The EMBO Journal vol.13 no.17 pp.4126–4136, 1994

Competition between nuclear localization and secretory signals determines the subcellular fate of a single CUG-initiated form of FGF3

Paul Kiefer, Piers Acland, Darryl Pappin, Gordon Peters and Clive Dickson[1]

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

[1]Corresponding author

Communicated by G.Warren

The presumed open reading frame for mouse FGF3, starting at the most 5′ AUG codon, predicts a hydrophobic N-terminus characteristic of a signal peptide for secretion. However, in reticulocyte lysates and transfected COS-1 cells, the full-length *Fgf-3* cDNA is translated almost exclusively from an upstream CUG codon. The resultant products are distributed in both the nucleus and the secretory pathway, implying that the single CUG-initiated form of FGF3 has dual fates. By analysing a series of deletion and replacement mutants and by linking parts of FGF3 to a heterologous protein, we show that secretion is mediated by cleavage adjacent to the previously defined signal peptide, whereas nuclear localization is determined primarily by a classical but relatively weak bipartite motif. In the context of FGF3, nuclear localization also requires the N-terminal sequences which lie upstream of the signal peptide. Thus, the subcellular fate of FGF3 is determined by the competing effects of signals for secretion and nuclear localization within the same protein, rather than by alternative initiation or processing.

Key words: fibroblast growth factors/nuclear localization/protein initiation/protein secretion

Introduction

In eukaryotic cells, the translation and subcellular localization of newly synthesized proteins can be determined by linear sequence motifs within the mRNA and the encoded protein. Thus, translation initiation generally conforms to the predictions of the ribosome scanning model in that the first in-frame methionine codon is preferred but the adjacent sequence context is critical (Kozak, 1989; reviewed in Cavener and Ray, 1991). Relatively rare exceptions have been reported where translation intitiates at non-AUG codons, particularly CUG (Kozak, 1991a). Although first reported in the human *MYC* gene, the fibroblast growth factor (FGF) family provides two notable examples of this phenomenon in which CUG initiation may have important biological consequences (Hann *et al.*, 1988; Florkiewicz and Sommer, 1989; Prats *et al.*, 1989; Acland *et al.*, 1990). In the mammalian genes for FGF2 and FGF3, the presumptive AUG at the start of the open reading frame is preceded by one or more in-frame CUG codons and it has been clearly demonstrated that these function as initiation sites for N-terminally extended products. Moreover, in both cases a proportion of the CUG-initiated protein appears to accumulate directly in the cell nucleus, rather than being exported into the extracellular compartment (Acland *et al.*, 1990; Renko *et al.*, 1990; Bugler *et al.*, 1991).

Proteins that function in the nucleus frequently contain specific nuclear localization signals (NLS) that are required for active translocation through the nuclear pore complex (reviewed in Dingwall and Laskey, 1991; Garcia-Bustos *et al.*, 1991). Typically, these signals comprise small clusters of basic amino acids, often flanked by proline residues, a classic example being the PKKKR motif in SV40 T-antigen (Kalderon *et al.*, 1984). However, more extensive surveys of nuclear proteins have suggested that the surrounding sequence context and the number of NLS motifs in the polypeptide can have important effects on the rate of nuclear uptake (Lanford *et al.*, 1986; Roberts *et al.*, 1987; Dworetzky *et al.*, 1988). Moreover, many NLS motifs function in pairs, perhaps the best characterized example being the bipartite signal in nucleoplasmin (Robbins *et al.*, 1991). As well as constraints on the distance between the basic domains, their amino acid composition can be critical. For example, substituting arginine for lysine in the SV40 T-antigen motif was shown to considerably reduce its influence. Conversely, adding extra copies of the weakened signal could compensate for these effects (Roberts *et al.*, 1987; Dworetzky *et al.*, 1988).

Although the mechanistic details have yet to be determined, it appears that NLS motifs are recognized by specific binding proteins (NSB) that bring otherwise cytoplasmic polypeptides to the nuclear pore complex (reviewed in Dingwall and Laskey, 1991; Garcia-Bustos *et al.*, 1991; Laskey and Dingwall, 1993). Both this step and subsequent translocation through the nuclear pore are energy dependent (Richardson *et al.*, 1988; Akey and Goldfarb, 1989; Moore and Blobel, 1992). Such processes therefore bear a number of similarities to the other major protein sorting system, namely the translocation of nascent polypeptides into the lumen of the endoplasmic reticulum (ER). A large body of literature has established that secretory proteins enter the ER by a mechanism that involves binding of the signal recognition particle (SRP) to an amino acid sequence in the primary translation product as it emerges from the ribosome (Water and Lingappa, 1986; reviewed in Gilmore, 1993). The most obvious characteristic of a signal peptide is a core of hydrophobic amino acids, but their variable length and composition suggest that they conform to a structural rather than a sequence consensus.

In this study, we have been investigating the influence that these classical signals might have on the expression and subcellular fate of FGF3. We previously showed that mouse FGF3 expressed from the most 5′ AUG codon in

the cDNA sequence enters the secretory pathway under the influence of a fairly typical signal peptide (Dixon *et al.*, 1989; Kiefer *et al.*, 1993a,b). This would be in line with its proposed role as a paracrine signalling molecule during embryogenesis, as well as its proven oncogenic properties in virally induced mouse mammary tumours and in transgenic models (reviewed in Peters, 1991). However, in both the mouse and human genes, the open reading frame extends upstream, by an additional 87 and 78 nucleotides respectively, to an in-frame CUG codon (Brookes *et al.*, 1989; Acland *et al.*, 1990). By mutating the respective CUG and AUG codons in mouse FGF3, we showed that the CUG served as an initiation codon both *in vitro*, in rabbit reticulocyte lysates, and *in vivo*, in COS-1 cells transfected with appropriate expression vectors (Acland *et al.*, 1990). Moreover, both cell fractionation and immunofluorescence assays demonstrated that a significant proportion of the CUG-initiated product was localized in the nucleus of transfected cells.

An obvious interpretation of these findings, based on precedents set by other proteins that have dual subcellular fates, was that the two forms of FGF3 were generated from alternative start sites: secreted FGF3 from the AUG codon and nuclear FGF3 from the CUG codon. By investigating the relevant signals in more detail, we now find that in vectors expressing the wild-type cDNA sequence, the CUG codon is by far the major site of initiation. This would be consistent with the fact that the AUG codon is in a particularly unfavourable context for translation, according to the ribosome scanning model (Dixon *et al.*, 1989; Kozak, 1989, 1991a). More importantly, these findings indicate that FGF3 may be the first example of a protein that can be directed to either the nucleus or the secretory pathway by a balance between classical signals for protein sorting within the same polypeptide. We propose a mechanism by which the influence of the competing signals can be modulated by changing their relative positions and strengths.

Results

Translation of FGF3 from alternative initiation codons

We previously reported that the mouse *Fgf-3* gene has the capacity to encode products that are initiated from a CUG codon located 87 nucleotides upstream of the most 5′ AUG in the protein-coding frame (Figure 1a; Acland *et al.*, 1990). The findings were based on two cDNA constructs, designated 3.2 and 4.1, which differed only at their 5′ ends (Figure 1b). In the former, sequences around the first AUG were optimized for translation initiation by removing most of the upstream DNA and by altering the adjacent nucleotides to match the most favourable context described by Kozak (reviewed in Cavener and Ray, 1991; Kozak, 1991a). In rabbit reticulocyte lysates, this cDNA was translated into a 28.5 kDa polypeptide (Figure 2a) and studies on the biosynthesis and secretion of FGF3 have largely focused on the processed and glycosylated forms of this translation prouct (Kiefer *et al.*, 1991, 1993b). With the 4.1 cDNA, in which the sequences around and preceding the AUG were left unchanged, the predominant product was a 32 kDa species, designated

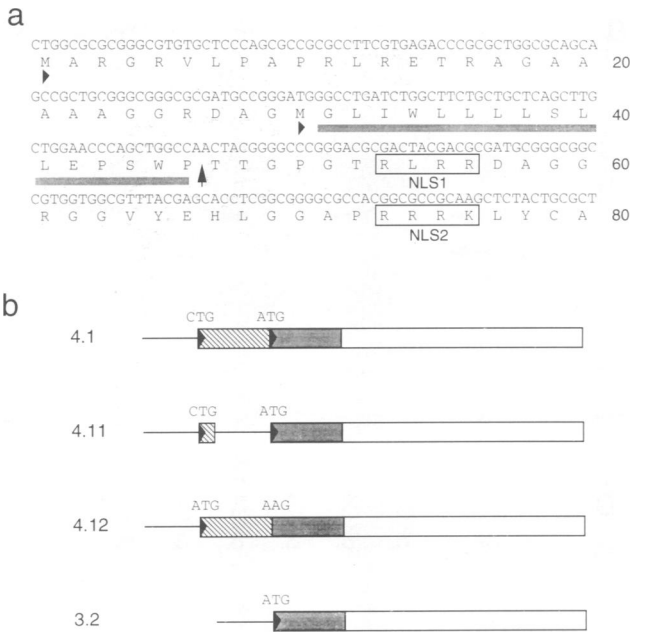

Fig. 1. Alternative initiation sites and the N-terminus of FGF3. (**a**) The DNA sequence of the N-terminus of FGF3 is shown with the predicted amino acid sequence in single letter code. The positions of the translation start sites are depicted as ▶. The hydrophobic core of the putative signal peptide is identified by the shaded line with the arrow indicating the signal peptide cleavage site. NLS1 and NLS2 are two basic regions implicated in nuclear localization. (**b**) The structure of four cDNAs used for expressing FGF3 in either reticulocyte lysates or in COS-1 cells. Boxes indicate sequences encoding FGF3 with non-coding sequences shown as a line. The CTG and ATG triplets that serve as initiation codons and their mutated forms are as indicated above each construct. ▶ represent translation start sites and the shaded box depicts the signal peptide.

p32, and only trace amounts of the AUG-initiated form were observed (Figure 2a).

