



# Molecular characterization of a zebrafish TCF ETS-domain transcription factor

Louise A Brown<sup>1,2,4</sup>, Shen-Hsi Yang<sup>1,3,4</sup>, Alan Hair<sup>1,3</sup>, Alex Galanis<sup>1,3</sup> and Andrew D Sharrocks<sup>\*1,3</sup>

<sup>1</sup>Department of Biochemistry and Genetics, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne, NE2 4HH, UK

**The ternary complex factor (TCF) subfamily of ETS-transcription factors represent key nuclear targets of the MAP kinase pathways. Members of this subfamily are classified by the presence of several conserved domains for DNA binding, interaction with SRF, interaction with MAP kinases and transcriptional activation. In this study we have isolated a further member of this subfamily (*TCF-1*) from zebrafish. The protein product of zebrafish *TCF-1* (zTCF-1), shares sequence similarity with the mammalian TCFs in all four conserved domains, with highest overall similarity to SAP-1. Zebrafish *TCF-1* is expressed throughout zebrafish embryonic development and exhibits typical TCF DNA binding characteristics, with the B-box being required for interaction with SRF. Of the mammalian TCFs, its DNA binding specificity resembles Elk-1. zTCF-1 is a target for both the growth factor/mitogen-activated and stress-activated MAP kinase cascades *in vitro* and *in vivo*. However, differential targeting occurs, with the profile of its activation closely resembling that of mammalian SAP-1. Together, our results demonstrate that the TCFs have been functionally conserved during vertebrate development.**

**Keywords:** ETS-domain; ternary complex factor; transcription factors; zebrafish

## Introduction

The ternary complex factors (TCFs) represent a subfamily of the ETS-domain transcription factors (reviewed in Treisman, 1994). In mammals, there are three known TCFs, Elk-1, SAP-1 and SAP-2/ERP/Net. Members of this subfamily are characterized by a high degree of sequence conservation in their ETS DNA-binding domains and in three additional conserved domains; the B-box, D-domain and C-domain (see Figure 1 and 3a). The B-box is required for efficient ternary complex formation with SRF and the *c-fos* SRE (Dalton and Treisman, 1992; Janknecht *et al.*, 1994; Hill *et al.*, 1993; Price *et al.*, 1995; Giovane *et al.*, 1994; Lopez *et al.*, 1994; Shore and Sharrocks, 1994). In Elk-1, the D-domain has been shown to act as a

docking motif for the ERK and JNK MAP kinases (Yang *et al.*, 1998a,b) and the C-domain is a MAP kinase-regulated transcriptional activation domain (TAD) (reviewed in Treisman, 1994).

In addition to forming ternary DNA-bound complexes with SRF, TCFs exhibit an autonomous, SRF-independent binding activity to higher affinity binding sites (reviewed in Treisman, 1994). In comparison to SAP-1, Elk-1 exhibits a more stringent DNA binding specificity (Shore and Sharrocks, 1995) whereas SAP-2 also differs and exhibits an intermediate specificity (Shore *et al.*, 1996; ALB, unpublished data). Both autonomous and SRF-dependent DNA binding by TCFs are controlled by cis-acting negative regulatory domains. In Elk-1 and SAP-1, the B-box appears sufficient to inhibit DNA binding although further inhibitory effects are exerted by C-terminal amino acids (Dalton and Treisman, 1992; Janknecht *et al.*, 1994). In the case of SAP-2, DNA binding appears to be inhibited by a combination of the C-terminal TAD and an additional domain, the net inhibitory domain (NID), which is located immediately C-terminal to the B-box (Giovane *et al.*, 1994; Lopez *et al.*, 1994; Price *et al.*, 1995; Maira *et al.*, 1996). However, ternary complex formation by SAP-2 is inefficient in comparison to Elk-1 and SAP-1 (Price *et al.*, 1995). In the case of Elk-1 and SAP-1, phosphorylation of their C-terminal TAD by MAP kinases, leads to the relief of inhibition and the activation of their DNA binding activity (Sharrocks, 1995; Gille *et al.*, 1992, 1995a,b; Whitmarsh *et al.*, 1995, 1997).

All three mammalian TCFs can be phosphorylated and activated by MAP kinases *in vitro* and *in vivo*. However, the response of each TCF to individual growth factor-regulated (ERK) and cytokine/stress-inducible (JNK and p38) MAP kinase pathways differs. Elk-1 is efficiently phosphorylated *in vitro* and activated *in vivo* by ERK, JNK and p38 $\alpha$ , p38 $\beta$ <sub>2</sub> and p38 $\gamma$  (reviewed in Treisman, 1996; Price *et al.*, 1996; Janknecht and Hunter, 1997; Whitmarsh *et al.*, 1997; Enslin *et al.*, 1998) although in the case of the p38 isoforms, targeting via the D-domain does not occur (Yang *et al.*, 1998b). In contrast, SAP-1 is a good target for ERK and p38 $\alpha$  but is a poor JNK substrate (Whitmarsh *et al.*, 1995, 1997; Price *et al.*, 1996; Strahl *et al.*, 1996). SAP-2 also appears to only be a good substrate for all the ERK and p38 $\alpha$  MAP kinase isoforms (Price *et al.*, 1996).

The *in vivo* role of the TCFs is currently poorly understood although one of their functions is to regulate the expression of immediate-early genes in response to various inductive stimuli (reviewed in Treisman, 1994; Cahill *et al.*, 1996). Elk-1 has been

\*Correspondence: AD Sharrocks

Current addresses: <sup>2</sup>Unilever Research, Colworth, Sharnbrook Bedford, MK44 1LQ, UK; <sup>3</sup>School of Biological Sciences, Stoppard Building, University of Manchester, Oxford Road, Manchester. M13 9PT, UK.

<sup>4</sup>The first two authors contributed equally to this work

Received 24 March 1999; revised 31 August 1999; accepted 1 September 1999

shown to be expressed in an active form in cells of the nervous system within the brain, suggesting a role in regulating neuronal gene expression (Sgambato *et al.*, 1998). In this study, we have isolated a cDNA encoding a member of the TCF subfamily from zebrafish (*TCF-1*) which is expressed during early development. The high degree of sequence similarity and DNA binding properties of the protein product of zebrafish TCF-1 (zTCF-1) demonstrate that it is conserved at both the sequence and functional levels with its mammalian counterparts. Moreover, in common with mammalian TCFs, zTCF-1 is activated and phosphorylated in a differential manner by a series of MAP kinase pathways and thereby can respond to diverse signals mediated by growth factors, mitogens, cytokines and stress. The identification and characterization of a TCF homologue from zebrafish is an important step in understanding the role of this subfamily of ETS-domain transcription factors in mediating the action of MAP kinase pathways during vertebrate development.

