

Determinants of T box protein specificity

Frank L. Conlon*, Lynne Fairclough, Brenda M. J. Price, Elena S. Casey[‡] and J. C. Smith^{§,¶}

Division of Developmental Biology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

We dedicate this paper to the memory of our friend and colleague Rosa Beddington

*Present address: Department of Genetics, Fordham Hall, UNC-Chapel Hill, Chapel Hill, NC 27599-3280, USA

[‡]Present address: Department of Molecular Biology, University of California, Berkeley, CA 94720, USA

[§]Present address: Wellcome CRC Institute, Tennis Court Road, Cambridge CB2 1QR, UK

[¶]Author for correspondence (e-mail: jim@welc.cam.ac.uk)

Accepted 6 July 2001

SUMMARY

Members of the T box family of transcription factors play important roles in early development. Different members of the family exert different effects and here we show that much of the specificity of the *Xenopus* T box proteins Xbra, VegT and Eomesodermin resides in the DNA-binding domain, or T box. Binding site selection experiments show that the three proteins bind the same core sequence, but they select paired sites that differ in their orientation and

spacing. Lysine 149 of Xbra is conserved in all Brachyury homologues, while the corresponding amino acid in VegT and Eomesodermin is asparagine. Mutation of this amino acid to lysine changes the inductive abilities of VegT and Eomesodermin to resemble that of Xbra.

Key words: T box genes, Brachyury, VegT, Eomesodermin, *Xenopus*

INTRODUCTION

Members of the T box family of transcription factors are required for formation of the basic vertebrate body plan and for normal development of organs such as the heart and limbs (reviewed by Kavka and Green, 1997; Papaioannou and Silver, 1998; Smith, 1999). T box genes are also implicated in human congenital malformations such as Holt-Oram syndrome (Basson et al., 1997; Li et al., 1997), ulnar-mammary syndrome (Bamshad et al., 1997; He et al., 1999) and DiGeorge syndrome (Jerome and Papaioannou, 2001; Lindsay et al., 2001; Merscher et al., 2001), and TBX2 proves to be amplified in a subset of human breast cancers (Jacobs et al., 2000). The founder member of the family, *Brachyury*, or *T*, encodes a sequence-specific DNA-binding protein that functions as a transcription activator (Conlon et al., 1996; Herrmann et al., 1990; Kispert and Herrmann, 1993; Kispert et al., 1995a). In mouse, *Xenopus*, zebrafish and chick embryos, *Brachyury* is expressed throughout the nascent mesoderm and transcripts are then restricted to the tailbud and notochord (Kispert et al., 1995b; Schulte-Merker et al., 1992; Smith et al., 1991; Wilkinson et al., 1990). Lack of *Brachyury* function, whether through genetic mutation in mouse (Chesley, 1935; Gluecksohn-Schoenheimer, 1938; Herrmann et al., 1990) and zebrafish (Halpern et al., 1993; Schulte-Merker et al., 1994), or by inhibiting the ability of the protein to activate transcription in *Xenopus* (Conlon et al., 1996), causes loss of posterior mesodermal structures and impairment of notochord differentiation. Furthermore, mis-expression of *Brachyury* in prospective ectodermal tissue of the *Xenopus* embryo causes those cells to activate mesoderm-specific genes and to form

mesodermal cell types such as muscle (Cunliffe and Smith, 1992; Cunliffe and Smith, 1994; O'Reilly et al., 1995). Together, these experiments indicate that *Brachyury* is both necessary and sufficient for normal mesoderm formation.

The first clue that *Brachyury* is a member of a family of proteins came from the observation that the DNA-binding domain of the protein (now referred to as the T box) shows extensive sequence homology with the product of the *Drosophila* gene *optomotor-blind* (Pflugfelder et al., 1992). Since then, over 50 such T box genes have been identified throughout the animal kingdom, and they prove to be expressed in, and to play roles in the development of, multiple cell types (see reviews cited above). Of the many issues raised by this work, one of the most important concerns the question of T box specificity. This is illustrated by results obtained with *Tbx4* and *Tbx5*, two of the most closely related members of the T box family. *Tbx4* is expressed at high levels in the hindlimb of the developing vertebrate embryo and *Tbx5* in the forelimb (Gibson-Brown et al., 1998; Isaac et al., 1998; Logan et al., 1998; Ohuchi et al., 1998). Mis-expression experiments suggest, remarkably, that limb identity is determined by which of the two T box genes is expressed in the developing limb bud (Logan and Tabin, 1999; Rodriguez-Esteban et al., 1999; Takeuchi et al., 1999). How do the different T box proteins exert these different effects?

In this paper, we address the question of T box specificity by studying three genes expressed during early *Xenopus* development: *Xenopus Brachyury* (*Xbra*) (Smith et al., 1991), *Eomesodermin* (Ryan et al., 1996) and *VegT/Antipodean* (Horb and Thomsen, 1997; Lustig et al., 1996; Stennard et al., 1996; Zhang and King, 1996). All three genes are expressed in the

mesoderm of the early gastrula and the function of each is required for proper patterning of the *Xenopus* embryo, with VegT likely to act both maternally and zygotically (Conlon et al., 1996; Horb and Thomsen, 1997; Ryan et al., 1996; Stennard et al., 1999; Zhang et al., 1998). The genes are also necessary for the normal development of other vertebrate species, including mouse and fish (Chesley, 1935; Gluecksohn-Schoenheimer, 1938; Halpern et al., 1993; Herrmann et al., 1990; Russ et al., 2000; Schulte-Merker et al., 1994).

Like Xbra, VegT and Eomesodermin are transcription activators and are capable of activating mesoderm-specific genes in isolated animal pole tissue (this work; Horb and Thomsen, 1997; Ryan et al., 1996; Tada et al., 1998). However, the types of mesoderm induced by each T box protein differ. In particular, Xbra induces posterior mesodermal cell types and activates posteriorly expressed genes while VegT and Eomesodermin can induce virtually the entire spectrum of mesodermal genes and of mesodermal cell types. In this study, we have used a series of chimeric proteins to investigate the basis of this inductive specificity. Our results show that much of the specificity resides within the T boxes of the proteins, but also that the C-terminal region of Xbra is capable of restricting the inductive abilities of the VegT and Eomesodermin T boxes.

The different inducing activities of Xbra, VegT and Eomesodermin suggest that the proteins might recognise different DNA target sequences. To address this question, we have carried out a series of binding site selection experiments. All three proteins prove to recognise the same core sequence of TCACACCT with some differences in flanking nucleotides. Significantly, however, further rounds of selection tend to select repeats of the core sequence, and the spacing and orientation of the repeats are different for each protein. For example, as reported by Kispert and Herrmann (Kispert and Herrmann, 1993), Brachyury selects the palindromic sequence TCACACCTAGGTGTGA while Eomesodermin frequently selects two direct repeats of the core motif separated by four nucleotides. It is possible that differences such as these underlie the different effects of the different T box proteins. Finally, we show that at least some aspects of specificity are associated with an asparagine residue in the T boxes of VegT and Eomesodermin; mutation of this residue to the lysine present in the equivalent position in Brachyury causes the two proteins to behave more like Xbra.

MATERIALS AND METHODS

Plasmid constructs and RNA synthesis

VegT was a gift from Mary Lou King (Zhang and King, 1996) and Eomesodermin was a gift from John Gurdon (Ryan et al., 1996). pSP64T-Xbra (Cunliffe and Smith, 1992) and pSP64T-Xbra-HA (Tada et al., 1997) have been described previously. The analogous pSP64T-VegT-HA and pSP64T-Eomesodermin-HA constructs were created by PCR; details are available on request. For T box VP16 fusions, amino acids 1-147 of yeast GAL4 were first fused in frame to the T boxes of VegT (amino acids 47-238), Eomesodermin (amino acids 210-469) or Xbra (amino acids 17-227). Each construct was then fused to the transcriptional activation domain of VP16 (amino acids 413-454) via a lambda linker (Brickman et al., 2000). Constructs were cloned into pSVGVP1 for transient transfections and pGEM-3Zf (Promega) for RNA injections.