The size of this product is consistent with initiation at the in-frame CUG, but to substantiate this conclusion we conducted N-terminal sequencing of p32 and generated mutants in which its expression was either impaired or enhanced. For example, in mutant 4.11, a termination codon was introduced 12 nucleotides downstream of the CUG by changing the arginine codon CGT to TGA (Figure 1). Although small amounts of p28.5 were obtained upon translation, there was no detectable 32 kDa product, indicating that the stop codon had blocked its synthesis (Figure 2a). Conversely, with mutant 4.12 only p32 was produced (Figure 2a). This mutant was optimized for initiation at the upstream site by changing the CUG to an AUG and adjusting the adjacent nucleotides to the Kozak consensus (GGCAUGG). Alternative initiation was also prevented by changing the downstream AUG to AAG (Figure 1b).

To facilitate the N-terminal sequencing, one additional mutant was constructed. Since methionine residues are frequently removed from newly synthesized proteins unless there is a large positively charged amino acid in the second position (Flinta *et al.*, 1986), the relevant alanine codon in 4.1 was changed to lysine (Figure 1a). When this contruct was translated in the presence of [^{35}S]methionine and [^{3}H]leucine, sequential Edman degradation of the resultant p32 revealed methionine in

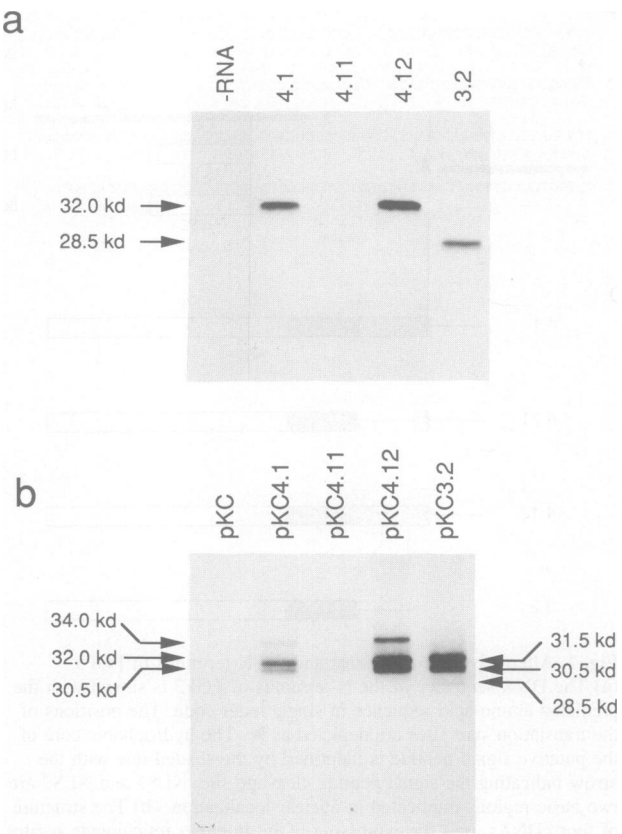

Fig. 2. The CUG codon is the major initiation site of FGF3 translation. (**a**) The wild-type and mutant FGF3 cDNAs described in Figure 1b were ligated into the pGEM3 vector and used to direct transcription with SP6 polymerase followed by translation in rabbit reticulocyte lysates. The labelled products were fractionated by SDS−PAGE in a 12.5% gel and detected by autoradiography. The sizes of the major products, indicated in kDa, were calculated relative to ^{14}C-labelled standards (Amersham). −RNA shows a control containing no input RNA. (**b**) The same cDNAs, as indicated, were cloned into the pKC vector and introduced into COS-1 cells by electroporation. After 48 h, the cultures were labelled with [^{35}S]methionine for 30 min and equivalent amounts of cell lysate were immunoprecipitated with a polyclonal antiserum against the C-terminus of FGF3. The precipitated proteins were fractionated by SDS−PAGE in a 12.5% gel and the FGF3-related products detected by autoradiography. The sizes and characterization of the various products have been described elsewhere (Acland *et al.*, 1990; Kiefer *et al.*, 1991).

fraction 1 with leucine in fractions 7 and 12, exactly as expected for initiation at the CUG codon (Figure 1a and data not shown). A similar experiment with the wild-type 4.1 construct yielded leucines in fractions 6 and 11, indicating that the terminal methionine had been removed. These data formally confirmed that translation was initiating at the CUG and that in this context it is read as a methionine codon.

CUG is the major initiation site for FGF3 in transfected cells

The existence of alternative initiation codons in the FGF3 sequence provided a mechanism for generating products with different subcellular fates. However, the *in vitro* translation data suggested that the CUG is the major initiation site for translation. It was therefore important to examine the efficiency of CUG-initiated translation relative to the proportions of nuclear and secreted forms of FGF3 in transfected cells. The cDNAs described above were transferred into an expression vector (pKC), based on the SV40 early promoter, and introduced into COS-1 cells by electroporation. After metabolic labelling with [^{35}S]methionine, the FGF3-related proteins were analysed by immunoprecipitation and SDS−PAGE. No FGF3-related proteins were detected in cells transfected with the vector alone (pKC), whereas the pKC3.2 vector, optimized for translation from the AUG, generated the expected gp31.5, gp30.5 and p28.5 products (Figure 2b). These correspond respectively to glycosylated and non-glycosylated forms of FGF3 before and after signal peptide cleavage, as previously described (Dixon *et al.*, 1989). With the vector containing the wild-type cDNA, pKC4.1, the predominant product was p32, as expected for initiation at the CUG codon (Figure 2b). Small amounts of 34 and 30.5 kDa species were also observed, which we interpret as glycosylated forms of p32 before and after cleavage. Thus, the 30.5 kDa form would be structurally equivalent to gp30.5 expressed from pKC3.2 (see below). The alternative explanation would be that some of these products are derived by initiation at the AUG codon. However, with the pKC4.11 mutant, where translation from the CUG is prematurely terminated and the AUG is in its normal context, only small amounts of the AUG-initiated products were detectable (Figure 2b). From the relative levels of the FGF3-related products expressed from pKC4.1 and pKC4.11, which differ by only two nucleotides, we conclude that initiation from the CUG is ~10-fold more efficient than initiation from the AUG codon. With pKC4.12, where the CUG has been changed to an AUG in an optimized context, even higher levels of the 34, 32 and 30.5 kDa products were obtained (Figure 2b).

CUG-initiated FGF3 is found in the nucleus/nucleoli and secretory pathway

Having established that the wild-type cDNA sequence is preferentially translated from the CUG, with little contribution from the AUG, it was essential to confirm that the nuclear and secreted forms of FGF3 were in fact derived from the same primary translation product. The distribution of FGF3 was therefore analysed by indirect immunofluorescence. In COS-1 cells expressing the wild-type pKC4.1 vector, the staining was predominantly nuclear and juxtanuclear (Figure 3a). As well as uniform staining of the nucleus, there was a significant concentration in the nucleoli. Since the juxtanuclear staining was superimposed on a more general reticular pattern, it can be taken as indicative of the secretory pathway with most of the signal concentrated in the Golgi complex (Figure 3a). In other work, we have clearly shown that the AUG-initiated forms of FGF3 expressed from pKC3.2 accumulate in the Golgi complex of transfected COS-1 cells (Kiefer *et al.*, 1993b) and an example of this type of staining is included in Figure 3b. To confirm that both patterns reflect the distribution of the CUG-initiated form of the protein, Figure 3c and d shows the corresponding immunofluorescence and phase contrast images for COS-1 cells transfected with pKC4.12, where translation from the upstream codon has been optimized and internal