**Results**

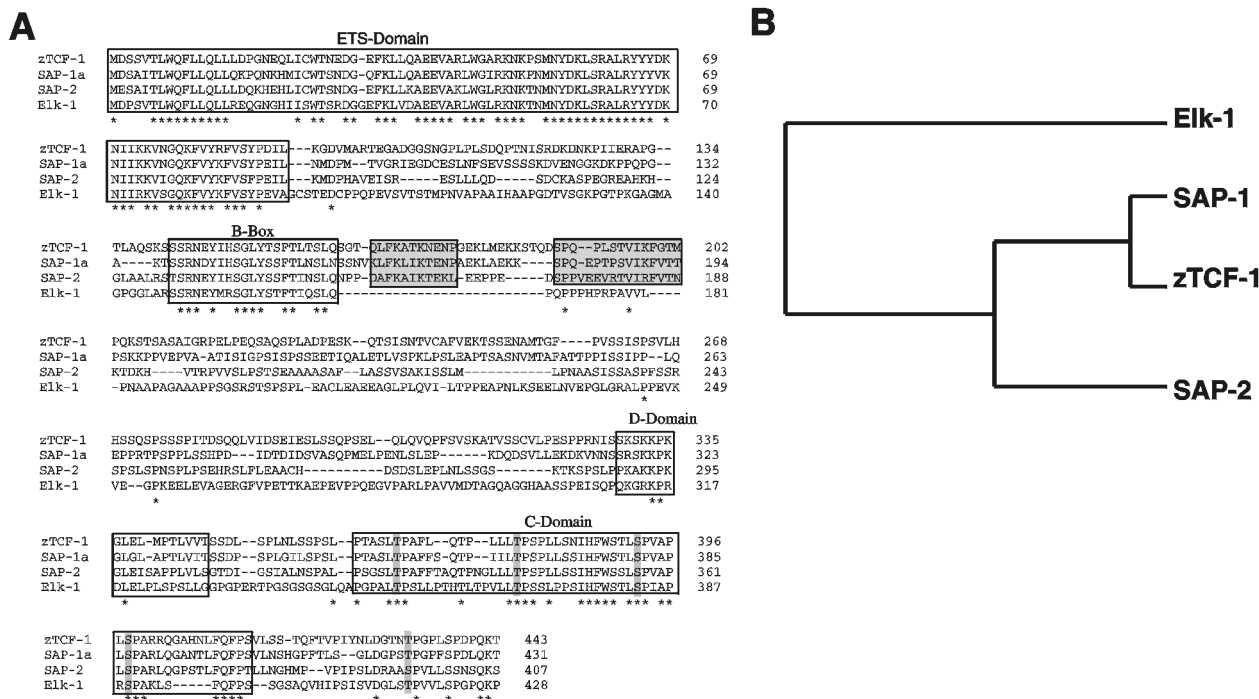
*Cloning of a member of the TCF subfamily of ETS-domain proteins from zebrafish*

A cDNA encoding part of the conserved ETS DNA-binding domain of a putative zebrafish TCF homologue (*TCF-1*) was amplified from an embryonic zebrafish cDNA pool using degenerate oligonucleotides and used to screen an embryonic cDNA library. One of the

isolated cDNAs ( $\approx 1.9$  kb) was sequenced and shown to contain an open-reading frame of 1329 nucleotides encoding a protein of 443 amino acids with a predicted molecular mass of 47 kDa (Figure 1a).

The amino acid sequence of the protein product of zebrafish *TCF-1* (zTCF-1) shares strong homology with the mammalian members of the TCF subfamily of ETS-domain proteins (Figure 1a). In particular, high amino acid sequence conservation is exhibited within the ETS-domain (82, 80 and 76% identical to SAP-1, SAP-2 and Elk-1 respectively) with further strong similarity within the other three functional domains conserved in this subfamily (the B-box and D- and C-domains). The mammalian TCFs are known targets of several MAP kinase cascades (reviewed in Treisman, 1996; Whitmarsh and Davis, 1996; Janknecht and Hunter, 1997; Price *et al.*, 1996; Whitmarsh *et al.*, 1997; Strahl *et al.*, 1996). Five potential phosphoacceptor sites for MAP kinases are conserved amongst zTCF-1 and the mammalian family members (Figure 1a). All these sites in the mammalian TCFs have been shown to be phosphorylated by one or more of the MAP kinases *in vitro* and *in vivo*.

Overall, the protein product of zebrafish *TCF-1* shows strongest identity to human SAP-1a (56% overall identity) with 47 and 38% overall identity observed between SAP-2 and Elk-1 respectively. Furthermore, phylogenetic analysis suggests that zebrafish *TCF-1* is more closely evolutionarily related to *SAP-1* than to the other two mammalian TCFs (Figure 1b). From primary sequence comparisons, zebrafish *TCF-1* is therefore most likely to represent a homologue of mammalian *SAP-1*.



**Figure 1** (a) Alignment of the amino acid sequences of zebrafish TCF-1 and human Elk-1 (Rao *et al.*, 1989), SAP-1a (Dalton and Treisman, 1992) and SAP-2 (Price *et al.*, 1995). Residues which are identical in all four proteins are indicated by asterisks. The highly conserved B-box and ETS-, D-, and C-domains are boxed. An additional region conserved amongst zTCF-1, SAP-1 and SAP-2 is boxed and shaded. Potential MAP kinase phosphorylation sites conserved amongst all four proteins are highlighted. (b) Clustal analysis (PC/GENE) of the phylogenetic relationship amongst zebrafish *TCF-1*, and mammalian *Elk-1*, *SAP-1* and *SAP-2*

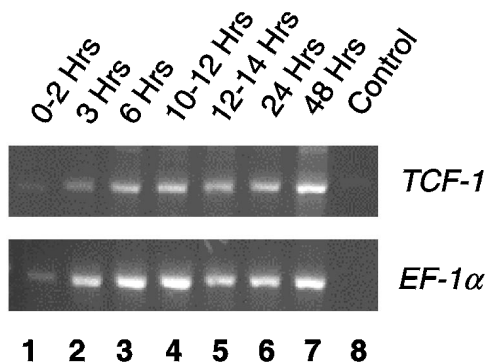
*Expression of zebrafish TCF-1 during embryonic development*

In order to demonstrate that zebrafish *TCF-1* is expressed during early development, its expression was detected using RT-PCR (Figure 2). As a control, its expression profile was compared to *EF-1 $\alpha$*  which is a ubiquitously expressed translation elongation factor. Initial experiments were carried out to ensure that the PCR reaction was in the linear range (Hair and Sharrocks, data not shown). Zebrafish *TCF-1* transcripts could be detected at low levels during the earliest stages examined (0–2 h post fertilization (hpf)). The expression of *TCF-1* increased substantially after 3 hpf and was maintained during the first 48 h of development (Figure 2, lanes 2–7). The relative expression levels of zebrafish *TCF-1* at each stage closely mirrored that of *EF-1 $\alpha$* , albeit at a lower level. In order to detect its spatial and temporal expression pattern, whole mount *in situ* hybridization experiments were carried out. However, no specific zebrafish *TCF-1* expression pattern could be detected (data not shown; see Discussion).