For T box 'swap' constructs (see Fig. 4) VXV and XEX were generated by replacing the T box of Xbra with that of VegT or Eomesodermin, respectively. VXV was generated by replacing the T box of VegT with that of Xbra. Truncations of Xbra, VegT and Eomesodermin (Fig. 2) occurred at amino acids 232, 375 and 578, respectively. Cloning details are available on request. Constructs were cloned into pcDNA3.1 (InVitrogen) for transient transfections and pCR2.1 (InVitrogen) for RNA injections.

Point mutations in Eomesodermin and VegT were generated by PCR. For both proteins, an asparagine residue in the T box (N155 in VegT and N353 in Eomesodermin) was changed to lysine, the amino acid present in the corresponding position in Xbra. Cloning details are available on request. Constructs were cloned into pcDNA3.1 (InVitrogen) for transient transfections and pCR2.1 (InVitrogen) for RNA injections.

All constructs were sequenced and gave proteins of the correct size after *in vitro* translation (data not shown). RNA from each construct was generated as described (Smith, 1993).

Embryos, microinjection and dissection

Xenopus embryos were obtained by *in vitro* fertilisation (Smith and Slack, 1983). They were maintained in 10% Normal Amphibian Medium (NAM: Slack, 1984) and staged according to Nieuwkoop and Faber (Nieuwkoop and Faber, 1975). *Xenopus* embryos were injected at the one-cell stage with 0.5 ng RNA in 10 nl water. For animal cap assays embryos were dissected in 75% NAM, and caps were cultured in the same medium until early gastrula stage 10.

RNA isolation and RNAase protection assays

RNAase protection assays were carried out as described (Jones et al., 1995). Each RNAase protection shown is representative of at least two independent experiments. Probes were as follows: *Xbra* (Smith et al., 1991), *Xwnt11* (Ku and Melton, 1993), *Bix4* (Tada et al., 1998), *goosecoid* (Cho et al., 1991), *chordin* (Sasai et al., 1994), *Xwnt8* (Christian et al., 1991; Smith and Harland, 1991), *Mix.1* (Rosa, 1989), *Pintallavis* (Ruiz i Altaba and Jessell, 1992) and *Xsox17α* (Hudson et al., 1997).

DNA gel-shift assays

Proteins used in electrophoretic mobility shift assays (EMSA) were prepared from DNA using the TNT *in vitro* translation kit (Promega). Binding reactions contained 1 µl of *in vitro* translated protein, 1× buffer and 20,000 cpm probe in a total volume of 12 µl. Control reactions (data not shown) contained a 100-fold excess of unlabelled specific or nonspecific oligonucleotide. The 1× buffer was either (i) 50 mM KCl, 1 mM EDTA, 20 mM Hepes pH 7.9, 10% glycerol, 100 µg/ml bovine serum albumin (BSA), 1 mM DTT, 0.3 mM PMSF plus Roche Complete minitabs protease inhibitors; or (ii) 60 mM KCl, 15 mM Tris pH 7.5, 7.5% glycerol, 250 µg/ml BSA, 0.05% NP40, 1 mM DTT, 4 mM spermine, 4 mM spermidine and protease inhibitors as above. Complexes were allowed to form at room temperature for 15-20 minutes after addition of probe. Oligonucleotides used in EMSA were annealed for 10 minutes at 88°C and cooled slowly to room temperature; they were then labelled by 3' filling with ³²P-dCTP (3,000 Ci/mmol) using the Klenow fragment (Promega).

Transient transfection analyses

Transient transfection assays were carried out as described (Conlon et al., 1996). Effector constructs are described above. The CAT reporter construct pBLCAT2 (Luckow and Schutz, 1987) was modified such that the sequence TTTCACACCT was inserted upstream of the promoter region (Fig. 2). *MLVlacZ* was co-transfected as a control for transfection efficiency (Hill et al., 1993).

PCR binding site selection assays

Binding site selection was carried out as described (Pollock and Treisman, 1990) using *in vitro* translated protein from pSP64T-

Xbra-HA, pSP64T-VegT-HA or pSP64T-Eomesodermin-HA. DNA fragments obtained after five or seven rounds of selection were PCR amplified and cloned into the vector MP19. After five rounds, 62 sequences were examined for Xbra, 60 sequences for VegT and 61 sequences for Eomesodermin. After seven rounds, the numbers were 97, 64 and 63, respectively. Previous work has shown that the sequence TCACACCT interacts with T box proteins (Casey et al., 1998; Casey et al., 1999; Kispert and Herrmann, 1993; Tada et al., 1998), and this motif, or variations of it, was observed in all the selected DNA fragments. Further analysis was carried out manually. This revealed that after seven rounds of selection some of the sequenced clones were identical, such that the numbers of different clones studied for Xbra, VegT and Eomesodermin were 92, 42 and 38, respectively.

RESULTS

Different effects of Xbra, VegT and Eomesodermin

Past studies suggest that the T box genes *Xbra*, *VegT* and *Eomesodermin* (Fig. 1A), all of which are expressed in the marginal zone of the *Xenopus* early gastrula (Fig. 1B-D), have different mesoderm-inducing activities. For example, *VegT* and *Eomesodermin* can induce expression of dorsoanterior markers such as *gooseoid*, while *Xbra* cannot (Cunliffe and Smith, 1992; Cunliffe and Smith, 1994; O'Reilly et al., 1995; Ryan et al., 1996). To confirm this finding, we have dissected animal pole regions from embryos previously injected with RNA encoding Xbra, VegT or Eomesodermin, cultured these animal caps to the equivalent of the early gastrula stage, and assayed them for expression of a panel of mesodermal and endodermal markers. Our results confirm that *Xbra* activates its own expression (data not shown), and that of *Xwnt11* and *Bix4*, but cannot induce *gooseoid*, *chordin*, *Xwnt8* or *Mix.1*, and it induces *Pintallavis* and *Xsox17 α* only weakly (Fig. 1E). By contrast, *VegT* and *Eomesodermin* induce the expression of all markers tested (Fig. 1E).

These differences between the T box proteins appear to be qualitative rather than quantitative. We have found no concentration of *Xbra* RNA, for example, that can induce expression of *gooseoid* (data not shown, but see Cunliffe and Smith, 1992; Cunliffe and Smith, 1994; O'Reilly et al., 1995; Tada et al., 1997).