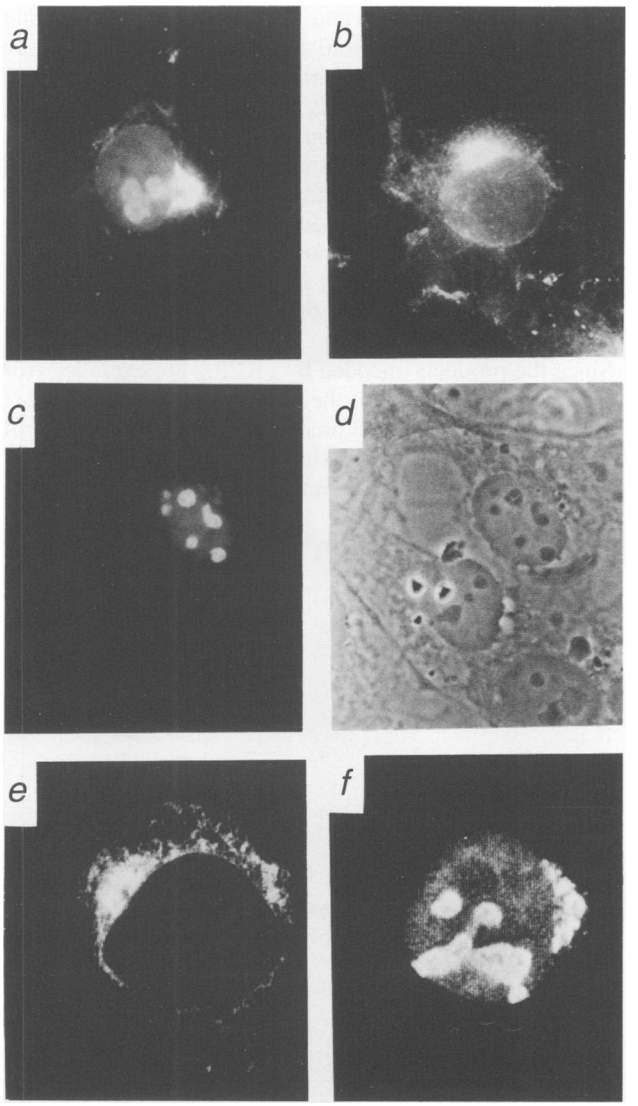

Fig. 3. Intracellular localization of FGF3 by immunofluorescence microsocopy. COS-1 cells transfected with pKC4.1 (**a**), pKC3.2 (**b** and **e**) and pKC4.12 (**c**, **d** and **f**) were grown on coverslips for 48 h and fixed in 4% paraformaldehyde. The coverslips were then stained with a rabbit polyclonal antibody against FGF3, except for panel (d) which shows the phase contrast picture of panel (c). The antibody complexes were visualized using goat anti-rabbit-Ig tagged with Texas red. In panels (e) and (f), the stained cells were examined by confocal microscopy with appropriate filters.

initiation precluded. Here, very distinctive nuclear and nucleolar staining was evident, as well as fluorescence indicative of the Golgi complex. Such conclusions were verified by generating confocal images of cells transfected with either pKC3.2, where only the AUG-initiated form was synthesized, or pKC4.12 (Figure 3e and f). Taken together, these immunofluorescence data demonstrate that products generated from the 32 kDa FGF3 precursor have a dual fate, entering both the nucleus and the secretory pathway of transfected cells.

CUG- and AUG-initiated forms of FGF3 are cleaved at the same site

As noted in Figure 2, pKC4.1-transfected COS-1 cells contained a significant amount of a 30.5 kDa product that co-migrated with the cleaved and glycosylated form of

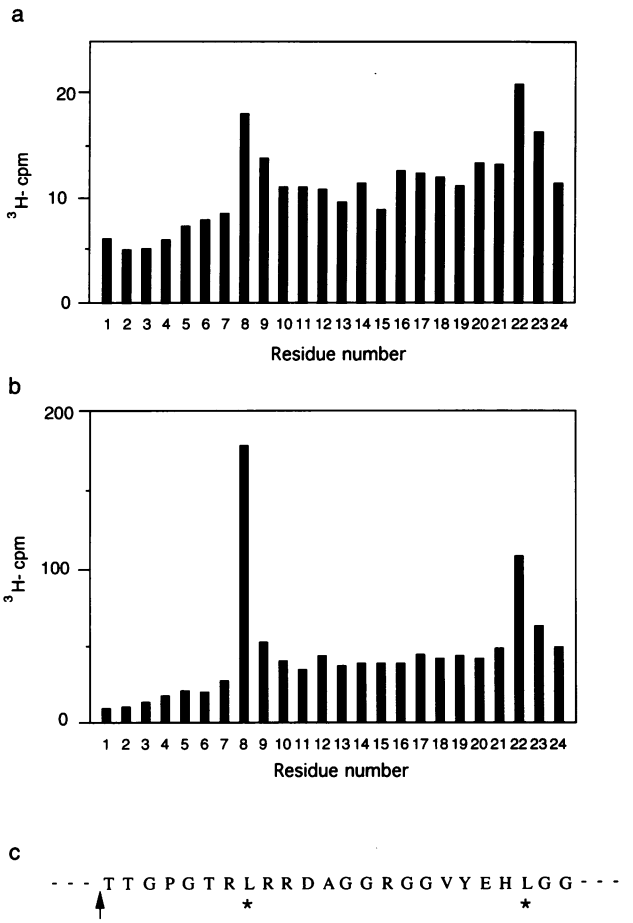

Fig. 4. N-terminal sequence analyses of processed forms of FGF3 initiated at CUG and AUG. COS-1 cells transfected with pKC4.12QMYC (**a**) or pKC3.2QMYC (**b**) were labelled with [^{3}H]leucine and [^{35}S]methionine and the MYC-tagged products were recovered by immunoprecipitation and SDS–PAGE. The proteins were electroblotted onto PVDF membranes and the cleaved and non-glycosylated 28.5 kDa products were located by autoradiography. The two proteins were subjected to sequential Edman degradation and the histograms plot the amount of [^{3}H]leucine in each of the recovered fractions. Leucine was detected at residues 3 and 17 from the N-terminus of both proteins. These correspond to residues 54 and 68 in the CUG-initiated form of FGF3 (equivalent to residues 25 and 39 in the AUG-initiated protein).

FGF3 encoded by pKC3.2. The most likely explanation is that p32 can be processed at the same signal peptide cleavage site, despite having the hydrophobic domain embedded within the protein. To address this possibility, the cleaved proteins derived from the CUG- or AUG-initiated precursors were compared by N-terminal sequencing. For practical reasons, the vectors were modified to improve the recovery of FGF3-related proteins from metabolically labelled cells. An epitope from the human MYC protein was appended to the C-terminus so that the products could be efficiently precipitated with a monoclonal antibody, the available FGF3 antisera being less effective for immunopreciptiation. Secondly, the single consensus site for asparagine-linked glycosylation was removed by changing the appropriate asparagine codon (AAC) to glutamine (CAG) as previously described (Acland *et al.*, 1990). By blocking carbohydrate addition, the products of interest were concentrated in a single band. Finally, the pKC4.12 vector was used in preference

to pKC4.1 to optimize translation from the upstream initiation site.

COS-1 cells transfected with this modified vector, pKC4.12QMYC, and the equivalent version for AUG-initiation, pKC3.2QMYC, were labelled with [^{35}S]-methionine and [^{3}H]leucine and cell extracts were immunoprecipitated with the monoclonal antibody 9E10 (Evan *et al.*, 1985). The MYC-tagged proteins were fractionated by SDS−PAGE, transferred to PVDF membranes and the cleaved and non-glycosylated 27.5 kDa products were recovered. Sequential Edman degradation revealed [^{3}H]leucine in fractions 8 and 22 in each case (Figure 4a and b). Thus, cleavage of the CUG- and AUG-initiated forms of FGF3 occurs at the same place, despite the differing extents of upstream sequence. By comparison with the predicted amino acid sequence (Figure 4c), this spacing of leucine residues was only compatible with cleavage between Pro46 and Thr47 (equivalent to residues 17 and 18 numbered from the AUG codon; see Figure 1). A computer-based algorithm based on known signal peptides indicates that this is the most favourable cleavage site in the N-terminal sequence of FGF3 (von Heijne, 1986). However, the CUG-initiated form of FGF3 presents a relatively unusual situation in that the hydrophobic core of the signal peptide (residues 32–41) is a considerable distance from the N-terminus.