*DNA binding properties of zTCF-1*

The mammalian TCFs Elk-1 and SAP-1 efficiently form ternary complexes with SRF and the *c-fos* SRE. Ternary complex formation requires a combination of protein-DNA interactions by the ETS-domain and binding to SRF via the B-box (reviewed in Treisman, 1994). SAP-2 also exhibits ternary complex forming ability but its efficiency is much reduced in comparison to SAP-1 and Elk-1 unless C-terminal truncations are made (Price *et al.*, 1995). All the TCFs can bind to the high affinity E74-like binding motifs in an SRF-independent manner (Dalton and Treisman, 1992; Janknecht *et al.*, 1994; Lopez *et al.*, 1994; Giovane *et al.*, 1994; Whitmarsh *et al.*, 1995; Price *et al.*, 1995).

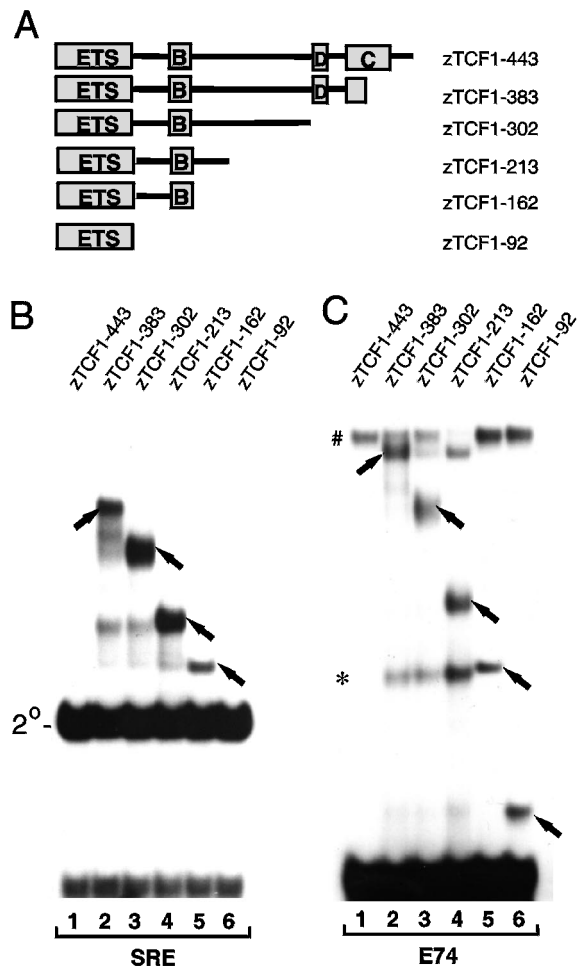
The DNA binding properties of zTCF-1 were investigated on the *c-fos* SRE and the E74 site (Figure 3). A series of truncated zTCF-1 proteins were produced to analyse the contribution of the conserved domains to DNA binding (Figure 3a).



**Figure 2** Expression of zebrafish *TCF-1* and *EF-1 $\alpha$*  during early embryonic development. RT-PCR was carried out with specific primer pairs for zebrafish *TCF-1* (top panel) and *EF-1 $\alpha$*  (bottom panel) with RNA samples from embryos at preMBT (0–2 h), blastula (3 h), gastrula (6 h), segmentation1(10–12 h), segmentation (12–14 h), 24 h, 48 h. The times given are hours post fertilization (hpf). The negative control in lane 8 was carried out in the absence of RNA

Autonomous binding of the zTCF-1 proteins to the SRE in the absence of SRF could not be detected (data not shown). However, in the presence of SRF, ternary zTCF-1-SRF-SRE complexes could be detected (Figure 3b). Full-length zTCF-1 formed weak ternary complexes (not visible in Figure 3b; see Figure 5). However, a series of C-terminally truncated derivatives efficiently formed ternary complexes with the exception of zTCF1-92 in which the B-box region was removed (Figure 3b lanes 1–6). In the absence of SRF, all the zTCF-1 derivatives efficiently bound the E74 site (Figure 3c) except the full-length protein which bound this site with reduced affinity (not visible in Figure 3c; see Figure 5).

Together these results demonstrate that the ETS-domain is sufficient to mediate zTCF-1 binding to high affinity sites but that the B-box is required to form



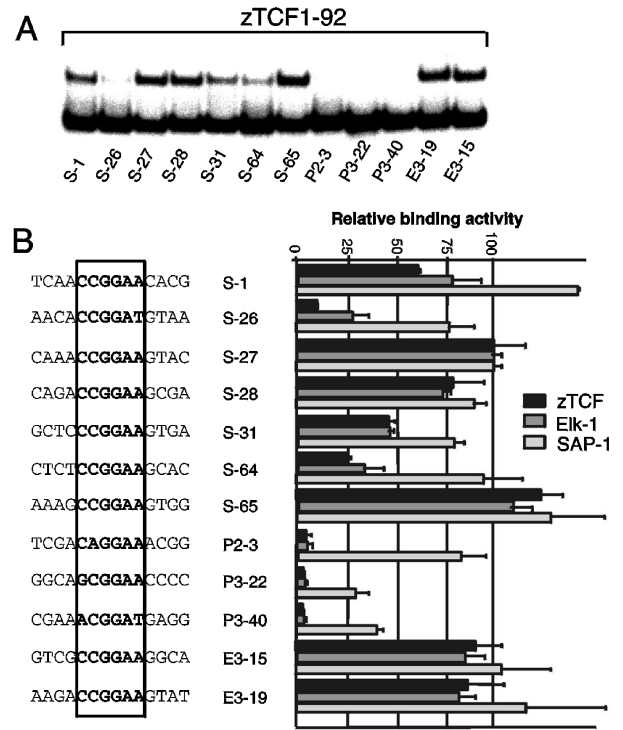
**Figure 3** DNA binding by zTCF-1. (a) Schematic illustration of the domain structure of the full-length and truncated proteins used in (b) and (c). Numbers indicate the positions of residues with respect to the full-length proteins. (b) and (c) Gel retardation analysis of wild-type and truncated zTCF-1 proteins bound to the *c-fos* SRE (b) or E74 (c) binding sites. Binding to the SRE was carried out in the presence of core<sup>SRF</sup>. The position of binary SRF-SRE complexes is indicated (2°). Arrows represent ternary zTCF-1-SRF-SRE (b) and binary zTCF-1-E74 (c) complexes formed by the indicated ‘full-length’ zTCF-1 derivatives. (\*) Represents a complex formed by a C-terminally degraded/truncated zTCF-1 protein whereas (#) represents complexes formed by a protein from the rabbit reticulocyte lysate. Equal molar quantities of each protein are used in the binding reactions

ternary complexes with SRF and the *c-fos* SRE. The C-terminal end of the protein appears to represent an inhibitory domain as its removal stimulates complex formation on both the E74 site and the SRE. zTCF-1 therefore shares many of the DNA binding properties exhibited by mammalian TCFs (see Discussion).