Xbra, VegT and Eomesodermin are transcriptional activators

The results described above show that the inductive effects of Xbra differ from those of VegT and Eomesodermin. As a first step towards understanding these differences, we sought to confirm, as would be inferred from previous work (Casey et al., 1999; Conlon et al., 1996; Horb and Thomsen, 1997; Ryan et al., 1996; Zhang and King, 1996), that all three T box proteins function as transcription activators. To this end, plasmids encoding Xbra, VegT or Eomesodermin were transfected into COS cells along with a reference plasmid and a reporter construct in which the T box binding site derived from the *eFGF* promoter, TTTCACACCT (Casey et al., 1998), is positioned upstream of a minimal promoter that drives chloramphenicol acetyl transferase (CAT). All three gene products activate CAT activity (Fig. 2). Levels of activation differ between the three T box proteins, but no significance can be attached to this observation at present

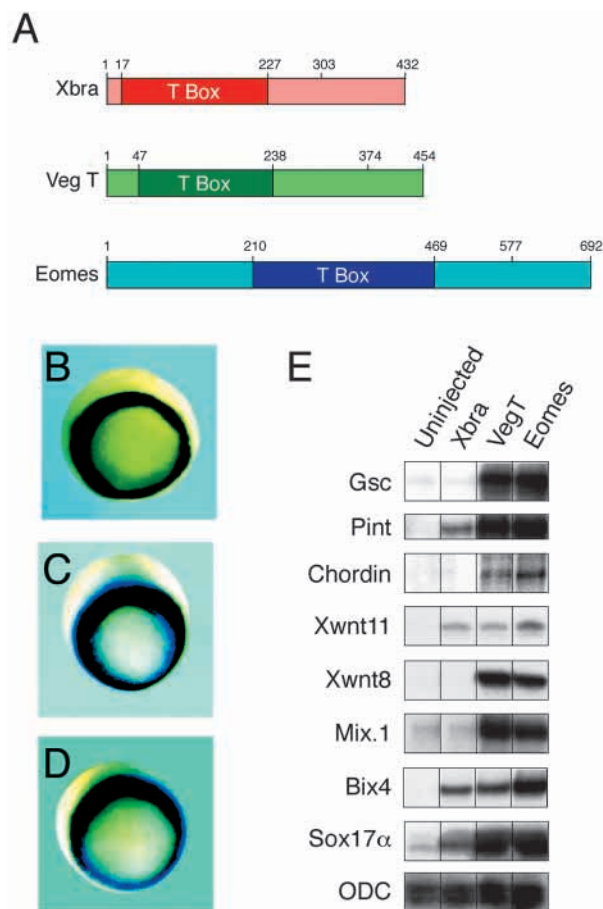


Fig. 1. The T box proteins Xbra, VegT and Eomesodermin are expressed in similar patterns but induce the expression of different genes. (A) The structures of Xbra, VegT and Eomesodermin. Note that Eomesodermin has a larger N-terminal domain than Xbra or VegT. (B) Expression of *Xbra* at the early gastrula stage analysed by whole-mount in situ hybridisation. (C) Expression of *VegT* at the early gastrula stage. (D) Expression of *Eomesodermin* at the early gastrula stage. (E) Different inducing properties of Xbra, VegT and Eomesodermin. Note that all three T box proteins induce expression of *Wnt11* and *Bix4*, that VegT and Eomesodermin induce higher levels of *Pintallavis* and *Sox17 α* than does Xbra, and that Xbra cannot activate *Gooseoid*, *chordin*, *Xwnt8* or *Mix.1*. The expression domains of the marker genes are as follows: *Xwnt11*, pan-mesodermal (Tada and Smith, 2000); *Bix4*, pan-mesendodermal (Casey et al., 1999; Tada et al., 1998); *Pintallavis*, dorsal mesoderm (Ruiz i Altaba and Jessell, 1992); *Sox17 α* , endodermal (Hudson et al., 1997); *Gooseoid*: dorsal mesoderm (Cho et al., 1991); *chordin*: dorsal mesoderm (Sasai et al., 1994); *Xwnt8*, ventral and lateral mesoderm (Christian et al., 1991; Smith and Harland, 1991); and *Mix.1*, pan-mesendodermal (Rosa, 1989).

because their levels of expression and affinities for the target site may differ.

The activation domain of Xbra is contained within the C-terminal half of the protein (Conlon et al., 1996; Kispert et al., 1995a), and removal of the C termini of Eomesodermin and VegT demonstrated that the same is true of these proteins, although VegT did retain some activity (Fig. 2). It is unlikely that the loss of transcriptional activation is due to instability of the truncated proteins, or to loss of a nuclear localisation

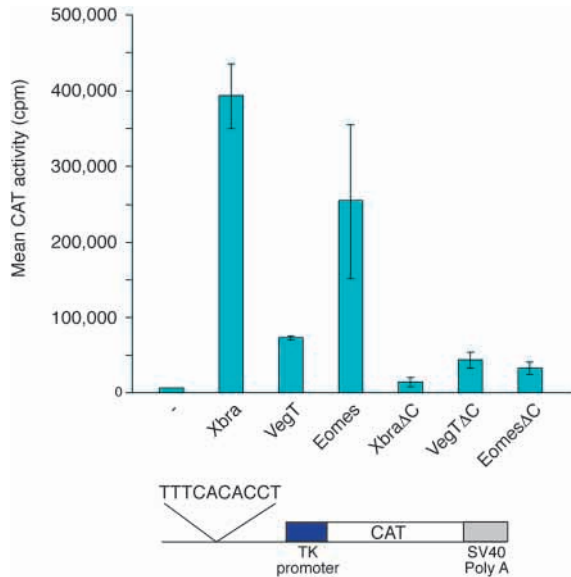


Fig. 2. Xbra, VegT and Eomesodermin can function as transcriptional activators. Expression constructs encoding Xbra, VegT or Eomesodermin, or truncated versions of the proteins (see Materials and Methods), were transfected into 3T3 cells along with a reporter plasmid in which the sequence TTTCCACACCT is placed upstream of a minimal promoter (below). All three T box proteins activated transcription, and activation was reduced in the truncated versions.

signal, because a similar truncated version of Xbra is both stable and nuclear (Walter Lerchner and JCS, unpublished work).

T box protein specificity resides in part in the T box

The different inductive effects of Xbra, VegT and Eomesodermin (Fig. 1E) might derive from differences in the T boxes of these proteins or in domains outside of the T boxes. For example, the proteins might activate different genes because their T boxes bind different DNA motifs or they may do so because they recruit different accessory proteins via non T box sequences. To address this question we have created fusion proteins in which the T boxes of the three proteins are fused to the activation domain of VP16 (Fig. 3A). The fusion proteins also contain, at their N termini, the GAL4 nuclear localisation signal; nuclear localisation of Xbra, and perhaps other T box proteins, requires amino acids within the C terminal half of the protein, which has been removed in these experiments (Kispert et al., 1995a). As predicted, all three VP16 constructs behaved as powerful transcription activators when tested with a reporter construct containing the *eFGF* T box binding site (data not shown).

The inductive effects of the three VP16 constructs resembled those of their parent proteins. For example, Xbra cannot induce expression of *gooseoid* or *chordin*, and nor can Xbra-VP16. VegT and Eomesodermin, however, can induce these genes, and so can VegT-VP16 and Eomesodermin-VP16 (Fig. 3C,D). We note that Xbra-VP16 induces higher levels of expression of *Pintallavis*, *Xwnt11* and *Bix4* than does Xbra (Fig. 3B). This suggests that the VP16 activation domain has stronger activity than the endogenous Xbra activation domain, and it reinforces the view that the inability of wild-type Xbra to activate

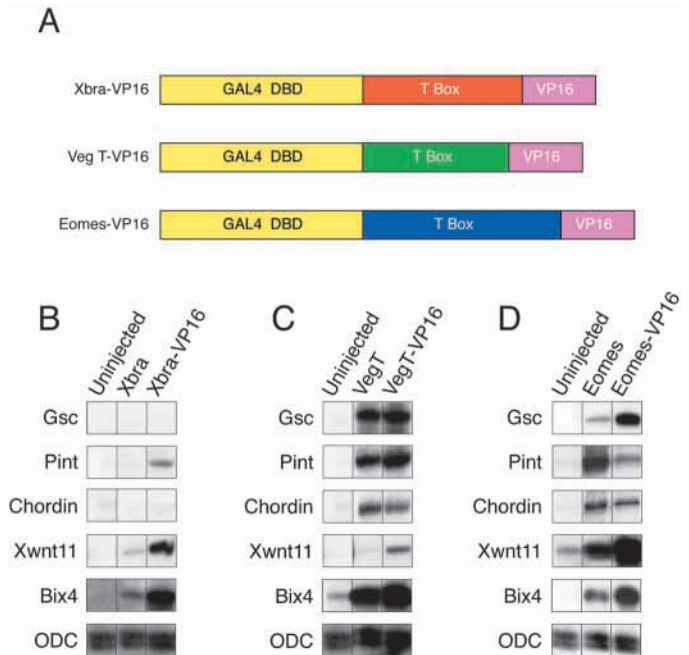


Fig. 3. Xbra, VegT and Eomesodermin specificity resides mainly in the T box. (A) Chimeric proteins comprising the T boxes of Xbra, VegT and Eomesodermin fused to the GAL4 DNA-binding domain and the VP16 activation domain. Proteins were expressed in early *Xenopus* embryos and their abilities to activate gene expression were assessed by RNAse protection. (B) Xbra and Xbra-VP16. (C) VegT and VegT-VP16. (D) Eomesodermin and Eomesodermin-VP16. Note that the inductive abilities of the chimeric constructs resemble those of the parent molecule.

expression of *gooseoid* represents a qualitative difference between Xbra and the other T box proteins, and that the structural basis of this difference resides in the T box.