The N-terminal domain of p32 is necessary for nuclear localization

Since the products encoded by pKC3.2 are excluded from the nucleus (Figure 3b), the additional 29 amino acids in the CUG-initiated FGF3 precursor presumably play a role in nuclear localization. Although this region does not contain a consensus NLS motif as defined for other nuclear

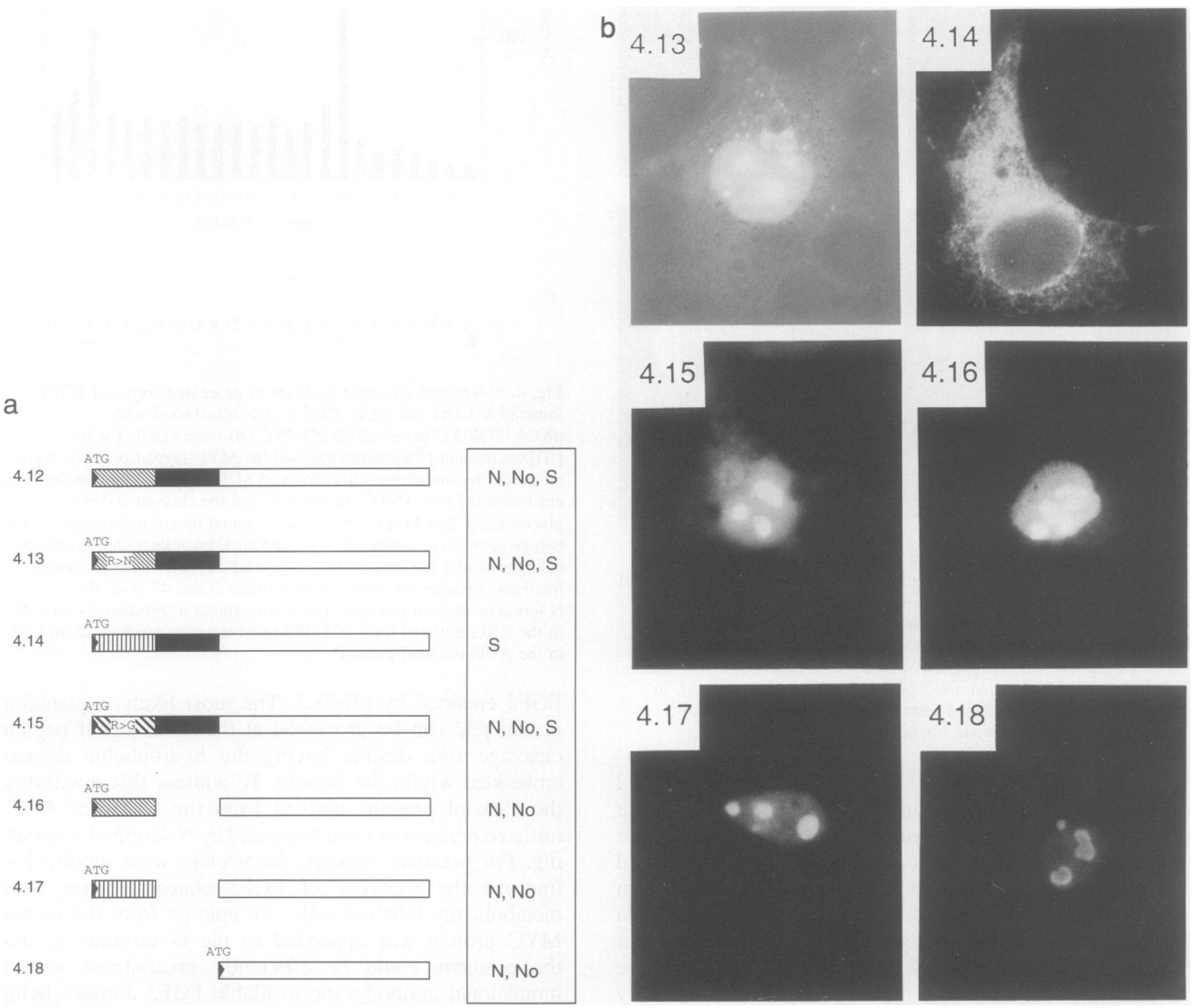

Fig. 5. Influence of N-terminal sequences on subcellular localization of FGF3. (**a**) A series of mutant cDNAs, all derivatives of the pKC4.12 plasmid, are depicted schematically using the same conventions as in Figure 1a. The signal peptide is represented by the shaded box with the N-terminal domain lightly striped. Vertical stripes identify the altered N-terminal domain generated by introducing a frameshift at codon 3 with a compensating mutation at codon 30. The bold stripes indicate the equivalent region from human FGF3. R>N and R>G refer to point mutations in the N-terminal regions designed to reduce the effect of potential nuclear localization signals in the N-terminal domain. The different plasmids were introduced into COS-1 cells by electroporation and 48 h later the subcellular distribution of FGF3-related products analysed by immunofluorescence. Representative examples of the staining are shown in (**b**) and the overall patterns are summarized alongside each cDNA in (a). N = nuclear, No = nucleolar and S = juxtanuclear and reticular staining indicative of the secretory pathway.

proteins, it does include an exact analogue of the sequence $RGRX_5R$, a suggested nuclear targeting motif for FGF2 (see Figure 1a and Bugler *et al.*, 1991). To investigate the influence of this motif and the N-terminal domain, a further series of mutants was constructed based on pKC4.12, which was optimized for translation of p32. COS-1 cells transfected with each plasmid were analysed by immunoprecipitation and immunoblotting, to verify the production and provenance of the mutant proteins (not shown), and by immunofluorescence, to determine their intracellular localization (Figure 5). Amino acid substitution in the $RGRX_5R$ motif had no discernible effect on the distribution of the products (pKC4.13, Figure 5), suggesting that it does not have a major role in nuclear localization of FGF3. In contrast, changing the reading frame of the N-terminal region, by introducing two nucleotides in codon 3 and making a compensating deletion from codon 30, had a dramatic effect and completely abolished nuclear localization (pKC4.14, Figure 5). Since the frameshift changed the protein sequence but conserved any essential secondary structure features in the nucleotide sequence, the results imply that the amino acid composition of this domain is critical. In this context, it is interesting that the corresponding segment of human FGF3 was able to substitute functionally for the N-terminal domain of mouse FGF3 (pKC4.15, Figure 5), despite being <30% identical in the primary sequence. In the example shown in Figure 5, an arginine in the human sequence had been changed to glycine to test whether the short cluster of basic residues in this region of human FGF3 had any influence, but there is no equivalent of the $RGRX_5R$ motif. Taken together, these findings imply that the N-terminal region contributes significantly to nuclear localization, but that its influence is mediated by a property other than a linear sequence motif.

The pKC4.13, 4.14 and 4.15 mutants all retained the hydrophobic core of the signal peptide, which was presumably competing for nuclear localization by directing some of the products into the secretory pathway. By deleting residues 30–46 of p32, encompassing the hydrophobic domain, this competition was ablated and the FGF3-related products were found to be exclusively nuclear and nucleolar (pKC4.16, Figure 5). Moreover, immunoprecipitation and immunoblotting verified that the major product encoded by this mutant had not undergone the cleavage and glycosylation expected for a secreted protein (not shown), confirming that the hydrophobic domain is absolutely necessary for translocation into the ER. Significantly, an identical result was obtained when the same residues were deleted from pKC4.14, where the influence of the N-terminal domain was abrogated by a frameshift (pKC4.17, Figure 5). This implies that the N-terminal region is not the only determinant for nuclear localization. Altering the sequences preceding the signal peptide cleavage site had generated an exclusively nuclear protein, suggesting that sequences downstream of this site contribute significantly to nuclear localization. Indeed, a protein in which all 46 residues preceding the cleavage site were removed and an AUG initiation codon introduced at residue 47 was efficiently translocated to the nucleus (pKC4.18, Figure 5).

Identifying nuclear localization signals in FGF3

With few exceptions, NLS comprise short, highly basic domains, typically flanked by one or more proline residues. As previously noted, the sequence of mouse FGF3 downstream of the signal peptide cleavage site contains several clusters of arginine residues and at least five regions could be considered candidate NLS motifs. The most persuasive (RLRR and RRRK; see Figure 1a) are in the N-terminal half of the protein and have been designated NLS1 and NLS2 respectively. To assess the influence of these sequences on the subcellular distribution of FGF3, additional mutations were introduced into the pKC4.12 vector, which encodes both nuclear and secreted products (see Figure 3c). The nature of these mutations and their effect on the location of FGF3 are summarized in Figure 6. Removal of both NLS1 and NLS2 effectively abolished

