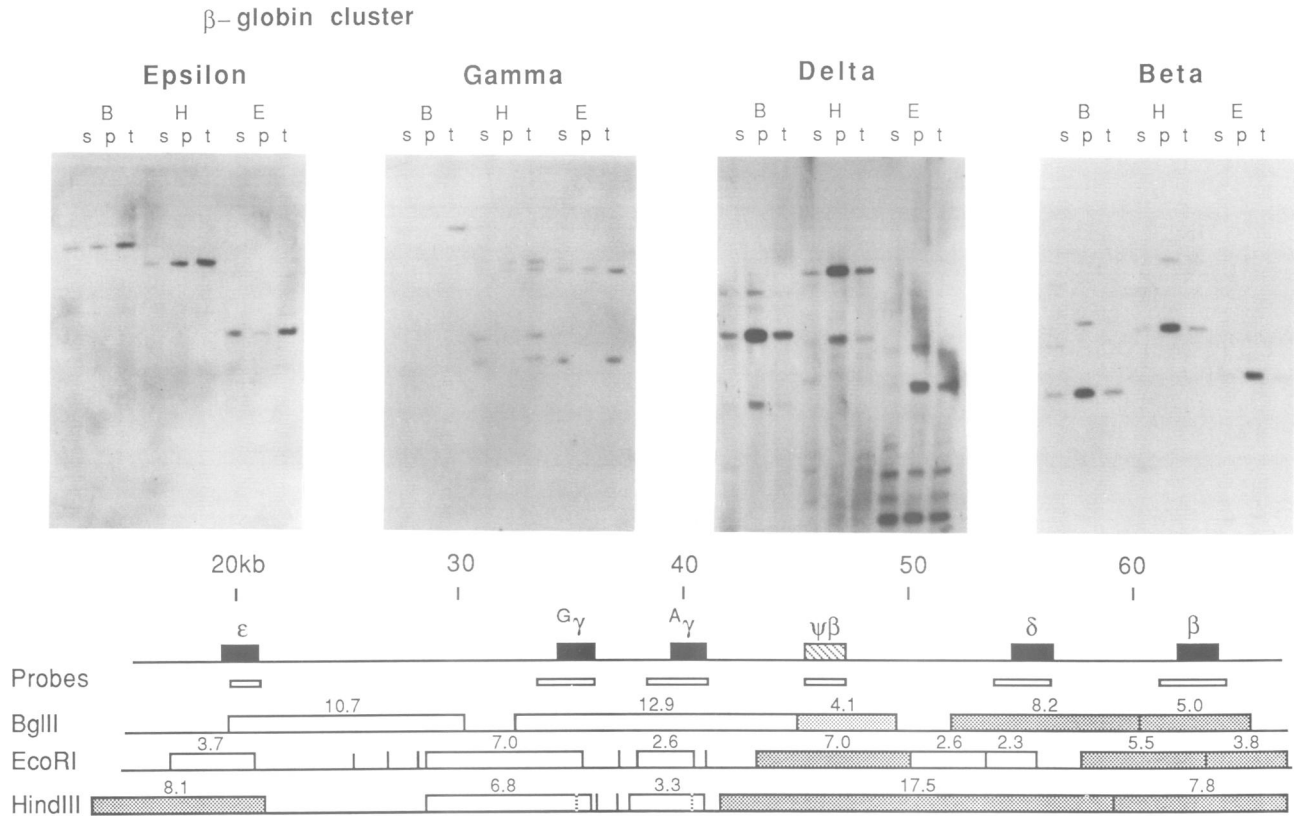

Fig. 2. Screening the vicinity of the β-like globin genes for SARs. The autoradiographs are presented as indicated in Figure 1. When probed with sequences from the β-like globin genes a range of attached and released fragments are detected. Whereas a number of fragments are released (such as those detected by the γ-globin probe, indicated by open boxes) certain fragments are clearly pellet associated and thus contain SAR elements (indicated on the map by shaded boxes). The 4.1-kb *Bgl*II fragment detected by ψβ is equivocally bound in this assay. The β-globin autoradiograph is of a hybridization with a 1.9-kb *Bam*HI fragment which detects only a single 5.5-kb *Eco*RI fragment (other bands are cross-hybridization with the δ-globin gene). A *Pst*I fragment of the β-globin gene also detects the strongly bound 3.8-kb *Eco*RI fragment (data not shown). Note that K562 is heterozygous for a double *Hin*dIII polymorphism close to the $^{G}\gamma$ and $^{A}\gamma$-globin genes. Map coordinates are shown with the origin as given by the upstream sequence data of Qiliang *et al.* (1985).

*Hin*dIII fragment is pellet enriched. In summary, as judged from the three bound *Bgl*II fragments and upstream *Hin*dIII fragment, at least four SARs have been detected around the β-globin complex using this assay. The length and overlap of the fragments detected do not allow more precise determinations except that the ε-gene-associated SAR (the ε SAR) is 5′ of the gene itself.