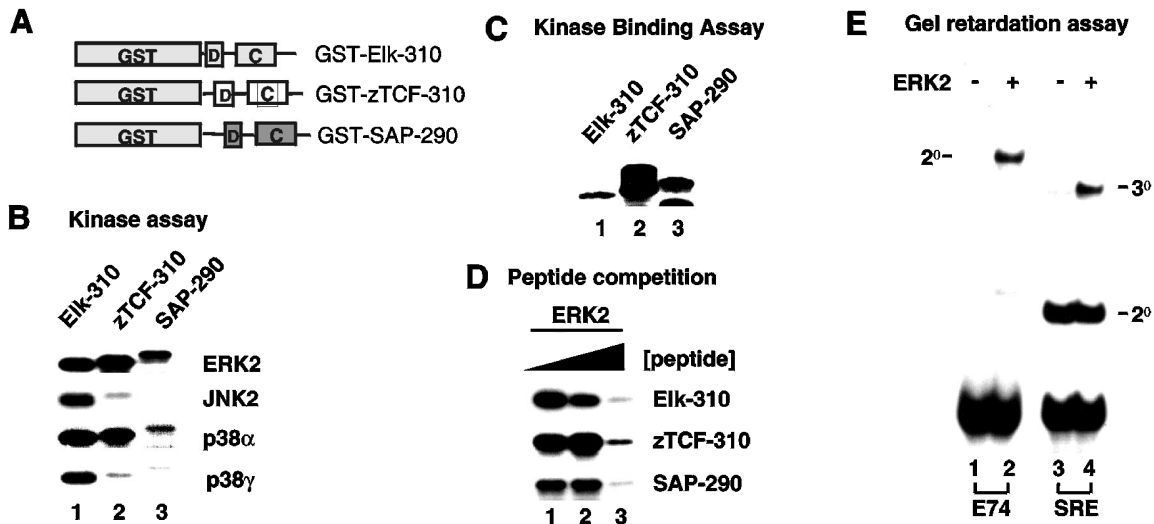
The mammalian TCFs exhibit different DNA binding specificities with Elk-1 showing a more stringent specificity than either SAP-1 or SAP-2 (Shore and Sharrocks, 1995; Shore *et al.*, 1996). The binding of zTCF-1 to a series of sites containing ets-motifs was therefore tested. zTCF-1 bound to these sites with differing efficiencies (Figure 4a). In comparison to Elk-1 and SAP-1, the relative binding efficiencies of zTCF-1 closely mirrors those exhibited by Elk-1 (Figure 4b). SAP-1 exhibits a lower site-selectivity and binds all the tested sites better than either zTCF-1 or Elk-1 (Figure 4b; Brown *et al.*, 1998a). The DNA binding specificity of zTCF-1 is therefore most similar to that of human Elk-1 (see Discussion).

*In vitro phosphorylation and activation of zTCF-1 by MAP kinases*

The mammalian TCFs are known targets of both the mitogen-activated (ERK) and stress-inducible (JNK and p38) MAP kinase cascades (reviewed in Treisman, 1996; Whitmarsh and Davis, 1996; Janknecht and Hunter, 1997; Price *et al.*, 1996; Whitmarsh *et al.*, 1997; Strahl *et al.*, 1996). zTCF-1 was therefore tested as a potential substrate for representative members of



**Figure 4** The DNA binding specificity of zTCF-1. (a) Gel retardation analysis of zTCF1-92 binding to a panel of ets binding sites. (b) Quantification of the data from part (a) in comparison to the binding of the equivalent DNA-binding domains from Elk-1 (Elk-1-93) and SAP-1 (SAP1-92) to the same series of sites (Brown *et al.*, 1998a). Data are quantified relative to the binding of each protein to the S27 site (taken as 100%). The central six nucleotides of the ets-motifs are boxed



**Figure 5** zTCF-1 is an *in vitro* target for MAP kinases. (a) Schematic illustration of the domain structure of the GST-TCF fusion proteins which contain the C-terminal regions of Elk-1, zTCF-1 and SAP-1. (b) The phosphorylation of the indicated GST-TCF fusion proteins by MAP kinases was examined using either activated ERK2, JNK-2, p38 $\alpha$  or p38 $\gamma$ . The activity of each protein kinase towards the substrate GST-Elk310 was standardized with respect to activated ERK2 (0.1U, NEB). Kinase assays were performed for 15 min at 30°C with equal molar quantities (5pmoles) of GST fusion proteins as substrates. (c) Binding and phosphorylation assays. Equal molar quantities (37.5 pmol) of the indicated GST fusion proteins were immobilized onto glutathione agarose beads and incubated with ERK2 (1U) for 4 h at 4°C. After extensive washing, the beads with the remaining bound kinase were incubated with [ $\gamma$ -<sup>32</sup>P]ATP in kinase buffer for 2 h at 30°C. (d) Phosphorylation of the indicated GST-TCF fusions by ERK2 in the presence of a competitor peptide corresponding to the Elk-1 D-domain. Kinase assays were carried out as described in part (b) except the MAP kinases were preincubated in the absence (lane 1) or presence of competitor peptides (100–1000-fold excess over the indicated GST-fusion substrates); 1 nmoles (lane 2) and 10 nmoles (lanes 3). (e) Gel retardation analysis of non-phosphorylated (lanes 1 and 3) and phosphorylated (lanes 2 and 4) zTCF-1 binding to the *c-fos* E74 (lanes 1–2) or SRE (lanes 3–4) binding sites. Binding to the SRE was carried out in the presence of core<sup>SRF</sup>. Labelling of the complexes is as described in Figure 3

the different classes of MAP kinases. Elk-1 is efficiently phosphorylated *in vitro* by ERK2, JNK2, p38 $\alpha$  and p38 $\gamma$  (Figure 5b, lane 1). However, in comparison, zTCF-1 and SAP-1 are only efficiently phosphorylated by ERK2 and p38 $\alpha$  (Figure 5b lanes 2 and 3).

ERKs have been demonstrated to interact directly with Elk-1 (Rao and Reddy, 1994; Cano *et al.*, 1995; Yang *et al.*, 1998a) by binding to the D-domain (Yang *et al.*, 1998a). The binding of ERK2 to the C-terminal domains of Elk-1, zTCF-1 and SAP-1 was investigated by incubation of the ERK2 with GST fusion proteins immobilised on glutathione-agarose beads. Specifically bound proteins were obtained by co-precipitation with the beads followed by removal of non-specifically bound proteins by extensive washing. Bound kinases were subsequently detected by incubation of the final precipitates with  $\gamma^{32}\text{P}$ -ATP to detect phosphorylation of the GST fusion proteins. Binding of ERK2 was detected to Elk-1, zTCF-1 and SAP-1 (Figure 5c lanes 1–3). Peptide competition assays using a peptide derived from the Elk-1 D-domain have been used to demonstrate the binding of ERK2 to its substrate occurs via the D-domain (Yang *et al.*, 1998b). Similarly, incubation of ERK2 with the D-domain peptide blocks its ability to efficiently phosphorylate zTCF-1 and SAP-1 (Figure 5d), indicating that targeting of ERK2 to zTCF-1 is required for efficient phosphorylation.