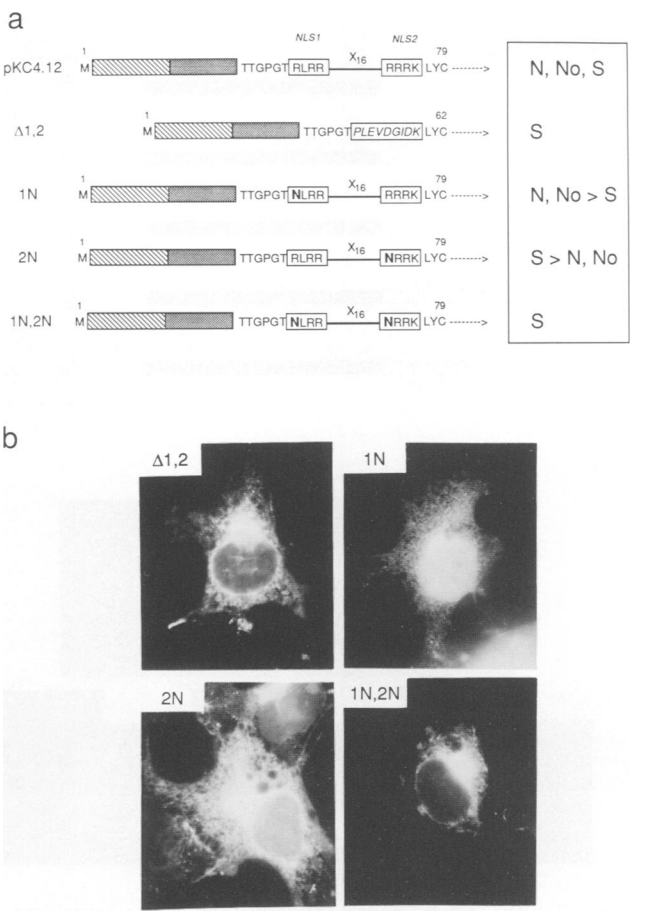

Fig. 6. Mutations affecting nuclear localization of FGF3. **(a)** The mutations introduced into the pKC4.12 plasmid are depicted schematically with the signal peptide represented by the striped box and the candidate nuclear localization signals NLS1 and NLS2 boxed, as in Figure 1a. In Δ1,2, both NLS1 and NLS2 have been deleted and nine random amino acids inserted. 1N has an arginine to asparagine mutation in NLS1, whereas 2N has an analogous mutation in NLS2. The 1N,2N plasmid has the mutation in both NLS motifs. The different plasmids were introduced into COS-1 cells by electroporation and 48 h later the subcellular distribution of FGF3-related products analysed by immunofluorescence. Representative examples of the staining are shown in **(b)** and the overall patterns are summarized alongside each cDNA. N = nuclear, No = nucleolar and S = secretory pathway.

nuclear localization (Δ1,2 in Figure 6). In the constuct shown, the region encompassing NLS1 and NLS2 was replaced with a stretch of nine amino acids that are predicted to have a random conformation. In the 1N mutant, the first arginine in NLS1 was changed to asparagine (Figure 6a). An analogous minimal change has been shown to diminish the effects of established NLS motifs in other proteins. In the context of FGF3, this mutation clearly reduced, but did not abolish, the accumulation of products in the nucleus (Figure 6b). A more dramatic effect was observed when a similar mutation was introduced into NLS2, but FGF3 was still detectable in the nucleus (2N in Figure 6). However, when the two mutations were combined (1N,2N), virtually all of the products were in the secretory pathway. Point mutations in the other candidate NLS motifs had little effect on the nuclear import of FGF3 (data not shown).

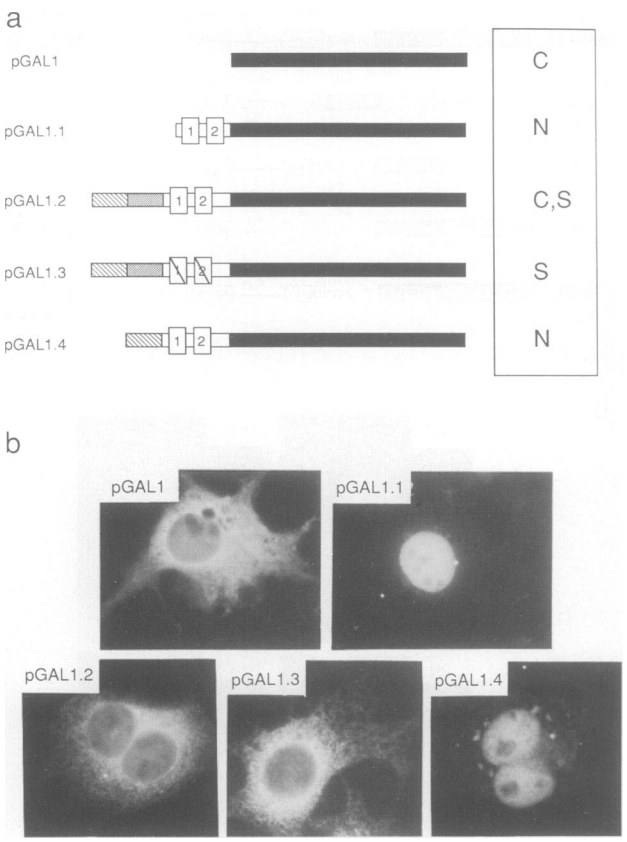

Fig. 7. Effect of FGF3 nuclear localization signals on a heterologous protein. Segments of the N-teminal region of FGF3, encompassing the first 79 residues, were fused to most of the coding domain of β-galactosidase (lacking seven amino acids) as depicted schematically in (**a**). The black bar represents β-galactosidase and the conventions used for FGF3 sequences are as described in previous figures. The pGAL1.1 plasmid contained only the candidate nuclear localization signals NLS2 and NLS2 (boxes 1 and 2), whereas pGAL1.2 included the N-terminal domain (striped box) and the signal peptide (shaded box), as well as NLS1 and NLS2. The pGAL1.3 contruct incorporated arginine to asparagine mutations in both NLS1 and NLS2 (analogous to 1N,2N in Figure 6). In pGAL1.4, the core of the signal peptide (residues 31–46) was deleted. (**b**) Following transfection of COS-1 cells with these plasmids, the location of β-galactosidase was determined by staining with a monoclonal antibody followed by a secondary antibody conjugated to Texas red.

NLS1 and NLS2 function as a bipartite nuclear localization signal

A more stringent test for the function of a nuclear localization signal is whether it is sufficient to direct a heterologous cytoplasmic protein into the nucleus. To explore this possibility, plasmids were constructed in which segments of the FGF3 coding sequences were fused with those for bacterial β-galactosidase (Figure 7). The latter offered a number of advantages as a marker protein: it is normally cytoplasmic, it is a large enzyme which is unlikely to enter the nucleus by diffusion and it can be readily detected with specific antibodies. For example, Figure 7b shows the appearance of COS-1 cells transfected with the β-galactosidase expression vector pGAL1 and stained with a monoclonal antibody. The product, β-galactosidase lacking seven residues from the N-terminus, was uniformly distributed in the cytoplasm and excluded from the nucleus.

When the NLS1 and NLS2 domains from FGF3 were fused to the N-terminus of β-galactosidase (pGAL1.1) the resultant products accumulated in the nucleus (Figure 7b). Thus, the combined effects of NLS1 and NLS2 are sufficient to direct nuclear localization on a heterologous protein, in agreement with the deletion and mutation analyses described in Figure 6. However, when the entire N-terminus of FGF3 was fused to β-galactosidase, including the signal peptide and both NLS1 and NLS2 (pGAL1.2), the resultant products showed a cytoplasmic and reticular pattern of staining, suggesting that a significant fraction of the fusion protein had been directed into the ER (Figure 7b). In this context, therefore, the core of the signal sequence in FGF3 appears able to override the effects of the nuclear localization signals, but may remain partially compromised as a secretory signal. Thus, the properties of the signal peptide became more convincing when the minimal arginine to asparagine mutations were introduced into NLS1 and NLS2 (pGAL1.3 in Figure 7). Conversely, upon removing the hydrophobic domain (residues 30–41) the combined effects of the N-terminus, NLS1 and NLS2 confer nuclear localization (pGAL1.4 in Figure 7).

Competition between secretory and nuclear localization signals

These effects on a heterologous protein suggested that the fate of FGF3 was determined by a balance between the secretory and nuclear localization signals. In their normal context, the signal peptide cleavage site and NLS1 are separated by only six residues and, since both these signals comprise linear motifs that are interpreted by the corresponding recognition particles, we considered the possibility that the proteins compete for binding. We therefore altered the distance between the two signals, by introducing a variable length spacer of random structure, and observed the effects on protein distribution. Figure 8a shows the normal disposition of the signal peptide, NLS1 and NLS2 in the pKC4.12 vector. As shown in Figure 3c, the resultant products are distributed equally between the nucleus and secretory pathway. Inserting 7, 10, 13 or 14 amino acids between the proline and glycine (Figure 8a) increased the proportion of products in the