Extended analysis of the gene complexes

The initial blot-hybridization analysis using ζ- and α-globin gene probes determined that the α-globin gene complex contrasted with the β-globin complex in not having SARs in close vicinity to the genes. In an attempt to identify the nearest SARs to the α-globin complex we have used a succession of probes from DNA cloned during a chromosomal walk up- and downstream of the genes. The probes used detected overlapping fragments spanning from 100 kb upstream of the ζ2-globin gene and 20 kb downstream of the α1-globin gene. For all the fragments from this 140-kb region, no pellet-enriched fragments above 50% were observed, and therefore no SARs exist (summarized in Figure 3a).

In extending the initial analysis of the β-globin gene complex, we have located further sites of attachment both upstream and downstream of those detected above (summarized in Figure 3b). The probe, 271BE, from 20 kb upstream of the ε-globin gene (−3 kb on the map) detects a short 2.2-kb *Hin*dIII attached fragment that localizes a SAR to within the sequence of the probe itself. Downstream, few probes are useful due to the presence of many highly repetitive DNA elements (Taramelli *et al.*, 1986), but we have used two probes from ~20 kb downstream of the β-globin gene. The probe, RK29 (at +80 kb), detects the attachment of a 2.3-kb *Bgl*II fragment. The 1.15-kb *Eco*RI fragment detected by this probe is released, localizing the SAR to an area just 3′ of the RK29 (at +82 kb). Complementing this, the probe, 264HE, 5 kb 3′ of RK29, detects an attached 5.6-kb *Eco*RI fragment that overlaps this area. The adjoining but non-overlapping 7-kb *Bgl*II fragment detected by this probe is also attached, either indicating a further SAR, or the single SAR straddles the *Bgl*II site.

Confirmation of the blot-hybridization results and localization of SARs

Whilst a number of SAR-containing fragments have been clearly detected, the blot-hybridization assay has a number of disadvantages. First is the equivocal results obtained in a few cases, particularly with larger fragments and also some $^{G}\gamma$ and ψβ-specific fragments. Also, for a finer analysis of the SARs, smaller fragments must be examined, and these can be more difficult to detect with the limited choice of unique probes available. Therefore, we have made use of

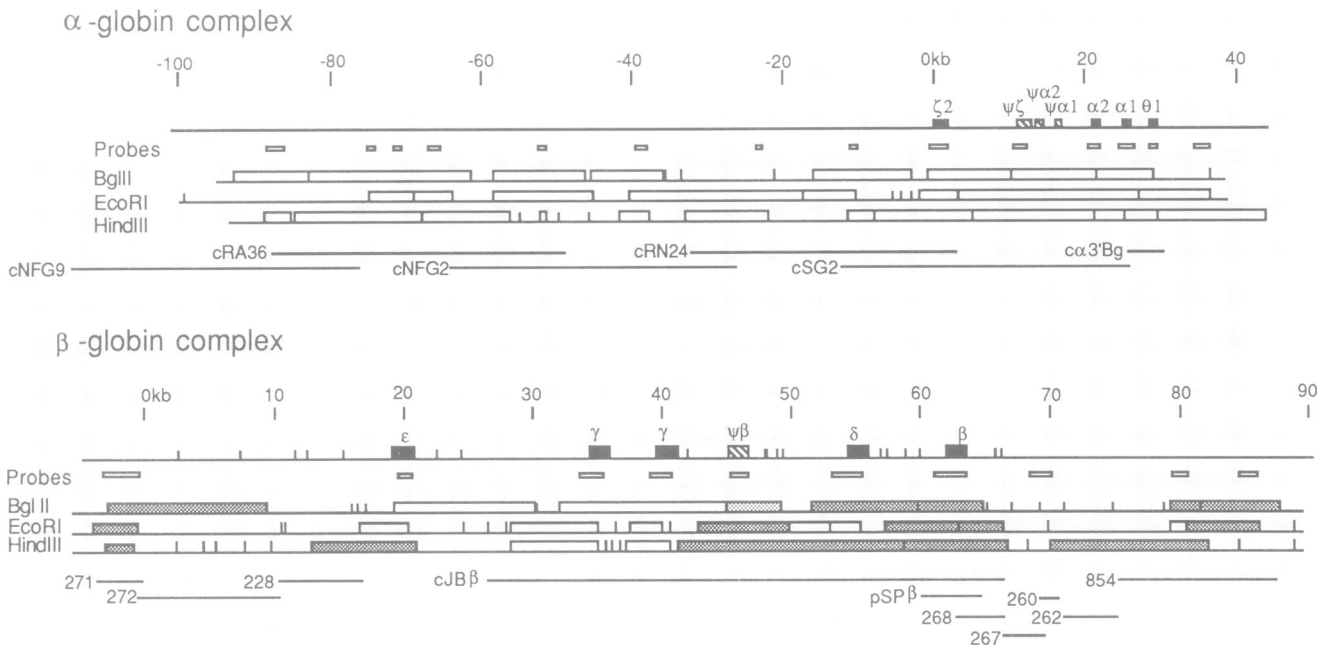

Fig. 3. Extended analysis of the globin complexes for SARs. The total areas covered in and around the two complexes are summarized. Fragments detected in the blot-hybridization assay are shown on the restriction maps as boxes, open and shaded boxes representing released and SAR-containing fragments respectively. Matches to the topoisomerase II consensus interaction sequence are shown as vertical lines on the β map. Probes used (indicated by small boxes) are as follows (left to right): for the α-complex, RA2.2, RA0.6, 5'HVR, RA1.0, RA330, RA1.4, L2, L1, *Sac*1.8(ζ) *Pst*α(α), R1, 3'HVR (Nicholls *et al.*, 1987); for the β-complex, 271BE (a 2.5-kb *Bgl*II−*Eco*RI fragment from pGSE271), ε,γ, ψβ, *Pst*δ, *Pst*β, RK29, 264HE (a 1.0-kb *Hin*dIII−*Eco*RI fragment from pGSE264) (see Antonarakis *et al.*, 1982; pGSE plasmids were provided by F.Grosveld). The clones used in the *in vitro* binding assays are shown below each map.