The Xbra C-terminal domain restricts the activation of target genes

An alternative explanation for the observation that Xbra-VP16 is a more potent activator of target genes than is Xbra, is that the C-terminal domain of Xbra somehow restricts target gene activation. To investigate this possibility, we placed the T boxes of VegT and Eomesodermin within the backbone of Xbra, thereby creating XVX and XEX, respectively (see Fig. 4A). Our reasoning was that non T box sequences of Xbra might restrict the activation of VegT and Eomesodermin target genes such as *gooseoid*, *Pintallavis* and *chordin*. Induction of these genes by the two chimeric proteins is indeed reduced, while activation of *Xwnt11* and *Bix4* is less affected (Fig. 4B). Thus, sequences outside the Xbra T box can restrict the activation of target genes. As might be predicted, insertion of the Xbra T box into VegT creates a protein whose inducing activity resembles that of Xbra-VP16, in that it cannot activate *gooseoid*, *Pintallavis* or *chordin* but can induce *Xwnt11* and *Bix4* (Fig. 4C).

Together, our results indicate that much of the biological specificity of the T box proteins Xbra, VegT and Eomesodermin resides within the T box, but that sequences outside the Xbra T box also restrict the activation of target genes.

Xbra, VegT and Eomesodermin bind the same core sequence but prefer double sites with different orientations and spacings

Much of the functional specificity of the T box proteins resides in their T boxes. It is possible that the different T boxes recognise different DNA sequences, and we have investigated this idea by carrying out PCR-based binding site selection experiments.

Binding site selection experiments were carried out essentially as described by Pollock and Treisman (Pollock and Treisman, 1990), using HA-tagged versions of Xbra, VegT and Eomesodermin. After five rounds of selection, we found that Xbra, VegT and Eomesodermin selected the same core sequence of TCACACCT with some differences in flanking nucleotides (Fig. 5). Of these differences, the most marked was the frequent selection by Xbra of a guanine nucleotide 5 bases

3' of the core sequence, and a concomitant preference for a T 5' of this guanine nucleotide and a T or a C 3' of the G (Fig. 5A and see Discussion). VegT and Eomesodermin had no preferred nucleotide at this position (Fig. 5B,C). However, we have been unable to design sequences that are specific for particular T box proteins in electrophoretic mobility shift assays.

Many of the sequences identified after five rounds of selection contained two core motifs. To quantitate this observation, we required that both motifs should contain at least six of the eight nucleotides of the core sequence TCACACCT. According to this criterion, double sites occurred in 14.5% of selected Xbra sequences, 38.5% of selected VegT sequences and 53.5% of Eomesodermin sequences. Double sites were observed much more frequently, however, after seven rounds of selection, with the corresponding figures being 39.2, 87.5 and 96.8%, respectively. Analysis of these sequences revealed very strong preferences for particular orientations and spacings of the two core sequences. In agreement with Kispert and Herrmann (Kispert and Herrmann,

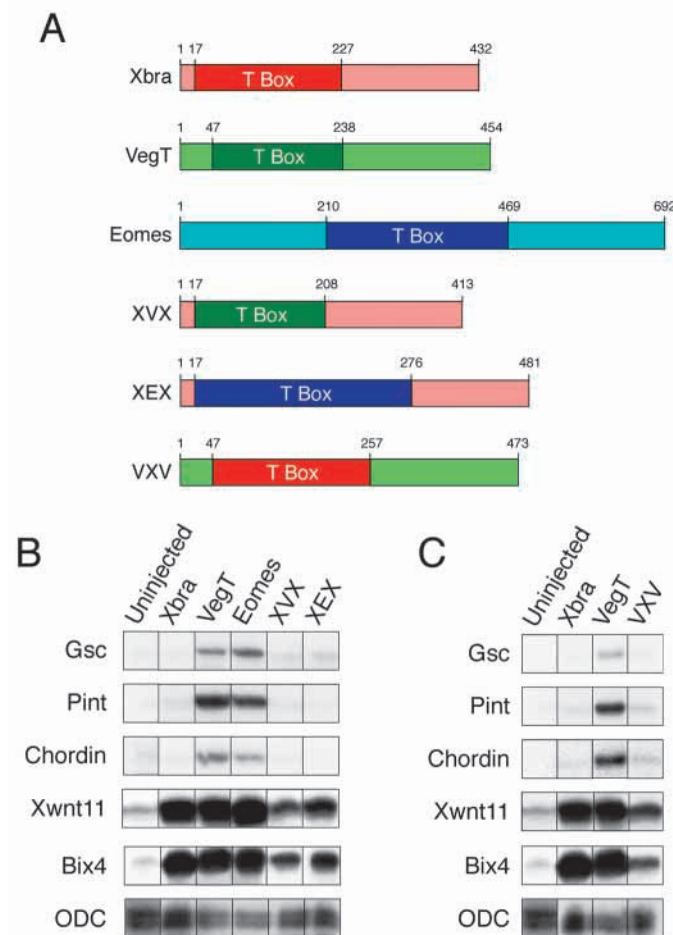


Fig. 4. Sequences outside the Xbra T box restrict induction by VegT and Eomesodermin. (A) The parent Xbra, VegT and Eomesodermin proteins, and chimeric versions thereof. XVX consists of the VegT T box surrounded by Xbra non-T box sequences; XEX contains the Eomesodermin T box surrounded by Xbra non-T box sequences; and VXV consists of the Xbra T box surrounded by VegT non-T box sequences. (B) Activation of *Gooseoid*, *Pintallavis* and *Chordin* by XVX and XEX is lower than activation of the same genes by the parent proteins, suggesting that the non-T box sequences of Xbra reduce levels of induction. (C) *Gooseoid*, *Pintallavis* and *Chordin* are not induced by a protein comprising the Xbra T box surrounded by VegT non-T box sequences.