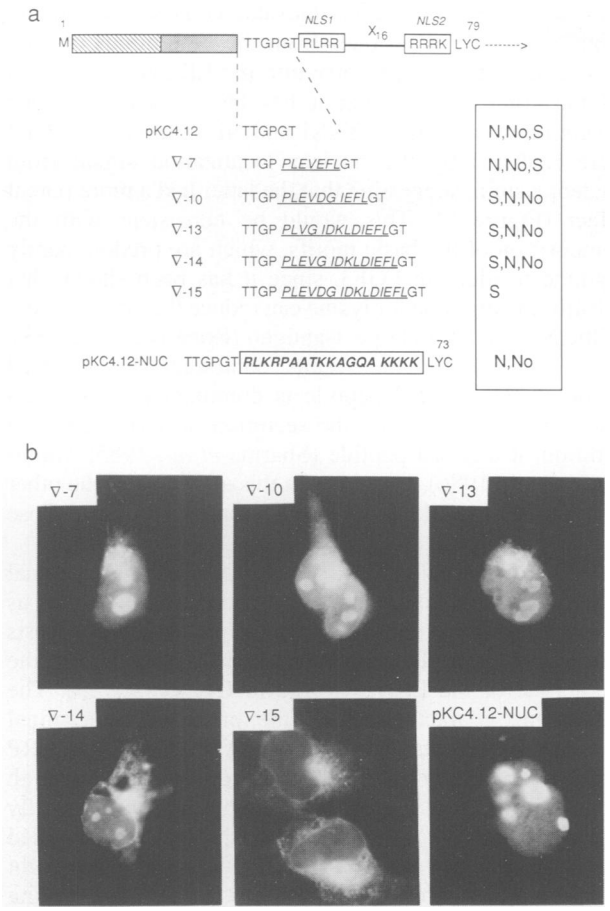

Fig. 8. Competition between secretory and nuclear localization signals in FGF3. (**a**) The structure of the pKC4.12 plasmid is shown schematically to indicate the six amino acids separating the signal peptide (shaded box) and NLS1 (RLRR). Variants of pKC4.12 were generated, designated by the ∇ symbol, in which this distance was increased by 7, 10, 13, 14 or 15 residues as indicated. In pKC4.12-NUC, the region of FGF3 encompassing NLS1 and NLS2 was replaced with the bipartite nuclear localization signal from nucleoplasmin. The different plasmids were introduced into COS-1 cells by electroporation and 48 h later the subcellular distribution of FGF3-related products analysed by immunofluorescence. Representative examples of the staining are shown in (**b**) and the overall patterns are summarized as before. N = nuclear, No = nucleolar and S = secretory pathway, with the order of these symbols signifying the degree of staining in each compartment. Although difficult to demonstrate photographically, the staining of the nucleus and nucleoli decreased steadily as the size of the insert increased.

secretory pathway, but did not completely abolish nuclear localization (Figures 8b). Significantly, the insertion of 15 amino acids tilted the balance in favour of the signal peptide and virtually all the staining was in the ER and Golgi complex (Figure 8b). Thus, increasing the distance between the competing motifs favoured recognition of the signal peptide over the NLS.

Comparison of the NLS1 and NLS2 motifs with known bipartite nuclear localization signals suggested that the FGF3 signal may be relatively weak, since the basic residues are predominantly arginine. More potent NLS motifs generally have a higher proportion of lysine residues. We therefore asked whether substituting the bipartite NLS domain from nucleoplasmin for the corresponding region of FGF3 would change the relative distribution of the protein (Robbins *et al.*, 1991). The resultant products were exclusively nuclear (Figure 8b), despite the presence of the FGF3 signal peptide. These data suggest that the more effective the nuclear localization domain, the more readily it competes with the FGF3 signal peptide in determining the subcellular fate of the protein.

Discussion

The ribosome scanning model for the initiation of protein synthesis predicts that translation will generally commence at the first AUG codon encountered in the RNA, provided that it is in a favourable context (Kozak, 1987, 1991b; reviewed in Cavener and Ray, 1991). This context has been defined by surveying known initiation sites and by site-directed mutagenesis of nucleotides adjacent to the AUG codon. While these predictions appear to hold true for the vast majority of eukaryotic genes, a number of exceptions have been reported in which initiation can occur at non-AUG codons, most commonly CUG. Examples include human *MYC, FGF2* and *LTK* and the mouse *Pim-1* and *Hck* genes, but in almost every case the CUG codon is viewed as an alternative to the major initiation event that occurs at a subsequent AUG codon (reviewed in Kozak, 1991b). The data we describe challenge this assumption, by showing that in the case of the mouse *Fgf-3* gene, the upstream CUG is the preferred site of initiation, both *in vitro* and *in vivo* (Figure 2). We have also established that in the context of the *Fgf-3* gene, the CUG is interpreted as a methionine rather than a leucine codon (data not shown).

Comparison of the 4.1 and 4.11 constructs, which differ at only two nucleotides, indicated a 10-fold preference for initiation at the CUG and a number of factors might contribute to this preference. The first is that the CUG in *Fgf-3* is in a favourable context for initiation, at least with respect to nucleotides in the −3 and +4 positions. Secondly, the sequences downstream of the CUG have the potential to form a secondary structure that may impede the progress of ribosomes and facilitate initiation at the CUG (Kozak, 1990). A similar feature occurs in other examples of CUG-initiated genes. Thirdly, the in-frame AUG codon in the *Fgf-3* gene is in a particularly unfavourable context for initiation, since it is immediately preceded by an alternative AUG in the +1 reading frame (Dixon *et al.*, 1989). The potential for ribosomes to initiate in the wrong frame (the resultant peptide would terminate after five residues) has been shown to interfere with translation of FGF3, and this feature is conserved in the mouse, human and *Xenopus* homologues (Brookes *et al.*, 1989; Dixon *et al.*, 1989; Kiefer *et al.*, 1993a).

On the other hand, the conservation of the in-frame AUG codon suggests that it may still function and it is interesting to recall that the mouse *Fgf-3* gene can be expressed from three distinct promoters, generating transcripts of different length (Mansour and Martin, 1988; Smith *et al.*, 1988; Grinberg *et al.*, 1991). It is therefore possible that the distance between the cap site and the CUG could influence the efficiency with which it can function as an initiation codon. In the cDNA used to construct the pKC4.1-based plasmids, at least 300 nucleotides were retained upstream of the CUG, reflecting

transcription from the P2 promoter of the mouse *Fgf-3* gene. Although P2 is less active in some cells, it is the dominant promoter in mammary tumours in which *Fgf-3* is activated by MMTV integration. In other cells, however, transcripts originating from the so-called P3 promoter would have only 12 nucleotides preceding the CUG codon and this may not be enough for effective ribosome loading. If such transcripts promote initiation from the AUG, then the mouse *Fgf-3* gene may have evolved in such a way that proteins with different fates are generated by alternative start sites for transcription, rather than translation.

Having established that intiation occurs predominantly at the CUG, it was important to examine what effect the upstream initiation and the additional 29 amino acids might have on the processing, glycosylation and eventual fate of the translation product. The data in Figures 2 and 4 indicated that at least a portion of the p32 precursor is cleaved adjacent to the hydrophobic signal peptide, despite its internal location, and that asparagine-linked glycosylation is unaffected. Thus, upstream initiation does not prevent the production of the secreted forms of FGF3 that were described in earlier reports (Dixon *et al.*, 1989; Kiefer *et al.*, 1991, 1993b). It does, however, divert a significant proportion of the product to the nucleus (Figure 3). *Fgf-3* may therefore be the first example of a gene that specifies both nuclear and secreted products from the same precursor, a feat that is more commonly achieved via alternative initiation or splicing (Maher *et al.*, 1989; Spence *et al.*, 1989; Lock *et al.*, 1991).

In terms of nuclear localization signals, the obvious candidates are located within the body of the mature FGF3, the first of which is only six residues downstream of the signal peptide cleavage site. The mutational analyses described in Figure 5 and the ability to translocate β-galactosidase to the nucleus (Figure 7) clearly established NLS1 and NLS2 as a bipartite nuclear localization signal, reminiscent of the well-characterized motifs in other proteins (reviewed in Dingwall and Laskey, 1991; Garcia-Bustos *et al.*, 1991). Since the composition and relative spacing of NLS1 and NLS2 is fairly typical of such signals, we have not attempted to change the distance between them. Such experiments with other bipartite signals have suggested that the spacing between the basic motifs can be varied considerably without affecting their function (Robbins *et al.*, 1991). However, it was also clear that in FGF3 the effects of NLS1 and NLS2 could be augmented by the presence of the 29 N-terminal residues in p32 (Figure 5). Thus, changing the amino acid sequence by shifting the reading frame drastically reduced nuclear uptake. On the other hand, the equivalent region of human FGF3 was able to substitute for that of the mouse protein, despite being <30% conserved at the primary sequence level. It therefore seems likely that the function of this region reflects some general feature of the polypeptide, rather than a linear sequence motif. In this context, we note that the N-termini of both mouse and human FGF3 are markedly more hydrophilic than the sequence generated by the frameshift mutation.

A plausible explanation for these effects would be that the N-terminal leader region can interfere with the function of the hydrophobic domain of the signal peptide. Although the configuration in FGF3 is unusual, with the hydrophobic motif located >30 residues from the N-terminus of p32, it is clear that it directs entry into the ER and operates in all the contexts in which it has been tested. The one exception was when the NLS1 and NLS2 motifs of FGF3 were replaced by the nuclear localization signal from nucleoplasmin, suggesting that the latter had a more potent effect (Figure 8). This would be consistent with the composition of the basic motifs, which are predominantly arginine residues in FGF3, since it has been shown that substituting arginine for lysine can reduce the effectiveness of the NLS in SV40 large T-antigen (Roberts *et al.*, 1987). More significantly, the results contradict the accepted wisdom that a signal peptide is dominant over an NLS motif, an example being the secretion of T-antigen upon addition of a signal peptide (Sharma *et al.*, 1985; Munro and Pelham, 1986). However, in this and most of the other systems analysed the hydrophobic domain is located close to the N-terminus.