a modified assay based on the observation that cloned SAR-containing DNA fragments will selectively bind to the scaffold *in vitro* if added during the enzyme digestion step (Gasser and Laemmli, 1986b). The basis of this is not fully understood since the endogenous DNA is not removed beforehand, although it seems likely that the exogenous DNA competes for vacant SAR binding sites.

The cosmid, cJBβ, contains a 39-kb *Asp*718 fragment including the region from upstream of the $^{G}\gamma$ gene to 2 kb downstream of the β-globin gene. When digested with *Bgl*II/*Asp*718, end-labelled, and used in this assay four fragments bind to the matrices, those containing the ψβ-, δ- and β-globin genes along with the 1.4-kb *Bgl*II−*Asp*718 fragment from 3′ of the β gene (which was only detectable in the blot-hybridization as part of the 3.8-kb *Eco*RI fragment) (Figure 4a). All other fragments (including vector fragments, and those of $^{G}\gamma$ globin that are equivocally released in the blot-hybridization assay) are associated to a very low degree. Thus the initial results are confirmed, there are four SARs in the region of cJBβ. A notable difference however is the binding of the 4.1-kb ψβ-globin fragment to the same high degree as the δ- and β-globin fragments in spite of an equivocal binding in the blot-hybridization assay. It is possible that an incomplete extraction of masking proteins from a SAR in this fragment is being observed in the blot-hybridization assay.

We have used this assay to confirm the absence of attachment sites from the 140-kb region around the α-globin gene complex. In order to cover such a large region some very large fragments were included in those analysed by the blot-hybridization assay, which yielded unclear results. We have repeated the analysis using an overlapping cosmid set in the *in vitro* binding assay. Thus, in separate experiments, we have added the following end-labelled cloned fragments to K562 nuclear scaffolds: cNFG9/*Hin*dIII, cRA36/*Eco*RI, cNFG2/*Hin*dIII, cRN24/*Hin*dIII, cSG2/*Hin*dIII, cα3′Bg/*Eco*RI (see Figure 3a). In all cases, it was clear that no specifically bound fragments were observed (data not shown).

We have also used this assay to localize further the SARs associated with the β and 3′ β *Bgl*II fragments. For the former we have analysed a subclone of the 5.0-kb *Bgl*II fragment from the cosmid, called pSPβ. Digesting with *Dra*I, end labelling and adding to scaffolds during *Bgl*II digestion of the endogenous DNA, we find that a single 1972-bp fragment binds strongly, which encompasses from just upstream of the β-globin gene to within its second intron (fragment b, Figure 4b). Further localization was possible using an end-labelled *Eco*RI/*Bam*HI digest of pSPβ. Here, the 795-bp *Bam*HI−*Eco*RI fragment, consisting of almost the entire IVS-2, was bound specifically (fragment b). The overlap between these two fragments thus localizes the SAR to a 520-bp region at the 5′ end of the β-globin gene IVS-2 (hereafter called the β-IVS2 SAR). The SAR just 3′ of the β-globin gene has been localized using pGSE268, a subclone containing the downstream 3.8-kb *Eco*RI fragment that includes the attached 1.4-kb *Bgl*II−*Asp*718 fragment. End-labelled *Bgl*II/*Eco*RI/*Taq*I fragments were incubated with scaffolds during a *Bgl*II digestion, and subsequently a single 807-bp *Bgl*II−*Taq*I fragment was observed to bind strongly (fragment c), the supernatant becoming depleted in the same way as for the parent *Bgl*II−*Asp*718 fragment (Figure 4b). Further mapping was achieved using a *Dde*I digest, from which a single 658-bp fragment bound (fragment a). Thus, the SAR has been localized to a 490-bp region 2 kb downstream of the β-globin gene (the 3′β SAR).