DNA binding by mammalian TCFs is activated by phosphorylation (Sharrocks, 1995; Gille *et al.*, 1992, 1995a, b; Whitmarsh *et al.*, 1995, 1997). *In vitro* translated full-length zTCF-1 binds DNA inefficiently (Figure 3), therefore we tested whether phosphorylation by MAP kinases stimulates its DNA binding activity. Phosphorylation of zTCF-1 by ERK2 stimulates its ability to bind autonomously to the high affinity E74 site (Figure 5e lanes 1–2) and to form ternary complexes with SRF and the *c-fos* SRE (Figure 5e lanes 3–4).

Together these results demonstrate that zTCF-1 is selectively phosphorylated by different classes of MAP kinases *in vitro* which leads to the activation of its DNA binding activity.

#### Activation of zTCF-1 by MAP kinase cascades *in vivo*

Mammalian TCFs also exhibit different responses to individual growth factor/mitogen-activated and stress-inducible MAP kinase cascades *in vivo* (reviewed in Treisman, 1996; Whitmarsh and Davis, 1996; Janknecht and Hunter, 1997; Price *et al.*, 1996; Whitmarsh *et al.*, 1997; Strahl *et al.*, 1996). For example, JNKs only poorly phosphorylate and activate SAP-1 in comparison to Elk-1 (Whitmarsh *et al.*, 1995; Gille *et al.*, 1995b; Strahl *et al.*, 1996; Price *et al.*, 1996). The response of zTCF-1 to different classes of MAP kinases *in vivo* was therefore tested (Figure 6).

EGF activates the endogenous ERK pathway in COS7 cells and stimulates transcriptional activation by Elk-1 (Yang *et al.*, 1998a). Similarly, both SAP-1 and zTCF-1 are activated by EGF stimulation (Figure 6a).

The stimulation of CHO cells with IL-1 leads to the activation of Elk-1 by the JNK pathway and SAP-1 by an unknown kinase (Figure 6b; Whitmarsh *et al.*, 1997). zTCF-1 is also activated in response to IL-1 stimulation (Figure 6b). However, this activation is not inhibited by the JNK inhibitor protein JIP-1 (Figure

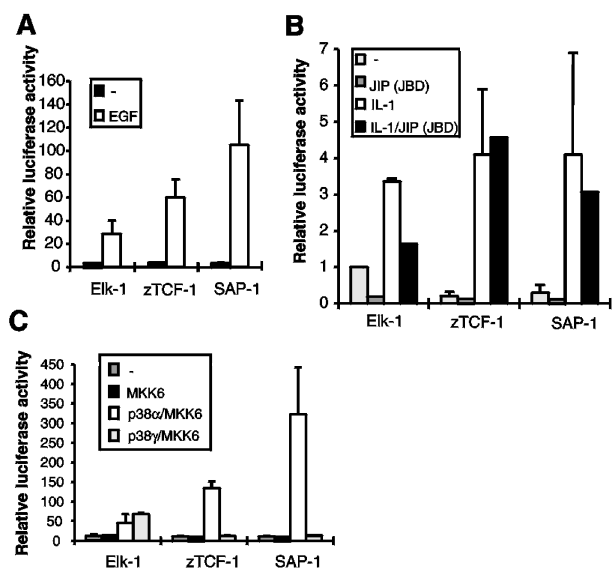
6b) or by a dominant negative form of MKK4 (data not shown), indicating that in common with SAP-1, zTCF-1 is activated by a novel IL-1-inducible kinase.

Elk-1 (Whitmarsh *et al.*, 1997; Price *et al.*, 1996) and SAP-1 (Price *et al.*, 1996; Janknecht and Hunter, 1997) can also be activated by the p38 MAPK kinase cascades *in vivo*. Elk-1 is activated by both p38 $\alpha$  and p38 $\gamma$  when cotransfected with a constitutively active version of the upstream kinase MKK6 (Figure 6c). In contrast, zTCF-1 and SAP-1 are activated by p38 $\alpha$  but neither is significantly activated by p38 $\gamma$  (Figure 6c).

Collectively, these results demonstrate that zTCF-1 is differentially activated by different classes of growth factor/mitogen- and stress-activated MAP kinase cascades *in vivo* in a manner which closely resembles the responses of mammalian SAP-1 to these pathways.

## Discussion

Members of the TCF subfamily of ETS-domain transcription factors are important nuclear mediators of the cellular responses to the activation of MAP kinase pathways (reviewed in Treisman, 1994, 1996; Whitmarsh and Davis, 1996; Sharrocks *et al.*, 1997). In this study, we have isolated a cDNA encoding a further family member from zebrafish (*TCF-1*) that exhibits sequence similarity in four conserved domains which are characteristic of this subfamily. In common with its mammalian counterparts, the protein product of zebrafish *TCF-1*, zTCF-1, represents a target for diverse MAP kinase cascades. Both sequence and



**Figure 6** zTCF-1 is an *in vivo* target for MAP kinase cascades. COS-7 (a and c) or CHO (b) cells were transfected with the indicated CMV promoter-driven GAL–TCF constructs and a GAL4-driven luciferase reporter plasmid. (a) Cells were serum starved for 12 h and either left unstimulated or stimulated with EGF. (b) Cells were left unstimulated or stimulated with IL-1. Expression vectors encoding the JNK pathway inhibitor protein JIP-1(DBD) were cotransfected where indicated. (c) Cells were cotransfected with a constitutively active form of MKK6 on its own or in combination with p38 $\alpha$  or p38 $\gamma$  expression vectors. Transfection efficiency was monitored by using the  $\beta$ -galactosidase expression vector pCH110. The luciferase activities relative to the unstimulated cells containing pCMV5-GAL-Elk-205 (means  $\pm$  s.d.;  $n = 3$ ) are presented

functional conservation of the TCFs points to a critical, fundamental role for these transcription factors during vertebrate development.

#### Conservation of TCF domain structure

Significant sequence similarity is observed between zTCF-1 and the mammalian TCFs, in all four of the conserved functional domains. Overall, the greatest sequence similarity is observed with SAP-1 (56% identity), whilst Elk-1 exhibits the least similarity (38%). This pattern is repeated within each conserved domain with SAP-1 exhibiting the greatest sequence identity in each case (ETS-domain, 82%; B-box, 81%; D-domain, 73%; C-domain, 87%). Furthermore, in addition to the five MAP kinase sites conserved in all the TCFs, five additional potential MAP kinase sites are conserved within the C-terminal domains of zTCF-1 and SAP-1. An additional region appears to be conserved amongst zTCF-1, SAP-1 and SAP-2 which acts as a negative inhibitory domain in Net (the murine SAP-2 homologue) and is thought to adopt a helix-loop-helix like structure (Maira *et al.*, 1996). The high overall similarity to SAP-1 and within the important functional domains of the protein, suggests that zTCF-1 is most likely a homologue of mammalian SAP-1.

#### DNA binding properties of zTCF-1

In common with the mammalian TCFs, zTCF-1 can bind DNA either autonomously or in an SRF-dependent manner on weaker binding sites. Autonomous binding is mediated by the ETS-domain alone and the binding specificity of zTCF-1 closely resembles that of Elk-1 which exhibits a more stringent binding specificity than either SAP-1 or SAP-2 (Figure 4, Shore and Sharrocks, 1995; Shore *et al.*, 1996; Brown *et al.*, 1998a). It was previously demonstrated that the presence of amino acid D69 in the Elk-1 ETS-domain (V68 in SAP-1) is mainly responsible for generating this stringent binding specificity (Shore *et al.*, 1996). Similarly, zTCF-1 contains the same residue at the equivalent position in the ETS-domain (D68; Figure 1).