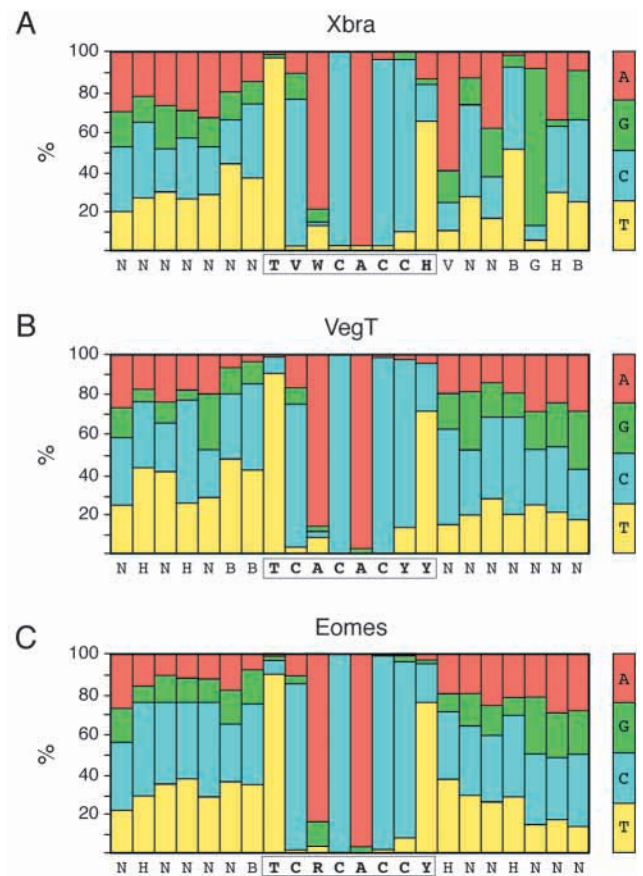


Fig. 5. Motifs selected by Xbra (A), VegT (B) and Eomesodermin (C) after five rounds of binding site selection. The consensus sequence is represented at the bottom of each histogram; if a nucleotide is present in greater than 10% of the selected sequences it is defined as being part of the consensus, and in this respect the sites selected by the three proteins differ. However, the core motif selected by all three T box proteins is clearly TCACACCT, and we have been unable to define sequences that are specific for a single member of the family. Note that Xbra shows a preference for a G positioned 5 nucleotides downstream of the core motif; see text for further details.

Table 1. Binding site selection data: round 7

Protein	Motif			
	→	→→	→←	←→
Xbra*	59	9	25	0
VegT	8	10	1	45
Eomes	2	27	33	1

*Three additional clones contained the motif →←→ and one contained →→→.

Table shows the frequency with which different types of motif were selected by each T box protein.

1993), double sites selected by Xbra are usually palindromic, with the two core sequences arranged in opposite orientations (Table 1) and with no intervening nucleotides (Table 2). Although double sites selected by VegT were also frequently palindromic, these sites are in the opposite orientations compared with those selected by Xbra (Table 1), and they are almost invariably separated by four nucleotides instead of being immediately juxtaposed (Table 2). Finally, sites selected by Eomesodermin are either in the same orientation as those observed with Xbra, or are arranged as direct repeats (Table 1). The spacing in the former case is usually four nucleotides, but three and five nucleotides are often observed. The spacing in the latter case is usually five nucleotides, but a four nucleotide spacing is also common.

The abilities of Xbra, VegT and Eomesodermin to interact with oligonucleotides containing one or two core motifs were investigated in electrophoretic mobility shift assays. Xbra, unlike VegT and Eomesodermin, interacted only very weakly with oligonucleotides containing just a single motif (data not shown). In this respect, it contrasts with proteins comprising just the Xbra T box, which interact strongly with a single motif (Casey et al., 1998). This apart, we were unable to demonstrate any specificity of the T box proteins for oligonucleotides containing just a single motif.

By contrast, electrophoretic mobility shift assays do suggest that the different T box proteins display preferences for different paired motifs. Typical results are presented in Fig. 6, and the data from over 20 such experiments are summarised in Table 3. The palindromic sequence selected by Xbra (→←) interacted only with Xbra and VegT, with optimum binding of Xbra occurring in the presence of EDTA (data not shown). The higher mobility of the VegT complex suggests that this T box protein may bind to the →← site almost exclusively as a monomer. By contrast, the ←NNNN→ sequence selected by VegT interacts with all three T box proteins, with the existence of lower-mobility forms of VegT and (to a much lesser extent) of Eomesodermin, suggesting that binding may occur as a dimer. It is surprising that Xbra recognises this site in

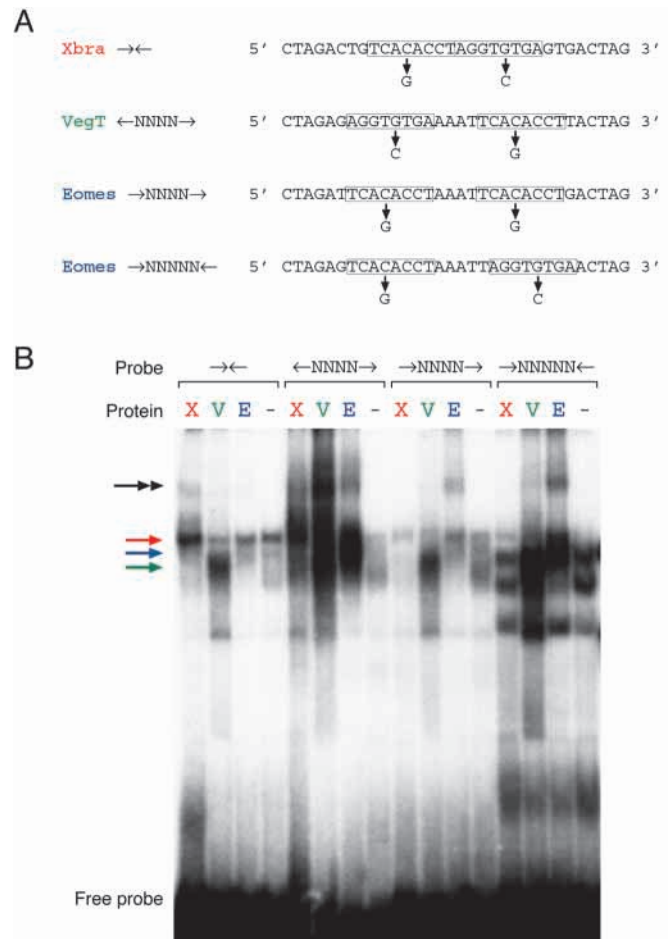


Fig. 6. Electrophoretic mobility gel shift assays demonstrate differences between different T box proteins to interact with different oligonucleotides. (A) Oligonucleotides used in electrophoretic gel mobility shift assays. Only one strand is shown and core motifs are boxed. Arrows indicate mutations in control oligonucleotides. Use of these oligonucleotides in electrophoretic gel mobility shift assays prevented binding (data not shown). (B) Band shift assay. Single-headed arrows indicate positions of high mobility complexes (red, Xbra; green: VegT, blue: Eomesodermin). Double-headed arrow indicates low mobility complexes.

electrophoretic mobility shift assays, because no ←NNNN→ sites were selected during the binding site selection procedure (Table 1). Finally, the two Eomesodermin sites, →NNNN→ and →NNNNN← do not bind Xbra but do interact with both VegT and Eomesodermin (Fig. 6). VegT interacts to form predominantly a high mobility complex, again suggesting that

Table 2. Spacing of clones

Motif	Base pairs between core sequences												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Xbra →←	25	0	0	0	0	0	0	0	0	0	0	0	0
VegT ←→	0	0	0	1	42	0	2	0	0	0	0	0	0
Eomes →→	0	0	1	7	8	6	0	0	0	0	0	0	5*
Eomes →←	0	0	0	1	11	16	1	4	0	0	0	0	0

*In these cases, the intervening 12 base pairs invariably contained part of the core T-box binding site.

Table shows the numbers of base pairs between core sequences selected by each T box protein.

Table 3. Specificity of different sites for different T-box proteins

Motif	Protein		
	Xbra	VegT	Eomesodermin
→←	+	+ (High)	-
←NNNN→	+	++ (Low)	++ (High)
→NNNN→	-	+ (High)	++ (Low)
→NNNN←	-	+ (High)	++ (Low)

–, no significant binding; +, moderate binding; ++, strong binding.
‘High’ refers to a high-mobility complex.
‘Low’ refers to a low-mobility complex.

it binds the Eomesodermin sites as a monomer. These results are summarised in Table 3.