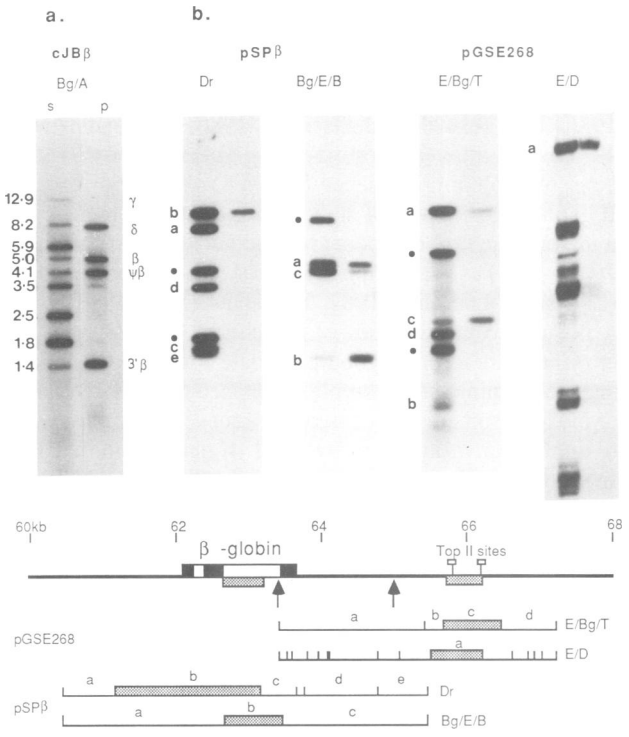

Fig. 4. Use of the *in vitro* binding assay in SAR localization. (**a**) The assay was tested using cosmid, cJBβ (see legend of Figure 3), cut with *Bgl*II and *Asp*718, and end labelled. The four specifically bound fragments correspond to those detected as containing SARs in the blot-hybridization assay as shown in Figure 2. (**b**) The *in vitro* binding assay allowed the localization of the SARs detected near the β-globin gene. Plasmids pSPβ and pGSE268 were digested as indicated, end labelled and used in this assay. The fragments binding strongly in each case are indicated on the map by shaded boxes. A dot beside a fragment indicates that it is vector-derived. The 3'β SAR is located between, but not including, the two Alu repetitive sequence elements downstream of the β-globin gene. Topoisomerase II sites are indicated as are the positions of two regulatory sequences (arrowed). For the *Dra*I digest of pSPβ the end fragments of the insert (a and e) include vector sequences and consequently are of different size on the autoradiograph. For the *Dde*I digest of pGSE268 numerous small fragments are given that are clearly released and are not shown on the diagram. Abbreviations for restriction enzymes are: Bg, *Bgl*II; A, *Asp*718; E, *Eco*RI; B, *Bam*HI; Dr, *Dra*I; D, *Dde*I; T, *Taq*I.

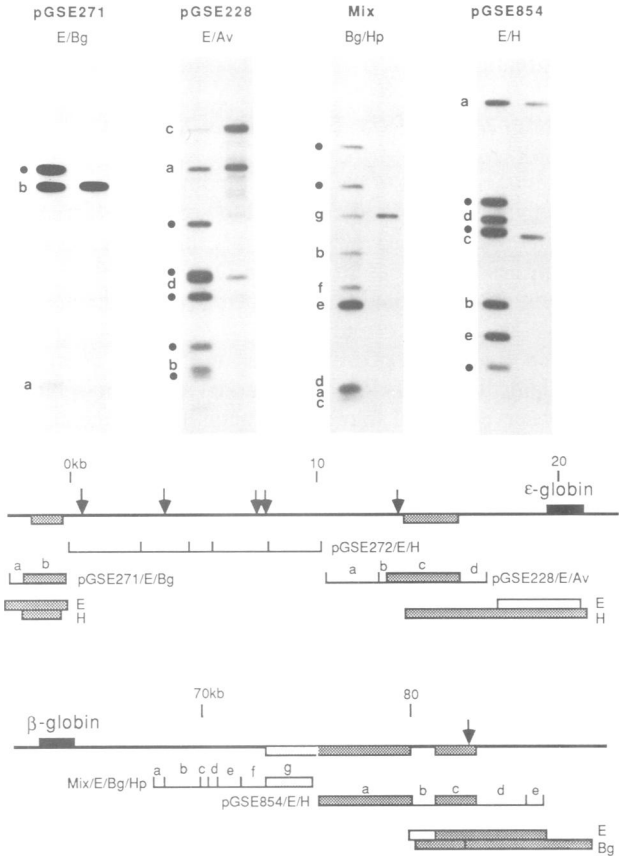

Fig. 5. *In vitro* analysis of the β-globin complex flanking regions. Maps of the upstream and downstream regions analysed are shown, the arrows indicating the strong hypersensitive sites. Relevant results from the blot-hybridization assay using the probes indicated by the small boxes are summarized from Figure 3. As in Figure 4, the clones used in the *in vitro* assay are indicated with matrix-binding fragments as shaded boxes. 'Mix' indicates a digest of three plasmids together, pGSE260, 262, 267, and these are also shown together on the map. The pellet track of 'mix' is overloaded ~4-fold to show the weakly binding fragment. The clones used are: pGSE271, a 3.3-kb *Eco*RI fragment; pGSE272, 10.4-kb *Eco*RI; pGSE228, 6.5-kb *Eco*RI; pGSE267, 3.3-kb *Eco*RI; pGSE260, 1.8-kb *Bgl*II; pGSE262, 3.8-kb *Bgl*II; pGSE854, 12-kb *Hpa*I−*Bam*HI. Abbreviations for restriction enzymes are as follows: E, *Eco*RI; H, *Hin*dIII; Bg, *Bgl*II; Av, *Ava*II; Hp, *Hpa*I.

The $\psi\beta$- and δ-globin SARs have not been mapped in detail here. In a provisional experiment for the $\psi\beta$-SAR, the 4.9-kb *Bgl*II fragment was isolated by electro-elution from a digest of the cosmid, cJBβ. This was re-cut with *Xba*I, end labelled and incubated with scaffolds, upon which a 1900-bp fragment bound. This fragment is located downstream of the $\psi\beta$-globin gene in the $\psi\beta-\delta$ intergenic region (data not shown).