Ternary complex formation with SRF and the *c-fos* SRE requires the presence of the B-box region of zTCF-1 (Figure 3b). However, it is interesting to note that the full-length protein inefficiently forms DNA-bound complexes both in the presence or in the absence of SRF. Such negative regulation has been observed in all the mammalian TCFs although different regions have been implicated in mediating this repressive effect (Dalton and Treisman, 1992; Janknecht *et al.*, 1994; Price *et al.*, 1995; Maira *et al.*, 1996). In the case of zTCF-1, deletion of the C-terminal 60 amino acids is sufficient to relieve this inhibition. In contrast to SAP-2/Net, no further negative regulatory regions appear to be present towards the N-terminus of the protein. In the case of Elk-1 and SAP-1, MAP kinase-mediated phosphorylation activates their DNA binding (Sharrocks, 1995; Gille *et al.*, 1992, 1995a,b; Whitmarsh *et al.*, 1995, 1997) although the mechanism of SAP-2 activation is unclear. Similarly, phosphorylation of zTCF-1 stimulates the binding activity of the full-length protein (Figure 5e).

#### Activation of zTCF-1 by MAP kinase cascades

The mammalian TCFs are known targets of both the mitogen-activated (ERK) and stress-inducible (JNK and p38) MAP kinase cascades (reviewed in Treisman, 1996; Whitmarsh and Davis, 1996; Janknecht and Hunter, 1997; Price *et al.*, 1996; Whitmarsh *et al.*, 1997; Strahl *et al.*, 1996). zTCF-1 is also an *in vitro* MAP kinase substrate and a target for these cascades *in vivo* (Figures 5 and 6). However, in common with the mammalian TCFs, zTCF-1 appears to be differentially targeted by different classes of MAP kinases. In comparison to Elk-1, zTCF-1 is efficiently phosphorylated by ERK2, and p38 $\alpha$  but is a poor substrate for JNK2 and p38 $\gamma$ . Similarly, *in vivo*, zTCF-1 is only efficiently activated by ERK2 and p38 $\alpha$ . This activation profile closely resembles that of SAP-1. In addition, IL-1 activation of both zTCF-1 and SAP-1 takes place by a cascade which is distinct from the JNKs (Figure 6b; Whitmarsh *et al.*, 1997). As the highest conservation of potential MAP kinase sites in and around the C-terminal activation domain is also observed with SAP-1, the signalling pathways that converge on these transcription factors and their mechanisms of activation are likely to be very similar. Peptide competition assays (Figure 5d) indicate that in common with Elk-1 (Yang *et al.*, 1998b), phosphorylation of both SAP-1 and zTCF-1 by ERK is facilitated by the presence of the D-domain. Therefore docking of the kinase onto a site distinct from the phosphoacceptor motifs appears to be an evolutionarily common mechanism in the TCF subfamily.

#### zTCF-1 is a homologue of which mammalian TCF?

The sequence conservation and activation profile by MAP kinase cascades suggest that zTCF-1 represents a direct homologue of mammalian SAP-1. However, the DNA binding specificity more closely resembles that of Elk-1. It is possible that other mammalian TCFs exist which exhibit stronger sequence similarity to zTCF-1. In this regard, it is interesting to note that the other family members cloned from zebrafish (PEA3, 61% identity and Fli-1, 75% identity) (Brown *et al.*, 1998a,b) exhibit substantially higher sequence similarity than zTCF-1 to their closest mammalian counterparts. The existence of other zebrafish TCFs is suggested from the sequence of cDNAs encoding part of the ETS-domain isolated using degenerate PCR primers (our unpublished data). These additional TCFs might represent homologues of other mammalian TCFs.

#### Potential role of zTCF-1 during development

To date, little is known about the roles of TCFs in regulating transcription during vertebrate development although one of their functions is thought to be to regulate the expression of immediate-early genes in response to various inductive stimuli (reviewed in Treisman, 1994; Strahl *et al.*, 1996). In adults, mammalian TCFs appear to be ubiquitously expressed with each tissue expressing the individual TCFs to different relative amounts (Price *et al.*, 1995; Magnaghi-Jaulin *et al.*, 1996). Recently, Elk-1 has been shown to be expressed in an active form in cells of the nervous

system within the brain (Sgambato *et al.*, 1998). Zebrafish *TCF-1* is expressed throughout early development. However, we could not detect a specific expression pattern by whole-mount *in situ* hybridization, indicating that its expression levels may be too low (data not shown). Similarly, mammalian TCFs are expressed at low levels (Price *et al.*, 1995). As the relative expression levels of *TCF-1* closely resembles that of the ubiquitously expressed *EF-1 $\alpha$*  (Figure 2), *TCF-1* may also be expressed ubiquitously in the zebrafish embryo, as demonstrated for its mammalian counterparts in a variety of cell lines. A low level ubiquitous expression pattern would explain our inability to detect specific expression patterns in the early embryo.

MAP kinase pathways and their nuclear substrates have been implicated in several important regulatory processes (reviewed in Treisman, 1996; Whitmarsh and Davis, 1996). Furthermore, several developmental events in lower eukaryotes have been shown to be regulated by MAP kinase cascades *via* ETS-domain transcription factors. However, the role of such pathways and transcription factors in vertebrate development has not been addressed. The identification of a zebrafish TCF homologue provides the basis for future studies to investigate the role of these transcription factors in transmitting signals from MAP kinase pathways into a transcriptional response during early vertebrate development.

## Materials and methods

### Cloning of zebrafish TCF

A random-primed  $\lambda$ Zap cDNA library, prepared from 20–40 h-old zebrafish embryos, was screened with a mix of two partial zebrafish cDNAs encoding the central region of the ETS DNA-binding domain of two potential TCF homologues as described elsewhere (Brown *et al.*, 1998b). One  $\lambda$  clone was purified which contained a cDNA insert corresponding to zebrafish *TCF-1*. A clone encoding a second TCF was not isolated. pAS253 (encoding full-length zTCF-1) was created by excision of the pBluescriptSK<sup>+</sup> vector by co-infection with helper phage. pAS256 is identical to pAS253 except that the zebrafish *TCF-1* cDNA is inserted into the *EcoRI* site of pBluescriptSK<sup>+</sup> in the opposite orientation.