Mutation of a single amino acid can change T box protein function

Our data show that the functional specificities of Xbra, VegT and Eomesodermin reside, in large part, in their T boxes. In an effort to identify amino acids that might determine specificity, we examined the sequences of the T boxes of Brachyury, VegT and Eomesodermin from a variety of species (Fig. 7A). Superposition of the VegT and Eomesodermin sequences onto the crystal structure of the Xbra T box (Fig. 7B) revealed only two predicted protein-DNA contact points, positions 149 and 214, at which the sequence of Brachyury differs from that of VegT or Eomesodermin. The amino acid substitution at position 214 is a conserved change replacing the alanine in Xbra with a glycine in VegT and Eomesodermin. However, the amino acid substitution at position 149 is a much more dramatic substitution, which replaces a basic lysine residue in Brachyury with the neutral polar residue asparagine in VegT and Eomesodermin. This residue comes at the end of a stretch of highly conserved amino acids that are predicted to form a pleated sheet structure. Lysine 149 is conserved in all Brachyury homologues and contacts the phosphate backbone of the DNA (Fig. 7B).

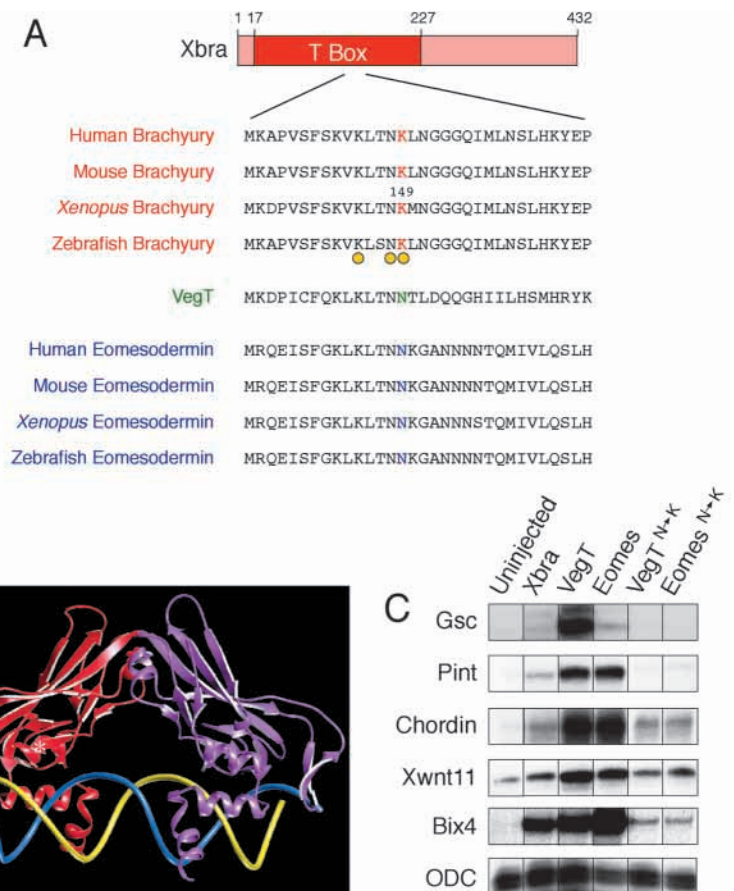
To investigate the significance of this amino acid in T box functional specificity, the asparagine

Fig. 7. Identification of lysine 149 as an amino acid which may confer T box specificity. (A) Diagram of Xbra and part of the T box sequences of human, mouse, *Xenopus* and zebrafish Brachyury. Yellow circles indicate amino acids which contact DNA (Muller and Herrmann, 1997). The sequences are aligned with the equivalent regions of *Xenopus* VegT and human, mouse, *Xenopus* and zebrafish Eomesodermin. Note that the K residue in Xbra is replaced by an N in VegT and Eomesodermin. (B) Asterisk (on the left-hand T box) marks the position of K149 on the crystal structure of the Xbra T box bound to DNA (Muller and Herrmann, 1997). (C) Mutation of an asparagine residue in VegT and Eomesodermin to lysine causes those proteins to resemble Xbra in their inducing activities; VegT^{N→K} and Eomesodermin^{N→K} lose the ability to induce *Gooseoid* and levels of *Chordin* and *Pintallavis* are significantly reduced.

of VegT and Eomesodermin was mutated to lysine. The effects of these mutant VegT and Eomesodermin constructs prove more to resemble those of Xbra (Fig. 7C), although we also note a general reduction in inducing activity. These results suggest that part of the specificity of T box proteins resides in K¹⁴⁹ of Xbra, whose equivalent residue in VegT and Eomesodermin is an asparagine.

DISCUSSION

T box proteins are transcription factors that control the specification and morphogenesis of many cell types during vertebrate and invertebrate development (reviewed by Kavka and Green, 1997; Papaioannou and Silver, 1998; Smith, 1999). In vertebrates, at least three members of the T box family – Brachyury, VegT and Eomesodermin – are involved in the induction and patterning of the mesoderm (Chesley, 1935; Gluecksohn-Schoenheimer, 1938; Griffin et al., 1998; Herrmann et al., 1990; Horb and Thomsen, 1997; Kimmel et al., 1989; Lustig et al., 1996; Russ et al., 2000; Ryan et al., 1996; Stennard et al., 1996; Zhang et al., 1998; Zhang and King, 1996). Although all three proteins contain T box domains and are expressed in the marginal zone of the embryo, previous studies and our present results show that they play different roles in mesoderm induction and patterning. Mis-expression of *Xbra* in animal pole explants induces expression of the mesodermal markers *Xwnt11* and *Bix4* but not markers of anterior or dorsal mesoderm such as *gooseoid*, *Pintallavis*



or *chordin*. By contrast, mis-expression of either VegT or Eomesodermin is able to induce expression of all these markers. We have used this observation (Fig. 1) as the basis of an in vivo assay to identify determinants of T box specificity.

T box specificity resides in large part in the T box

Our experiments show that all three T box proteins function as activators of transcription (Fig. 2). We have taken advantage of this observation to construct chimeric proteins comprising the Xbra, VegT or Eomesodermin T box fused to the VP16 activation domain. Expression of these constructs in *Xenopus* embryos reveals that the specificity of the three proteins resides in the T box (Fig. 3).

One significant qualification to this conclusion is that sequences outside of the Xbra T box restrict the inducing activities of VegT and Eomesodermin (Fig. 4). The mechanism by which this occurs is unknown, but it may be significant that full-length Xbra binds DNA rather poorly, while the T box domain alone binds strongly (see below and Casey et al., 1998).

DNA binding specificity

To investigate the molecular basis of T box specificity, binding site selection experiments were carried out. As described above, after five rounds of selection all three proteins selected predominantly single sites, defined by the core motif TCACACCT. This represents half of the palindromic sequence previously identified by Kispert and Herrmann (Kispert and Herrmann, 1995a). There were no dramatic differences between the sequences selected by the three proteins, save the frequent selection by Xbra of a G positioned 5 nucleotides 5' of the core motif (Fig. 5). The significance of this observation is not clear, although it may represent the first step towards the selection of a palindromic sequence: the $\rightarrow\leftarrow$ sequence selected by Xbra contains a G at the same position relative to the first core motif (Fig. 6A). Consistent with this suggestion, we observe that in single Xbra sites such G residues are frequently flanked (in 26% of cases) by two Ts, creating the triplet TGT, which is also present in the palindromic sequence selected by Xbra. In addition, we note that 23% of the G residues are flanked by T and C, giving the triplet TGC. If these observations do provide a clue as to the preference of Xbra for particular DNA sequences, they suggest that the G positioned five nucleotides downstream of the core motif are particularly important, followed by a 3' T and then a 5' T. This suggestion does not explain, however, the frequent occurrence (28% of cases) of the triplet CGA; the interaction of Xbra with DNA clearly requires further study.