The *in vitro* binding assay has enabled us to analyse the flanking regions of the complex more completely using a contiguous set of plasmid clones. The upstream region was analysed using clones of three contiguous *Eco*RI fragments—pGSE271, 272 and 228 (Figure 5). The region shown by blot hybridization to contain the ϵ SAR is included in pGSE228. From an *Eco*RI/*Ava*II digest of pGSE228 a single 2.0-kb fragment from this region binds strongly to isolated scaffolds, confirming its presence, and the absence of other SARs in this area. SARs are also absent from the area of the 10.4-kb *Eco*RI fragment in pGSE272, no fragments bind specifically in *Eco*RI/*Hin*dIII digests of this plasmid (data not shown). Further upstream, the blot-hybridization result is also confirmed. A 2.5-kb *Bgl*II/*Eco*RI fragment from pGSE271 (fragment b) corresponding to the probe 271BE binds specifically in the *in vitro* binding assay. In summary, two SARs are located in the 23-kb flanking region upstream of the ϵ-globin gene. The region downstream from the 3'β SAR (in pGSE268) has been analysed using the contiguous/overlapping clones, pGSE267, 260, 262, 854 (Figure 5). From an *Eco*RI/*Hin*dIII digest of pGSE854 two fragments bind (Figure 5). The bound 2.2-kb *Hin*dIII fragment (fragment c) localizes further the +82-kb SAR detected by the probe RK29. The 5' 5.8-kb fragment is also bound (fragment a), locating a new SAR. Although not localized further it seems likely that this SAR is at the 5' end of the 5.8-kb fragment, as in an *in vitro* experiment using pGSE267, 260 and 262 together in an *Eco*RI/*Bgl*II/*Hpa*I digest (shown as 'mix' in Figure 5) a single 2.8-kb *Hpa*I−*Bgl*II fragment from pGSE262 is consistently but very weakly bound, suggesting the presence of a single SAR

at the junction of pGSE262 and pGSE854 (Figure 5). In summary, two SARs are located in the 20-kb flanking region downstream of the 3′β SAR.

Characterization of the SARs within the β-globin gene complex

An interesting result of the *in vitro* binding assay using cJBβ is that the 3′β SAR is often bound to a greater degree than the other sites. We have pursued this further by observing the abilities of various exogenous DNA species to act as competitors of the binding of cloned SAR fragments from the *Bgl*II-digested cosmid.

When linearized pGSE268 was used as a competitor (containing the 3′β SAR) a very marked reduction of binding was observed for all four SAR-containing fragments (data not shown). In contrast, competing with non-specific DNA in the form of sheared *Escherichia coli* DNA or linearized pSP64 produced very little reduction in binding (up to 25 $\mu g/A_{260}$ unit of nuclei), although the background of non-specific association of non-SAR fragments was reduced. We conclude that all four sites are specific associations, and furthermore involve the same underlying interactions. When competing with linearized pSPβ (containing the β-IVS2 SAR), however, a steady reduction in binding of all fragments except the 3′β SAR was observed. This SAR reduced in binding only slightly. These observations suggest that the β-IVS2- and 3′β-SARs have lower and higher affinities for the binding sites, or else that the β-IVS2-SAR-type sites are a subset of general sites that are used by the 3′β SAR. If non-specific competitor DNA (pSP64) is added in native supercoiled form, a small but steady depletion is observed from the bound fraction, although this is less marked for the 3′β SAR. This suggests that DNA supercoiling may be an important factor in the association with the matrix *in vivo*.

All these experiments have been carried out using scaffolds prepared from K562 nuclei. Although tissue-specific differences in SAR binding have not been found to be a feature of these sites in general we have looked at SAR binding of the cJBβ/*Bgl*II/*Asp*718 fragments to scaffolds prepared from HeLa nuclei and those from an Epstein–Barr virus (EBV)-transformed lymphocyte cell line. In each case the same four *Bgl*II fragments bound (as in Figure 4a), so that the same potential binding sites must be present. There are differences in the strength of binding (although the 3′β SAR is always strongest) but this probably reflects different optimum isolation conditions for nuclei and scaffolds of various cell types.