### RT-PCR

Embryos of appropriate stages were isolated and dechlorinated before being frozen in liquid nitrogen. Total RNA was extracted using the RNeasy extraction kit according to the manufacturer's instructions (QIAGEN). In each case, 30 mg of tissue was used (representing 20–35 embryos). Samples were normalized based on the concentration of the 18S rRNA band and equal amounts ( $\approx 0.5 \mu\text{g}$ ) were used in random hexanucleotide primed reverse transcription reactions followed by a 32 cycle PCR (Plant *et al.*, 1996) using the following primer pairs: EF-1 $\alpha$  [EFzebra1 (5'-TTCCGTCTGCCACTTCAGGAT-3') and EFzebra2 (5'-CAACGGTCTGCCTCATGTCA-3')].  $T_a = 57^\circ\text{C}$ , giving a product of 563 bp] and zTCF-1 [ADS603 (5'-CCAACCAAACATCTCC-3') and ADS604 (5'-CTCACTGTCAAT-CACCAGC-3')].  $T_a = 52^\circ\text{C}$ , giving a product of 526 bp]. PCRs were optimized to ensure that the amplification was in the linear range and the products visualized on an ethidium bromide stained 3% NuSieve agarose gel (Flowgen).

### Plasmid constructs

pAS351 (encoding GST fused to zTCF-1 amino acids 311–443) was constructed by ligating a *BamHI/XbaI* cleaved PCR product (primers ADS267/326) into the same sites in pGEX-KG (Guan and Dixon, 1991). pAS466 (encoding C-terminally Flag-tagged full-length zTCF-1) was constructed by ligating a *NcoI/XhoI* cleaved PCR product (primers ADS293/325) into the same sites in pET-nef-PFH (Zhao and Narayan, 1993). pAS545 (encoding GST-Elk-310; Elk-1 amino acids 310–428) (Yang *et al.*, 1998a) and pGEX-SAP-290 (encoding GST-SAP-290; SAP-1 amino acids 290–431) (Whitmarsh *et al.*, 1995) have been described previously.

The following plasmids were prepared for producing proteins by *in vitro* transcription/translation. pAS316 (encoding zTCF-1 amino acids 1–92; zTCF1-92) and pAS317 (encoding zTCF-1 amino acids 1–162; zTCF1-162) were constructed by ligating *NcoI/BamHI* and *NcoI/NotI* cleaved PCR products (primers ADS318/319 and ADS309-319 respectively) into the same sites in pAS37 (Sharrocks *et al.*, 1993a). All PCRs to generate zTCF-1-derived constructs were carried out using pAS253 as a template.

For transient transfection assays, the following constructs were used: pAS572 (encoding GAL-Elk-205 [Elk-1 amino acids 205–428] under the control of a CMV promoter) (Yang *et al.*, 1998a). pAS855 (encoding GAL-SAP-290 [SAP-1 amino acids 290–431] under the control of a CMV promoter), was constructed by inserting a *HindIII/XbaI* fragment from pGAL-SAP-1 (Whitmarsh *et al.*, 1997) into the same sites of pCMV5. pAS856 (encoding GST fused to zTCF-1 amino acids 310–443) was constructed by ligating a *BamHI/XbaI* cleaved PCR product (primers ADS267/268) into the same sites in pAS769. pAS857 (encoding the GAL4 DNA binding domain fused to zTCF-1 amino acids 310–443) was constructed by ligating a *BamHI/XbaI* fragment from pAS857 into the same sites in pSG424. pAS858 (encoding GAL-zTCF-310–443 driven by a CMV promoter) was constructed by inserting a *HindIII/XbaI* fragment from pAS857 into pCMV5.

pCMV5-F-p38 $\alpha$  (Raingeaud *et al.*, 1995), pCDNA3-F-p38 $\gamma$  (Enslin *et al.*, 1998), pCDNA3-F-MKK6(Glu) (encoding constitutive-active MKK6) (Whitmarsh *et al.*, 1996), pCDNA3-F-DN-MKK4 (encoding dominant-negative MKK4) (Whitmarsh *et al.*, 1995), pCDNA3-F-JIP-1(JBD) (encoding the JNK binding domain of JIP-1) (Dickens *et al.*, 1997), pG5E1b-luc (Seth *et al.*, 1992) have been described previously.

The nucleotide sequences of all the PCR-derived constructs were verified by di-deoxy sequencing.

### Protein production

Purification and expression of GST-fusion and hexahistidine-tagged proteins were carried out essentially as described previously (Shore *et al.*, 1995; Yang *et al.*, 1998a). *In vitro* translated proteins were synthesized using the TNT coupled *in vitro* transcription/translation system (Promega) from either pBluescriptKS<sup>+</sup>-derived plasmids or from linear PCR products under the control of a T3 promoter. The primer pairs ADS238/ADS261, ADS238/265 and ADS238/274 were used on the template pAS253 to synthesize PCR products encoding the proteins zTCF1-213, zTCF1-383 and zTCF1-303 respectively. <sup>35</sup>S-labelled proteins were analysed by electrophoresis through 0.1% sodium dodecylsulphate (SDS)-12% polyacrylamide gels before visualization and quantification by autoradiography and phosphorimaging (BAS 1500 Phosphorimager and TINA 2.08e software, Fuji).

### Gel retardation analysis

Gel retardation assays were performed as described previously (Sharrocks *et al.*, 1993b) using 5% polyacrylamide

gels cast in 0.5×Tris-borate-EDTA (TBE). DNA binding sites were derived from annealed synthetic oligonucleotides (E74 and SRE) (Shore and Sharrocks, 1994) or from cloned sites from binding site-selection procedures (Shore and Sharrocks, 1995; Brown *et al.*, 1998a). Individual DNA-binding sites derived from the site-selection procedure were synthesized and labelled by PCR as described previously (Shore and Sharrocks, 1995). Binding reactions were set up to achieve <50% binding of the free DNA to allow quantification of relative DNA binding affinities. To account for slight differences in the amounts of each binding site added in each experiment, results were normalized by dividing by the total DNA present in the binding reaction.

#### Cell culture, transfection and extract preparation

COS-7 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS) (Gibco BRL). CHO cells were maintained in F12 medium supplemented with 5% foetal bovine serum (FBS) (Gibco BRL). Transfection experiments were carried out using Superfect transfection reagent (Qiagen) as described previously (Yang *et al.*, 1998a).

For reporter gene assays, GAL4-driven promoters were co-transfected with various vectors encoding GAL4-fusion proteins. The activities of the GAL4 DNA binding domain (amino acids 1–147), and GAL4-fusion proteins (50 ng of plasmid DNA) were measured in cotransfection assays in all cell lines using 1 µg of reporter plasmid pG5E1bLuc. Transfection efficiencies were normalized by measuring the activity from a cotransfected plasmid (0.5 µg) which expresses β-galactosidase (pCH110, Pharmacia KB Biotechnology Inc.). Cell extracts, were prepared and luciferase and β-galactosidase assays were carried out as described previously (Yang *et al.*, 1998a).

#### Protein kinase assays

Recombinant JNK and p38 MAP kinases were prepared as described previously (Yang *et al.*, 1998b). Recombinant

active ERK2 was obtained from New England Biolabs (NEB). The kinase assays were routinely carried out in 20 µl reaction volumes using GST-fusion proteins as described previously (Yang *et al.*, 1998a). Full-length purified recombinant His-tagged zTCF-1 was used as a substrate for subsequent use in gel retardation assays. The phosphorylation of substrate proteins was examined following SDS-PAGE by autoradiography, and quantified by phosphorimaging (Fuji BAS1500; TINA 2.08e software). Peptide competition experiments were carried out essentially as described above except for preincubation of 50–5000 pmoles of the peptide competitor (N-KGRKPRDLELPLSPSLLGGPGPE-C) with MAP kinases before the kinase reactions. Final peptide concentrations were 5–50 µM (10–100-fold excess over GST-TCF substrates). Binding and phosphorylation assays were carried out as described previously (Yang *et al.*, 1998a).