A further two rounds of selection resulted in the isolation of a large number of paired T box binding motifs. The results of these experiments are summarised in Tables 2 and 3, which show that different T box proteins prefer different types of paired motifs and suggest that they bind some sites as dimers and some as a monomer. For example, VegT appears to bind the two sites selected by Eomesodermin (\rightarrow NNNN \rightarrow and \rightarrow NNNN \leftarrow) as a monomer, while Eomesodermin appears to bind as a dimer.

These observations provide a basis for T box protein specificity, and it will be of great interest to elucidate the structures of Xbra, VegT and Eomesodermin on their respective sites. We note, however, that the enhancer of no natural T box target gene has yet proved to contain the motifs

summarised in Table 3. For example, the enhancers of the *Xenopus* genes *eFGF* and *Xnr1* contain just the motif TCACACCT (Casey et al., 1998; Hyde and Old, 2000), and although *Bix4* contains three tandem motifs TGACACCT, TCACACCT and TCACACGT, the spacings between the motifs are 16 and nine nucleotides respectively (Tada et al., 1998).

T box target genes have also been identified in *Ciona intestinalis*, where the tropomyosin-like gene responds directly to Brachyury (Di Gregorio and Levine, 1999). Here, three Brachyury recognition sequences have been identified, one of which (Ci-Bra #3) is identical to the sequence identified in the enhancers of *Xenopus eFGF* and *Xnr1*. The other two, Ci-Bra Prox and Ci-Bra Dist, comprise two motifs, with the proximal element arranged as inverted repeats and the distal element arranged as tandem repeats. In neither case, however, do the motifs correspond exactly to the sequences isolated in our binding site selection experiments or those of Kispert and Herrmann (Kispert and Herrmann, 1993). Additional experiments are necessary to define the extent to which T box proteins can tolerate departures from the 'perfect' sites.

Finally, we note that the properties of the T box proteins Brachyury, TBX1 and TBX2 have recently been studied (Sinha et al., 2000). TBX1, like Brachyury, binds DNA as a dimer, while TBX2 appears to bind the same sequence as a monomer. This observation is reminiscent of the interactions of VegT and Eomesodermin with the \rightarrow NNNN \rightarrow and \rightarrow NNNN \leftarrow sites mentioned above. Also of interest is the fact that TBX2, unlike TBX1 and Brachyury, is a transcriptional repressor. Together with our results, these observations provide further insight into the functional specificities of the T box proteins.

A single amino acid can define the activity of T box proteins

Our data indicate that the different inducing activities of Xbra, VegT and Eomesodermin are mostly defined by their T boxes. Comparison of the presumed protein-DNA contact points of the three proteins, based on the crystal structure of the Xbra T box (Muller and Herrmann, 1997), suggested that lysine 149 of Xbra might be important in defining functional specificity. In support of this idea, mutation of the corresponding asparagine residue in VegT and Eomesodermin to lysine caused the modified proteins to behave more like Xbra, in that they could not induce high levels of *Pintallavis* or *chordin* and they could not activate *gooseoid* at all (Fig. 7C). Interestingly, exchange of a neutral polar residue for a basic amino acid also changes the DNA binding specificity of *Drosophila* homeodomain proteins (Hanes and Brent, 1989; Treisman et al., 1989). For example, replacing the neutral polar glutamine residue at position 9 in the recognition helix of Bicoid with the lysine found in the equivalent position of Antennapedia changes the specificity of Bicoid to that of Antennapedia (Hanes and Brent, 1989).

The mechanism by which a single amino acid substitution might change the specificity of the T box proteins is unclear. This difficulty is compounded because position 149 of Xbra contacts the phosphate backbone of DNA and is not predicted to make a base-specific contact. Indeed, our results show that Xbra, VegT and Eomesodermin select the same core sequence (Fig. 5). One possibility is that position 149 affects the affinity of protein-DNA interactions, but this is unlikely because even

the highest levels of Xbra fail to activate anterior markers such as *gooseoid* (Cunliffe and Smith, 1992; Cunliffe and Smith, 1994; O'Reilly et al., 1995; Tada et al., 1997). Another suggestion is that position 149 of Xbra might alter target specificity through protein-protein interactions, as occurs in Sox proteins (reviewed by Kamachi et al., 2000) and homeobox proteins (reviewed by Chariot et al., 1999; Mann, 1999). Consistent with this proposal, it was recently demonstrated that the transcriptional activity of the T box protein Tbr-1 is altered by its association with the guanylate kinase CASK/LIN-2 (Hsueh et al., 2000). Moreover, classical genetic studies carried out on the mouse Brachyury allele T^C are consistent with the presence of a Brachyury interacting protein (MacMurray and Shin, 1988). However, no interacting protein has been yet identified for Xbra, VegT or Eomesodermin. We plan to search for such proteins and to carry out structural analyses of T box proteins.

This work was supported by the Medical Research Council and the British Heart Foundation. E. S. C. was a Hitching-Elion Fellow. We thank Steve Smerdon for discussions on T box structure, Tim Mohun and Surendra Kotecha for help with binding-site selections, Caroline Hill for advice and help with band-shifts, and Josh Brickman for donation of constructs and help with experimental design. We are also grateful to Bob Duronio for critical comments on the manuscript.