Discussion

Structural analysis of the β-IVS2- and 3′β SARs

Of the SARs detected within the β-globin complex, two have been localized to specific short regions: the β-IVS2 SAR to a 520-bp region of the β-globin gene's IVS2 and the 3′β SAR to a 490-bp region that is 2 kb downstream of the gene. In common with previous studies in other species, both SARs are characterized by being A/T-rich, 75% in each case. Significantly, strong scaffold association is not determined by a simple A/T content of fragments. Within the β-globin complex we find that there are other regions of equally high A/T content that are not bound. Further, competition with non-specific DNA from *E.coli* (with many A/T-rich regions) had very little effect on the SAR associations. It would therefore seem that specific sequence elements determine SAR behaviour. Putative elements have been located in SARs from both *Drosophila* and mouse (Gasser and Laemmli, 1986a,b; Cockerill and Garrard, 1986; Kas and Chasin, 1987), the most interesting being matches to the 15-bp consensus binding sequence for *Drosophila* topoisomerase II. This is potentially of significance as topoisomerase II is a major structural component of the nuclear matrix (Berrios *et al.*, 1985) as well as the metaphase chromosome scaffold (Lewis and Laemmli, 1982; Earnshaw *et al.*, 1985) where it is situated at the bases of the chromosomal loops. The enzyme is therefore a good candidate for the SAR binding sites. Screening the β-complex sequence for >84% identity to the topoisomerase II consensus sequence with full identity to the 6-bp invariant core, we find 18 matches in the 50 kb from the ε-globin gene to the 3′β SAR (shown in Figure 3). Two of these fall within the 490-bp 3′β SAR sequence itself. Two other clusterings occur: a group of four sites falls in the region to which the ψβ SAR has been localized and seven sites cluster in a 3.5-kb region between the δ- and β-globin genes, which is part of the 8.1-kb δ-globin-containing *Bgl*II fragment that is matrix associated. At variance with this, however, no strong consensus sites are present within or near the β-IVS2-SAR. Perhaps topoisomerase II sites may constitute a variable part of the recognition requirements of SARs. The differences we have observed in the binding behaviours of the β-IVS2- and 3′β-SARs may be due to this. As well as a structural role of 'loop-fastener', topoisomerase II may also have an enzymatic role in manipulating the supercoiled state of chromatin loops, which may be important in both replication and gene expression.

Sequences resembling A- and T-boxes (Gasser and Laemmli, 1986b) and the hexanucleotide, ATATTT (Cockerill and Garrard, 1986) are also present in the two SARs, but their significance is difficult to gauge considering the AT-richness of the sequences.

SARs lie near functionally important regions

There is an accumulating body of evidence to suggest that SARs represent elements of physiological significance; in particular, there is a correlaton of SARs with known *cis*-acting regulatory sequences. In this context the two SARs localized close to the β-globin gene are particularly interesting. A *cis*-regulatory region within the β-globin gene second intron has been localized, which confers a β-like pattern of expression on the γ-globin gene (Behringer *et al.*, 1987). The element lies at the 3′ end of this intron (Antoniou *et al.*, 1988) and thus the β-IVS2 SAR is either within or very close to it. This kind of arrangement (termed 'co-habitation') has been frequently observed in *Drosophila*: the three highly tissue-specific genes, *ftz, Sgs*-4 and *Adh* each have nearby SARs that overlap with elements known to be important in their correct tissue-specific control (Gasser and Laemmli, 1986b).

Similarly, the 3′β SAR is closely associated with the tissue-specific enhancer of the β-globin gene (Kollias *et al.*, 1986, 1987), which has been localized to a 700-bp *Acc*I–*Dra*I fragment 3′ of the β-globin gene itself. The 3′β SAR is thus situated ~500 bp downstream from this, giving rise to a situation highly reminiscent of the mouse κ-chain gene, where a SAR is also closely juxtaposed with the

enhancer sequence in the J−C_x intron (E_x). Kollias *et al.* (1987) do not exclude the possibility that the complete β-gene enhancer extends further downstream. These instances of co-habitation support a putative role in gene expression of at least some SARs in which they bring nearby *cis*-regulatory sequences into close proximity with *trans*-acting regulatory factors residing at the nuclear scaffold, so facilitating their efficient interaction. The potential significance of the SARs associated with the other genes remains to be determined, but they may point to other regions in the β-globin complex of possible regulatory importance.

Deduced loop sizes (3−15 kb) are much smaller than mean sizes predicted for bulk DNA by biophysical methods (50−200 kb, Cook and Brazell, 1975; Benyajati and Worcel, 1976; Igo-Kemenes and Zachau, 1977). A possible explanation is that SARs are clustered in the immediate vicinities of genes to form 'mini-loops'. Secondly, at best, the assays used here detect regions of DNA with the potential for SAR behaviour (Gasser and Laemmli, 1986b). The observed lack of tissue specificity and the ability of exogenous cloned DNA to bind (the *in vitro* binding assay) suggest this. Many of these regions may not be ordinarily or permanently associated with the scaffold *in vivo* perhaps due to blocking by proteins that are removed on LIS extraction. The association of SARs with the scaffold 'compartment' may then be developmentally controlled.

Certain SARs flanking a gene complex may mark the limits of the independent functional domain unit. Such a domain unit has been defined for the *Drosophila ftz* and *Sgs*-4 genes, and in each case only constructs including both upstream and downstream SARs consistently show no position effects. The active chromatin domain of the chick lysozyme gene also has SARs at its boundaries (Phi-Yan and Stratling, 1988). Similarly, three of the SARs detected here (−3, ϵ and +82 SARs) are located at the boundaries of the β-globin complex functional domain (marked by super hypersensitive DNase I sites, Tuan *et al.*, 1985; see Figure 5). These boundary regions, including the SARs, have recently been shown to support the position-independent high-level expression of the β-globin gene in transgenic mice (Grosveld *et al.*, 1987). As putative domain boundary elements these SARs may be postulated to act as fixed boundaries, isolating and insulating a gene complex from the effects and activities of neighbouring gene complexes. A finer localization of the functional sequences in these areas is of great interest. The two enhancer elements co-habiting with the β-IVS2- and 3′β SARs were originally detected as DNase I-hypersensitive sites (Groudine *et al.*, 1983). As a corollary, three of the flanking SARs are very close to hypersensitive sites, possibly indicating the positions of important regulatory elements.