#### Acknowledgements

We thank Margaret Bell and Katherine Stewart for excellent technical and secretarial assistance, Bob Liddell for DNA sequencing and oligonucleotide synthesis. We are grateful to members of our laboratories for stimulating discussions and comments on the manuscript. We are also grateful to Alan Whitmarsh and Roger Davis for reagents. This work was supported by grants from the North of England Cancer Research Campaign, Cancer Research Campaign [CRC], the Wellcome Trust, a Newcastle University Research Committee Studentship to LA Brown and a Jeffcock Studentship to A Galanis. AD Sharrocks is a Research Fellow of the Lister Institute of Preventative Medicine.

#### Accession number

The Accession number for the zTCF-1 cDNA sequence is AC, AJ249170; ID, DRE249170.

#### References

- Brown AL, Amores A, Schilling TF, Jowett T, Baert J-L, de Launoit Y and Sharrocks AD. (1998a). *Oncogene*, **17**, 93–104.
- Brown AL, Schilling TF, Rodaway A, Jowett T, Ingham PW, Patient R and Sharrocks AD. (1998b). *Mech. Dev.* (Submitted).
- Cahill MA, Janknecht R and Nordheim A. (1996). *Curr. Biol.*, **6**, 16–19.
- Cano E, Hazzalin CA, Kardalidou E, Buckle RS and Mahadevan LC. (1995). *J. Cell. Science*, **108**, 3599–3609.
- Dalton S and Treisman R. (1992). *Cell*, **68**, 597–612.
- Dickens M, Rogers J, Cavanagh J, Raitano A, Xia Z, Halpern JR, Greenberg ME, Sawyers C and Davis RJ. (1997). *Science*, **277**, 693–696.
- Enslin H, Raingeaud J and Davis RJ. (1998). *J. Biol. Chem.*, **273**, 1741–1748.
- Gille H, Kortenjann M, Thomae O, Moomaw C, Slaughter C, Cobb MH and Shaw PE. (1995a). *EMBO J.*, **14**, 951–962.
- Gille H, Sharrocks AD and Shaw PE. (1992). *Nature*, **358**, 415–417.
- Gille H, Strahl T and Shaw PE. (1995b). *Curr. Biol.*, **5**, 1191–1200.
- Giovane A, Pintzas A, Maira S-M, Sobieszcuk P and Wasyluk B. (1994). *Genes Dev.*, **8**, 1502–1513.
- Guan K and Dixon JE. (1991). *Anal. Biochem.*, **192**, 262–267.
- Hill CS, Marais R, John S, Wynne J, Dalton S and Treisman R. (1993). *Cell*, **73**, 395–406.
- Janknecht R and Hunter T. (1997). *EMBO J.*, **16**, 1620–1627.
- Janknecht R, Zinck R, Ernst WH and Nordheim A. (1994). *Oncogene*, **9**, 1273–1278.
- Lopez M, Oettgen P, Akbarali Y, Dendorfer U and Liberman TA. (1994). *Mol. Cell. Biol.*, **14**, 3292–3309.
- Magnaghi-Jaulin L, Masutani H, Lipinski M and Harel-Bellan A. (1996). *FEBS Lett.*, **391**, 247–251.
- Maira S-M, Wurtz J-M and Wasyluk B. (1996). *EMBO J.*, **15**, 5849–5865.
- Plant KE, Hair A and Morgan GT. (1996). *Nucl. Acids Res.*, **24**, 3514–3521.
- Price MA, Cruzalegui FH and Treisman R. (1996). *EMBO J.*, **15**, 6552–6563.
- Price MA, Rogers AE and Treisman R. (1995). *EMBO J.*, **14**, 2589–2601.
- Raingeaud J, Gupta S, Rogers JS, Dickens M, Han J, Ulevitch RJ and Davis RJ. (1995). *J. Biol. Chem.*, **270**, 7420–7426.
- Rao VN, Huebner K, Isobe M, Ar-rusida A, Croce CM, Reddy ESP. (1989). *Science*, **244**, 66–70.
- Rao VN and Reddy ESP. (1994). *Oncogene*, **9**, 1855–1860.
- Seth A, Gonzalez FA, Gupta S, Raden DL and Davis RJ. (1992). *J. Biol. Chem.*, **267**, 24796–24804.

- Sgambato V, Vanhoutte P, Pagès C, Rogard M, Hippskind R, Besson M-J and Caboche J. (1998). *J. Neuroscience*, **18**, 214–226.
- Sharrocks AD. (1995). *FEBS Lett.*, **368**, 77–80.
- Sharrocks AD, Brown AL, Ling Y and Yates PR. (1997). *Int. J. Biochem. Cell Biol.*, **29**, 1371–1387.
- Sharrocks AD, Gille H and Shaw PE. (1993b). *Mol. Cell. Biol.*, **13**, 123–132.
- Sharrocks AD, Von Hesler F and Shaw PE. (1993a). *Nucl. Acids Res.*, **21**, 215–221.
- Shore P and Sharrocks AD. (1994). *Mol. Cell. Biol.*, **14**, 3283–3291.
- Shore P and Sharrocks AD. (1995). *Nucl. Acids Res.*, **23**, 4698–4706.
- Shore P, Bisset L, Lakey J, Waltho JP, Virden R and Sharrocks AD. (1995). *J. Biol. Chem.*, **270**, 5805–5811.
- Shore P, Whitmarsh AJ, Bhaskaran R, Waltho J, Davis RJ and Sharrocks AD. (1996). *Mol. Cell. Biol.*, **16**, 3338–3349.
- Strahl T, Gille H and Shaw PE. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 11563–11568.
- Treisman R. (1994). *Curr. Opin. Genet. Dev.*, **4**, 96–101.
- Treisman R. (1996). *Curr. Opin. Cell. Biol.*, **8**, 205–215.
- Whitmarsh AJ and Davis RJ. (1996). *J. Mol. Med.*, **74**, 589–607.
- Whitmarsh AJ, Shore P, Sharrocks AD and Davis RJ. (1995). *Science*, **269**, 403–407.
- Whitmarsh AJ, Yang S-H, Su MS-S, Sharrocks AD and Davis RJ. (1997). *Mol. Cell. Biol.*, **17**, 2360–2371.
- Yang S-H, Whitmarsh AJ, Davis RJ and Sharrocks AD. (1998b). *EMBO J.*, **17**, 1740–1749.
- Yang S-H, Yates P, Whitmarsh AJ, Davis RJ and Sharrocks AD. (1998a). *Mol. Cell. Biol.*, **18**, 710–720.
- Zhao L-J and Narayan O. (1993). *Gene*, **137**, 345–346.