Membrane translocation requires the binding of a signal recognition particle (SRP) to the hydrophobic domain as it emerges from the ribosome, a process that arrests translation until the complex becomes located in the membrane of the ER (reviewed in Gilmore, 1993). The length and characterisitcs of the preceding N-terminal sequence in p32 may therefore affect the binding of SRP and the accompanying translational pause. Thus, although an AUG-initiated form of FGF3 would proceed rapidly and exclusively into the secretory pathway, as exemplified by the results with pKC3.2, the CUG-initiated form might be less efficiently recognized by SRP. As well as delaying translocation to the ER, impaired binding of SRP could well facilitate the association of the NSB complex to NLS1 and NLS2. The net effect would be competitive binding of the two recognition complexes, whose target sequences in FGF3 are separated by only six residues. We tested this possibility by altering the relative positions of the competing signals, arguing that increasing the distance between them might enable the SRP to function more effectively. This was indeed the case, as shown in Figure 8. Insertion of additional amino acids steadily decreased the proportion of FGF3 in the nucleus, reaching a cut-off point with an insert of 15 residues. By separating the competing motifs in this way, the dominance of the signal peptide was restored and the products were entirely in the secretory pathway.

Competition for binding also implied that it should be possible to shift the balance in the other direction by altering the affinity for the different sequence motifs. Thus, replacing NLS1 and NLS2 in FGF3 with the bipartite nuclear localization signal from nucleoplasmin resulted in a protein in which NSB binding was dominant (Figure 8). Such a possibility may have been overlooked in other proteins, where the hydrophobic domain of the signal peptide is in a more favourable location. We presume that in the CUG-initiated form of FGF3, the reduced efficiency of the signal peptide makes the competing effects more apparent. In a wider context, the data we describe for FGF3 clearly show that it is possible for proteins to have dual fates determined by competition for binding to linear sequence motifs. Since the effects of these otherwise conventional signals can be finely balanced by their

relative positions and contexts, our findings have more general implications for other proteins that carry such motifs.

Materials and methods

Cell culture

COS-1 cells were maintained in Dulbecco's modified Eagles medium (DMEM) containing 10% fetal calf serum (FCS) and passaged once a week at a ratio of 1:10. For transient DNA transfections, 20 μg purified plasmid DNA was introduced into 5×10^5 COS-1 cells by electroporation (450 V/250 μF) using a Bio-Rad Gene-Pulser. Between 48 and 72 h after transfection the cells were either labelled with [^{35}S]methionine and [^{35}S]cysteine and harvested directly for protein analysis or analysed by indirect immunofluorescence (see below).

Plasmid constructions

Vectors expressing the 4.1 and 3.2 FGF3 cDNA constructs, either in an SV40-based expression vector (pKC) or an *in vitro* translation vector (pGEM3), have been described previously (Dixon *et al.*, 1989; Acland *et al.*, 1990). Where necessary, variant cDNAs were cloned into both vectors. For pGEM plasmids, the cDNAs were orientated for transcription with SP6 RNA polymerase. The variants are described in terms of their DNA sequences, rather than the corresponding RNA.

The 4.12 cDNA insert was based on 4.1, but the sequences surrounding the upstream CTG were changed from GCCGGC**CTG**G to GCCGCC**ATG**G and the internal ATG 87–89 nucleotides downstream were changed to AAG by site-directed mutagenesis (this plasmid was kindly provided by F.Fuller-Pace). The derivative plasmids pKC3.2QMYC and pKC4.12QMYC were obtained by changing the asparagine codon (residue 65 in the 3.2 construct) to glutamine (AAC to CAG) and using two complimentary oligonucleotides to replace the last 10 amino acids of FGF3 with those of a MYC epitope (EQELISEEDL). These oligonucloetides were substituted for sequences between the unique *Acc*I site near the C-terminus of FGF3 and an *Eco*RI site in the non-coding region. The 4.11 cDNA was generated from 4.1 by PCR amplification with a mis-matched oligonucleotide primer. The 5′ primer changed codon 5 from CGT to TGA, introducing a translational stop, and included nine nucleotides upstream of the CTG, containing a natural *Nae*I site. The 3′ primer was based on the 12 nucleotides following the signal peptide cleavage site, which included a natural *Apa*I site (Figure 1a). This enabled the PCR product to be substituted for sequences in 4.1 via these sites.

The 4.13–4.17 constructs were all derived from pKC4.12 by substituting fragments generated by PCR amplification. In 4.13, the 5′ primer introduced a unique *Xba*I site upstream of the start codon and changed codon 5 from arginine (CGT) to asparagine (AAT). The 3′ primer was based on the *Apa*I site as above. In 4.14, a frameshift was introduced by incorporating two extra nucleotides between codons 2 and 3 in the 5′ primer. Making a compensating deletion at the ATG codon (nucleotides 89 and 90) in the 3′ primer generated a novel *Stu*I site. In a separate PCR, this two nucleotide deletion was used to introduce a *Stu*I site in the FGF3 sequence in 4.12. The frameshifted product was then inserted into pKC4.12 by exploiting the novel *Stu*I site. In the 4.15 construct, a similar strategy was used to replace the N-terminal sequences in 4.12 with the corresponding region of human FGF3.

The deletion clone pKC4.16 was constructed by amplifying the N-terminal domain with a 3′ primer that extended upstream from the ATG codon and incorporated the 12 nucleotides and *Apa*I site distal to the signal peptide cleavage site. This effectively removed the core of the signal peptide (residues 31–46) by fusing the N-terminal domain to the remainder of FGF3 via the *Apa*I site. To generate pKC4.17, the 3′ overhang of the *Apa*I site was removed using T4 DNA polymerase. The body of the FGF3 sequences were then blunt-end ligated to the *Xba*I–*Stu*I fragment containing the 5′ end of the frameshift mutant pKC4.14. This deleted the signal peptide sequence and corrected the reading frame to encode the remainder of FGF3. The 4.18 mutant was created by filling in the *Nco*I site of pKC3.2 using T4 polymerase and ligating the blunt end to the natural *Sma*I site just downstream of the region encoding the signal peptide.

Mutations in the putative nuclear localization signals were also generated by PCR amplification using mis-matched oligonucleotides, changing either Arg53 (1N) or Arg73 (2N) or both (1N,2N) to asparagine. To produce the insertion mutants, the sequences between the *Apa*I and *Sma*I sites in FGF3 were replaced with an *Apa*I–*Sma*I segment from the multicloning site of the Bluescript KSII plasmid vector, introducing 15 additional in-frame amino acids. Further modifications were introduced using the sites within the polylinker to produce inserts of 7, 10, 13 and 14 residues, as illustrated in Figure 8. To create Δ1,2, the sequence between codons 35 and 77 was replaced by a linker encoding the nine amino acids illustrated in Figure 6. To construct pKC4.12.NUC, a *Hin*dIII site was introduced into pKC4.12 adjacent to the encoded sequence LYC by site-directed mutagenesis. This did not change the coding potential, but allowed the substitution of sequences between the *Hin*dIII site and the upstream *Apa*I site with a pair of complementary oligonucleotides encoding the nucleoplasmin bipartite NLS, as shown in Figure 8.

The β-galactosidase–FGF3 fusion proteins were based on the expression plasmid pKC4 using the β-galactosidase gene from pJ3-β-Gal (kindly provided by Hartmut Land). To construct pGAL1, PCR was used to amplify the N-terminal domain of β-galactosidase (residues 7–278) and to introduce a unique *Xho*I site in the 5′ primer. Segments of FGF3 were fused to β-galactosidase through this site. In pGAL1.1, for example, two complementary oligonucleotides containing half *Xba*I and *Xho*I sites at the 5′ and 3′ ends were used to duplicate the FGF3 sequences between codons 20 and 48 (GPG...RKL) and add a methionine codon to the N-terminus. A derivative of pKC4.12 was also available in which a single point mutation at codon 105 (numbered from the CTG) had generated a new *Xho*I site, without changing the amino acid sequence (provided by F.Fuller-Pace). This allowed the fusion of 105 N-terminal codons of pKC4.12 to the body of the β-galactosidase sequence, generating the plasmid designated pGAL1.2 in Figure 7. pGAL1.3 was created by replacing an *Apa*I–*Not*I fragment of pGAL1.2 with the corresponding region of the 1N,2N form of pKC4.12. In pGAL1.4, the *Xba*I–*Not*I fragment of pGAL1.2 was replaced by the corresponding fragment of 4.16.