No association between SARs and illegitimate recombination events

We find no clear associations between SARs and the break-points of the many deletions of the β-globin gene complex. This includes the group of equal-sized but staggered large deletions for which a model involving excision of a chromosomal loop across its base has been proposed (Vanin *et al.*, 1983; Taramelli *et al.*, 1986; Collins *et al.*, 1987). This is no surprise as the staggered nature of the deletions mitigates against the involvement of fixed attachment sites. It is likely that loop excision occurs during the transient association between the chromosome and the matrix at the point of replication (not detected by the assay used here, see Jackson and Cook, 1986; Razin, 1987).

A fundamental distinction between the α- and β-globin gene complexes

We have demonstrated the existence of at least eight SARs within 90 kb of the β-globin gene complex. In contrast, no SARs were located in 140 kb of the α-globin gene complex. How can this be reconciled with a putative role in gene expression? In this context, it is interesting that the combined data from expression assays (Treisman *et al.*, 1983; Charnay *et al.*, 1984; Wright *et al.*, 1984) together with observations on methylation (Bird *et al.*, 1987; Fischel-Ghodsian *et al.*, 1987b), long-range structure (Bernardi *et al.*, 1985; Fischel-Ghodsian *et al.*, 1987a) and replication timing (Holmquist, 1987) indicate that the processes underlying α- and β-globin gene expression may be different. It is tempting to speculate that the distribution of SARs may also reflect fundamentally different modes of gene expression control.

Materials and methods

The techniques of matrix isolation and analysis are derived from Mirkovitch *et al.* (1984) and Gasser and Laemmli (1986b).

Isolation of nuclei

Nuclei were isolated from cell cultures using the protocol and solutions given by Mirkovitch *et al.* (1984). The nuclei were generally stored at −20°C as 5 A_{260} unit aliquots in isolation buffer and EDTA made 50% (v/v) with glycerol.

LIS extraction of nuclei

Aliquots of nuclei were prepared for analysis by washing in isolation buffer without EDTA and stabilized by incubation at 37°C for 20 min. Removal of histones by extraction with 25 mM LIS was achieved according to the method of Mirkovitch *et al.* (1984) with modifications (Gasser and Laemmli, 1986a) except that 1.5 ml of LIS extraction buffer was used, extraction and subsequent washing being carried out in 1.6 ml Eppendorf tubes. Extracted nuclei were washed five times in digestion buffer.

Blot-hybridization assay

Washed, LIS-extracted nuclei from 2 A_{260} units per analysis were digested with restriction enzymes (Boehringer, Amersham) in a 400-μl total volume (1 U/μl enzyme concentration) with occasional agitation for 3−5 h. Afterwards, released and bound DNA fractions were separated by pelleting the remnant nuclei in a microfuge for 2 min, the supernatant set aside (S fraction) and the pellet washed before being resuspended in 400 μl digestion buffer (P fraction). DNA was purified from each fraction (Mirkovitch *et al.*, 1984) and its concentration checked by absorption at 260 nm and ethidium bromide visualization of aliquots run on an agarose gel.

Samples (5 μg) were run on 0.8% agarose gels with total genomic DNA aliquots alongside for comparison. Gel electrophoresis, blotting onto nylon filters and subsequent hybridizations with labelled probes were by standard methods (Old and Higgs, 1982; Maniatis *et al.*, 1982). The use of nylon filters (Hybond-N, Amersham) allowed multiple successive hybridizations with different probes, giving a controlled comparison between probes. The probes used are given in the legends to Figures 1, 2 and 4. All were labelled by multiprime extension using [α-^{32}P]dCTP (Amersham).

In vitro assay

In general, 100−500 ng of cloned DNA were digested with appropriate restriction endonucleases, and 3′-end labelled using Klenow enzyme and appropriate [α-^{32}P]nucleoside triphosphates. Unincorporated nucleotides were removed by centrifugation through a Sephadex G-50-80 column, and the labelled DNA purified by phenol/chloroform extraction and ethanol precipitation, finally resuspended in digestion buffer. 5−100 ng labelled DNA (1−2 × 10^5 c.p.m.) were added to LIS-extracted nuclei from 1.5 A_{260} units of intact nuclei at the commencement of the endonuclease digestion step as given above. Subsequent steps of DNA fractionation and isolation were also as given above. Aliquots of the S and P fractions were

electrophoresed on agarose gels (between 0.8 and 2%) in a typical ratio of two to five times S to P with respect to counts of ^{32}P loaded, the gels dried down and autoradiographed. In addition to the endogenous nuclear DNA present it was found that further inclusion of unlabelled non-specific competitor DNA (*E.coli* sheared chromosomal DNA) in the digestion step could reduce background in cases where non-specific probe binding was a problem. In the competition experiments themselves, up to 50 μg of unlabelled competitor DNA were added at the same time as the labelled probe DNA.

Acknowledgements

We thank W.G.Wood and R.W.Jones for helpful discussions and comments on the manuscript, R.Jones for providing the clones pSPβ and cJBβ and L.Roberts for tireless typing. The pGSE clones were a kind gift from F.G.Grosveld and D.R.Greaves. We also thank D.J.Weatherall for helpful comments and continued support.

References

Antonarakis,S.E., Boehm,C.D., Giardinia,P.J.V. and Kazazian,H.H. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 137.

Antoniou,M., deBoer,E., Habets,G. and Grosveld,F. (1988) *EMBO J.*, **7**, 377–384.

Behringer,R.R., Hammer,R.E., Brinster,R.L., Palmiter,R.D. and Townes,T.M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7056–7060.

Benyajati,C. and Worcel,A. (1976) *Cell*, **9**, 393–407.

Berezney,R. and Coffey,D. (1974) *Biochem. Biophys. Res. Commun.*, **60**, 1410–1419.

Bernardi,G., Olofsson,B., Filipski,J., Zerial,M., Salinas,J., Guny,G., Meumier-Rotival,M. and Rodier,F. (1985) *Science*, **228**, 953–958.

Berrios,M., Osheroff,N. and Fisher,P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4142–4146.

Bird,A.P., Taggart,M.H., Nicholls,R.D. and Higgs,D.R. (1987) *EMBO J.*, **6**, 999–1004.

Charnay,P., Treisman,R., Mellon,P., Chao,M., Axel,R. and Maniatis,T. (1984) *Cell*, **38**, 251–263.

Cockerill,P.N. and Garrard,W.T. (1986) *Cell*, **44**, 273–282.

Collins,F.S., Cole,J.L., Lockwood,W.K. and Iannuzzi,M.C. (1987) *Blood*, **70**, 1797–1803.

Cook,P.R. and Brazell,I.A. (1975) *J. Cell Sci.*, **19**, 261–279.

Earnshaw,W.C., Halligan,B., Cooke,C.A., Heck,M.M.S. and Liu,L.F. (1985) *J. Cell. Biol.*, **100**, 1706–1715.

Fischel-Ghodsian,N., Nicholls,R.D. and Higgs,D.R. (1987a) *Nucleic Acids Res.*, **15**, 6197–6207.

Fischel-Ghodsian,N., Nicholls,R.D. and Higgs,D.R. (1987b) *Nucleic Acids Res.*, **15**, 9215–9225.

Gasser,S.M. and Laemmli,U.K. (1986a) *EMBO J.*, **5**, 511–517.

Gasser,S.M. and Laemmli,U.K. (1986b) *Cell*, **46**, 521–530.

Grosveld,F., Van Assendelft,G.B., Greaves,D.R. and Kollias,G. (1987) *Cell*, **51**, 975–985.

Groudine,M., Kohwi-Shigematsu,T., Gelinas,R., Stamatoyannopoulos,G. and Papayannopoulou,T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7551–7555.

Holmquist,G.P. (1987) *Am. J. Hum. Genet.*, **40**, 151–173.

Igo-Kemenes,T. and Zachau,H.G. (1977) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 109–118.

Jackson,D.A. and Cook,P.R. (1986) *EMBO J.*, **5**, 1403–1410.

Kas,E. and Chasin,L.A. (1987) *J. Mol. Biol.*, **198**, 677–692.

Kollias,G., Wrighton,N., Hurst,J. and Grosveld,F. (1986) *Cell*, **46**, 89–94.

Kollias,G., Hurst,J., deBoer,E. and Grosveld,F. (1987) *Nucleic Acids Res.*, **15**, 5739–5747.

Lebkowski,J.S. and Laemmli,U.K. (1982) *J. Mol. Biol.*, **156**, 325–344.

Lewis,C.D. and Laemmli,U.K. (1982) *Cell*, **29**, 171–181.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mirkovitch,J., Mirault,M.-E. and Laemmli,U.K. (1984) *Cell*, **39**, 223–232.

Nicholls,R.D., Fischel-Ghodsian,N. and Higgs,D.R. (1987) *Cell*, **49**, 369–378.

Old,J.M. and Higgs,D.R. (1982) In Weatherall,D.J. (ed.), *Methods in Haematology*. Churchill Livingstone, London, pp. 74–102.

Phi-Van,L. and Stratling,W.H. (1988) *EMBO J.*, **7**, 655–664.

Qiliang,L., Powers,P.A. and Smithies,O. (1985) *J. Biol. Chem.*, **260**, 14901–14910.

Razin,S.V. (1987) *BioEssays*, **6**, 19–23.

Taramelli,R., Kioussis,D., Vanin,E., Bartram,K., Groffen,J., Hurst,J. and Grosveld,F.G. (1986) *Nucleic Acids Res.*, **14**, 7017–7029.

Treisman,R., Green,M.R. and Maniatis,T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7428–7432.

Tuan,D., Solomon,W., Qiliang,L.S. and Irving,M.L. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6384–6388.

Vanin,E.F., Henthorn,P.S., Kioussis,D., Grosveld,F. and Smithies,O. (1983) *Cell*, **35**, 701–709.

Wright,S., Rosenthal,A., Flavell,R. and Grosveld,F. (1984) *Cell*, **38**, 265–273.

Received on May 31, 1988; revised on July 20, 1988

Note added in proof

Since the submission of this work, it has been reported that SARs may be associated with the boundaries of the regulatory domain of the human interferon-β gene [Bode,J. and Maass,K. (1988) *Biochemistry*, **27**, 4706–4711].