Immunoprecipitation, immunoblotting and in vitro translation

Transfected cell cultures were incubated for 30 min in serum-free DMEM lacking cysteine and methionine (labelling medium) then labelled for 30 min with 500 μCi [^{35}S]methionine and [^{35}S]cysteine (ExpreSS Labeling Mix, New England Nuclear) in 2 ml labelling medium. The cells were lysed in 2 ml extraction buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate, 0.02% sodium azide, with 100 μg phenylmethylsulfonyl fluoride and 10 mg/μl aprotinin) at 4°C for 30 min. The samples were then centrifuged for 15 min in a microfuge and the supernatants subjected to immunoprecipitation with polyclonal antiserum specific for the C-terminus of mouse FGF3 (5 μl antiserum per ml cell extract). After 1 h at 4°C, 100 μl of a 25% suspension of protein A–Sepharose (Pharmacia) was added and the immune complexes were collected by centrifugation. The precipitates were washed twice in NET buffer (150 mM NaCl, 1 mM EDTA, 50 mM Tris–HCl, pH 7.5) containing 0.1% NP-40 and 0.25% gelatin and once in 10 mM Tris–HCl, pH 7.5, 0.1% NP-40. The washed complexes were boiled for 5 min in dissociation buffer and analysed by SDS–PAGE in a 15% gel. After fixing in 10% acetic acid and 50% methanol and treatment with AMPLIFY as described by the manufacturer (Amersham), the ^{35}S-labelled products were detected by fluorography.

In some experiments, extracts from unlabelled cells were fractionated directly by SDS–PAGE in 12.5 or 15% gels, transferred to nitrocellulose membranes (Schleicher and Schuell) and then processed with rabbit polyclonal antibody to the C-terminus of mouse FGF3. Immune complexes were visualized with ^{125}I-labelled protein A (Amersham) and autoradiography as described previously (Kiefer *et al.*, 1991).

Several of the cDNAs used in transfection experiments were transferred into the vector pGEM3 to allow the synthesis of cRNA using SP6 polymerase. Aliquots of each cRNA (5 μg) were translated for 1 h in 50 μl rabbit reticulocyte lysate in the presence of labelled precursors after which 4 μl of each reaction was fractionated by SDS–PAGE in a 12.5% gel and the labelled products visualized directly by autoradiography.

Immunofluorescence

COS-1 cells grown on glass coverslips were transfected with the appropriate plasmids and 48 h later the cells were fixed in 4% paraformaldehyde in PBS for 20 min. The cells were then permeabilized with 0.2% Triton X-100 for 4 min and treated with 3% bovine serum albumin in PBS (3% BSA–PBS) to block non-specific binding of the antibodies. The coverslips were then exposed to primary antibodies and fluorescently labelled secondary antibodies diluted in 3% BSA–PBS. After washing, the stained cells were mounted in 90% glycerol containing *p*-phenylenedi-

amine and viewed with a 100× oil immersion lens on a Zeiss microscope equipped with the appropriate barrier filters for Texas red optics. Confocal images were taken on a Zeiss confocal laser scan microscope.

Microsequencing

For N-terminal sequence analyses, transfected COS-1 cells were labelled for 8 h with 250 μCi/ml [^{3}H]leucine (120–190 Ci/mmol; Amersham) and with 200 μCi/ml [^{35}S]methionine (1000 Ci/mmol; New England Nuclear) and immunoprecipitated with the monoclonal antibody 9E10 (Evan *et al.*, 1985), The FGF3-related proteins were fractionated by SDS−PAGE, electroblotted onto polyvinylidene difluoride (PVDF) membranes (Millipore) and the segments of the membrane containing the relevant labelled proteins were identified by autoradiography. The radioactive proteins were then covalently immobilized on the PVDF membrane by treatment with poly(allylamine) and 1,4-phenylenediisothiocyanate (Pappin *et al.*, 1990) and sequenced for 25 or 30 cycles using a MilliGen 6600 solid-phase sequencer. The presence of [^{3}H]leucine and [^{35}S]methionine in each cycle fraction was determined by liquid scintillation counting.

Acknowledgements

We would like to thank Drs Colin Dingwall and Graham Warren for their useful comments and suggestions on the manuscript. We are also grateful to Anna-Maria Florence for technical help in the preparation of plasmid DNAs.

References

Acland,P., Dixon,M., Peters,G. and Dickson,C. (1990) *Nature*, **343**, 662–665.
Akey,C. and Goldfarb,D. (1989) *J. Cell. Biol.*, **109**, 971–982.
Brookes,S., Smith,R., Casey,G., Dickson,C. and Peters,G. (1989) *Oncogene*, **4**, 429–436.
Bugler,B., Amalric,F. and Prats,H. (1991) *Mol. Cell. Biol.*, **11**, 573–577.
Cavener,D. and Ray,S. (1991) *Nucleic Acids Res.*, **19**, 3185–3192.
Dingwall,C. and Laskey,A. (1991) *Trends Biochem. Sci.*, **16**, 478−481.
Dixon,M., Deed,R., Acland,P., Moore,R., Whyte,A., Peters,G. and Dickson,C. (1989) *Mol. Cell. Biol.*, **9**, 4896–4902.
Dworetzky,S., Lanford,R. and Feldherr,C. (1988) *J. Cell. Biol*, **107**, 1279–1287.
Evan,G., Lewis,G., Ramsey,G. and Bishop,M. (1985) *Mol. Cell. Biol.*, **5**, 3610–3616.
Flinta,C., Persson,B., Jornvall,H. and von Heijne,G. (1986) *Eur. J. Biochem.*, **154**, 193–196.
Florkiewicz,R. and Sommer,A. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 3978–3981.
Garcia-Bustos,J., Heitman,J. and Hall,M.N. (1991) *Biochim. Biophys. Acta*, **1071**, 83–101.
Gilmore,R. (1993) *Cell*, **75**, 589–592.
Grinberg,D., Thurlow,J., Watson,R., Smith,R., Peters,G. and Dickson,C. (1991) *Cell Growth Different.*, **2**, 137–143.
Hann,S., King,M., Bentley,D., Anderson,C. and Eisenman,R. (1988) *Cell*, **52**, 185–195.
Kalderon,D., Roberts,B., Richardson,W. and Smith,A. (1984) *Cell*, **39**, 499–509.
Kiefer,P., Peters,G. and Dickson,C. (1991) *Mol. Cell. Biol.*, **11**, 5929–5936.
Kiefer,P., Mathieu,M., Close,M., Peters,G. and Dickson,C. (1993a) *EMBO J.*, **12**, 4159–4168.
Kiefer,P., Peters,G. and Dickson,C. (1993b) *Mol. Cell. Biol.*, **13**, 5781–5793.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
Kozak,M. (1989) *J. Cell Biol.*, **108**, 229–241.
Kozak,M. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8301–8305.
Kozak,M. (1991a) *J. Cell Biol.*, **115**, 887–903.
Kozak,M. (1991b) *J. Biol. Chem.*, **266**, 19867–19870.
Lanford,R., Kanda,P. and Kennedy,R. (1986) *Cell*, **46**, 575–582.
Laskey,R. and Dingwall,C. (1993) *Cell*, **74**, 585–586.
Lock,P., Ralph,S., Stanley,E., Boulet,I., Ramsey,R. and Dunn,A. (1991) *Mol. Cell. Biol.*, **11**, 4363–4370.
Maher,D.W., Lee,B.A. and Donoghue,D.J. (1989) *Mol. Cell. Biol.*, **9**, 2251–2253.
Mansour,S. and Martin,G. (1988) *EMBO J.*, **7**, 2035–2041.
Moore,M.S. and Blobel,G. (1992) *Cell*, **69**, 939–950.
Munro,S. and Pelham,H. (1986) *Cell*, **46**, 291–300.
Pappin,D.J.C., Coull,J.M. and Köster,H. (1990) *Anal. Biochem.*, **187**, 10–19.
Peters,G. (1991) *Sem. Virol.*, **2**, 319–328.
Prats,H. *et al.*, (1989) *Proc. Natl Acad. Sci. USA.*, **86**, 1836–1840.
Renko,M., Quarto,N., Morimoto,T. and Rifkin,D. (1990) *J. Cell Physiol.*, **144**, 108–114.
Richardson,W.D., Mills,A.D., Dilworth,S.M., Laskey,R.K. and Dingwall,C. (1988) *Cell*, **52**, 655–664.
Robbins,J., Dilworth,S.M., Laskey,R.A. and Dingwall,C. (1991) *Cell*, **64**, 615–623.
Roberts,B.L., Richardson,W.D. and Smith,A.E. (1987) *Cell*, **50**, 465–475.
Sharma,S., Rodgers,L., Brandsma,J., Gething,M.-J. and Sambrook,J. (1985) *EMBO J.*, **4**, 1479–1489.
Smith,R., Peters,G. and Dickson,C. (1988) *EMBO J.*, **7**, 1013–1022.
Spence,A., Sheppard,P., Davie,J., Matuo,Y., Nishi,N., McKeehan,W., Dodd,J. and Matusik,R. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7843–7847.
von Heijne,G. (1986) *Nucleic Acids Res.*, **14**, 4683–4690.
Water,P. and Lingappa,V. (1986) *Ann. Rev. Cell Biol.*, **2**, 499–516.

Received on May 12, 1994; revised on June 20, 1994