REFERENCES

- Bamshad, M., Lin, R. C., Law, D. J., Watkins, W. C., Krakowiak, P. A., Moore, M. E., Franceschini, P., Lala, R., Holmes, L. B., Gebuhr, T. C. et al. (1997). Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nat. Genet.* **16**, 311-315.
- Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soultis, J., Grayzel, D., Kroumpouzou, E., Traill, T. A., Leblanc Straceski, J. et al. (1997). Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. *Nat. Genet.* **15**, 30-35.
- Brickman, J. M., Jones, C. M., Clements, M., Smith, J. C. and Beddington, R. S. P. (2000). Hex is a transcriptional repressor which contributes to anterior identity and suppresses the Spemann organizer. *Development* **127**, 2303-2315.
- Casey, E. S., O'Reilly, M. A., Conlon, F. L. and Smith, J. C. (1998). The T-box transcription factor Brachyury regulates expression of *eFGF* through binding to a non-palindromic response element. *Development* **125**, 3887-3894.
- Casey, E. S., Tada, M., Fairclough, L., Wylie, C. C., Heasman, J. and Smith, J. C. (1999). *Bix4* is activated directly by VegT and mediates endoderm formation in *Xenopus* development. *Development* **126**, 4193-4200.
- Chariot, A., Gielen, J., Merville, M. P. and Bours, V. (1999). The homeodomain-containing proteins: an update on their interacting partners. *Biochem. Pharmacol.* **58**, 1851-1857.
- Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. *J. Exp. Zool.* **70**, 429-459.
- Cho, K. W. Y., Blumberg, B., Steinbeisser, H. and De Robertis, E. M. (1991). Molecular nature of Spemann's organizer: the role of the *Xenopus* homeobox gene *gooseoid*. *Cell* **67**, 1111-1120.
- Christian, J. L., McMahon, J. A., McMahon, A. P. and Moon, R. T. (1991). *Xwnt-8*, a *Xenopus* Wnt-1/int-1-related gene responsive to mesoderm-inducing factors, may play a role in ventral mesodermal patterning during embryogenesis. *Development* **111**, 1045-1055.
- Conlon, F. L., Sedgwick, S. G., Weston, K. M. and Smith, J. C. (1996). Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. *Development* **122**, 2427-2435.
- Cunliffe, V. and Smith, J. C. (1992). Ectopic mesoderm formation in *Xenopus* embryos caused by widespread expression of a *Brachyury* homologue. *Nature* **358**, 427-430.
- Cunliffe, V. and Smith, J. C. (1994). Specification of mesodermal pattern in *Xenopus laevis* by interactions between *Brachyury*, *noggin* and *Xwnt-8*. *EMBO J.* **13**, 349-359.
- Di Gregorio, A. and Levine, M. (1999). Regulation of Ci-tropomyosin-like, a Brachyury target gene in the ascidian, *Ciona intestinalis*. *Development* **126**, 5599-5609.
- Gibson-Brown, J. J., Agulnik, S. I., Silver, L. M., Niswander, L. and Papaioannou, V. E. (1998). Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development. *Development* **125**, 2499-2509.
- Gluecksohn-Schoenheimer, S. (1938). The development of two tailless mutants in the house mouse. *Genetics* **23**, 573-584.
- Griffin, K. J., Amacher, S. L., Kimmel, C. B. and Kimelman, D. (1998). Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development* **125**, 3379-3388.
- Halpern, M. E., Ho, R. K., Walker, C. and Kimmel, C. B. (1993). Induction of muscle pioneers and floor plate is distinguished by the zebrafish no tail mutation. *Cell* **75**, 99-111.
- Hanes, S. D. and Brent, R. (1989). DNA specificity of the Bicoid activator protein Is determined by homeodomain recognition helix residue 9. *Cell* **57**, 1275-1283.
- He, M., Wen, L., Campbell, C. E., Wu, J. Y. and Rao, Y. (1999). Transcription repression by *Xenopus* ET and its human ortholog TBX3, a gene involved in ulnar-mammary syndrome. *Proc. Natl. Acad. Sci. USA* **96**, 10212-10217.
- Herrmann, B. G., Labeit, S., Poutska, A., King, T. R. and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**, 617-622.
- Hill, C. S., Marais, R., John, S., Wynne, J., Dalton, S. and Treisman, R. (1993). Functional analysis of growth factor-responsive transcription factor complex. *Cell* **73**, 395-406.
- Horb, M. E. and Thomsen, G. H. (1997). A vegetally localized T-box transcription factor in *Xenopus* eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. *Development* **124**, 1689-1698.
- Hsueh, Y. P., Wang, T. F., Yang, F. C. and Sheng, M. (2000). Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2. *Nature* **404**, 298-302.
- Hudson, C., Clements, D., Friday, R. V., Stott, D. and Woodland, H. R. (1997). Xsox17 α and - β mediate endoderm formation in *Xenopus*. *Cell* **91**, 397-405.
- Hyde, C. E. and Old, R. W. (2000). Regulation of the early expression of the *Xenopus* nodal-related 1 gene, Xnr1. *Development* **127**, 1221-1229.
- Isaac, A., Rodriguez-Esteban, C., Ryan, A., Altaber, M., Tsukui, T., Patel, K., Tickle, C. and Izpisua-Belmonte, J. C. (1998). Tbx genes and limb identity in chick embryo development. *Development* **125**, 1867-1875.
- Jacobs, J. J., Keblusek, P., Robanus-Maandag, E., Kristel, P., Lingbeek, M., Nederlof, P. M., van Welsem, T., van De Vijver, M. J., Koh, E. Y., Daley, G. Q. et al. (2000). Senescence bypass screen identifies TBX2, which represses cdkn2a (p19ARF) and is amplified in a subset of human breast cancers. *Nat. Genet.* **26**, 291-299.
- Jerome, L. A. and Papaioannou, V. E. (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat. Genet.* **27**, 286-291.
- Jones, C. M., Kuehn, M. R., Hogan, B. L. M., Smith, J. C. and Wright, C. V. E. (1995). Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651-3662.
- Kamachi, Y., Uchikawa, M. and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet.* **16**, 182-187.
- Kavka, A. I. and Green, J. B. A. (1997). Tales of tails: Brachyury and the T-box genes. *Biochim. Biophys. Acta.* **1333**, F73-F84.
- Kimmel, C. B., Kane, D. A., Walker, C., Warga, R. M. and Rothman, M. B. (1989). A mutation that changes cell movement and cell fate in the zebrafish embryo. *Nature* **337**, 358-362.
- Kispert, A. and Herrmann, B. G. (1993). The Brachyury gene encodes a novel DNA binding protein. *EMBO J.* **12**, 3211-3120.
- Kispert, A., Koschorz, B. and Herrmann, B. G. (1995a). The T protein encoded by *Brachyury* is a tissue-specific transcription factor. *EMBO J.* **14**, 4763-4772.
- Kispert, A., Ortner, H., Cooke, J. and Herrmann, B. G. (1995b). The chick *Brachyury* gene: developmental expression pattern and response to axial induction by localized activin. *Dev. Biol.* **168**, 406-415.
- Ku, M. and Melton, D. A. (1993). *Xwnt-11*, a maternally expressed *Xenopus* wnt gene. *Development* **119**, 1161-1173.

- Li, Q. Y., Newbury-Ecob, R. A., Terrett, J. A., Wilson, D. I., Curtis, A. R., Yi, C. H., Gebuhr, T., Bullen, P. J., Robson, S. C., Strachan, T. et al. (1997). Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat. Genet.* **15**, 21-29.
- Lindsay, E. A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H. F., Scambler, P. et al. (2001). *Tbx1* haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* **410**, 97-101.
- Logan, M. and Tabin, C. J. (1999). Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science* **283**, 1736-1739.
- Logan, M., Simon, H. G. and Tabin, C. (1998). Differential regulation of T-box and homeobox transcription factors suggests roles in controlling chick limb-type identity. *Development* **125**, 2825-2835.
- Luckow, B. and Schutz, G. (1987). CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. *Nucleic Acids. Res.* **15**, 5490.
- Lustig, K. D., Kroll, K. L., Sun, E. E. and Kirschner, M. W. (1996). Expression cloning of a *Xenopus* T-related gene (Xombi) involved in mesodermal patterning and blastopore lip formation. *Development* **122**, 4001-4012.
- MacMurray, A. and Shin, H.-S. (1988). The antimorphic nature of the TC allele at the mouse *Tlocus*. *Genetics* **120**, 545-550.
- Mann, R. S. (1999). The specificity of homeotic gene function. *BioEssays* **17**, 855-863.
- Merscher, S., Funke, B., Epstein, J. A., Heyer, J., Puech, A., Lu, M. M., Xavier, R. J., Demay, M. B., Russell, R. G., Factor, S. et al. (2001). TBX1 is responsible for cardiovascular defects in Velo-Cardio-Facial/DiGeorge Syndrome. *Cell* **104**, 619-629.
- Muller, C. W. and Herrmann, B. G. (1997). Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. *Nature* **389**, 884-888.
- Nieuwkoop, P. D. and Faber, J. (1975). *Normal Table of Xenopus laevis (Daudin)*. Amsterdam: North Holland.
- O'Reilly, M.-A., Smith, J. C. and Cunliffe, V. (1995). Patterning of the mesoderm in *Xenopus*: dose-dependent and synergistic effects of Brachyury and *Pintallavis*. *Development* **121**, 1351-1359.
- Ohuchi, H., Takeuchi, J., Yoshioka, H., Ishimaru, Y., Ogura, K., Takahashi, N., Ogura, T. and Noji, S. (1998). Correlation of wing-leg identity in ectopic FGF-induced chimeric limbs with the differential expression of chick *Tbx5* and *Tbx4*. *Development* **125**, 51-60.
- Papioannou, V. E. and Silver, L. M. (1998). The T-box gene family. *BioEssays* **20**, 9-19.
- Pflugfelder, G. O., Roth, H. and Poock, B. (1992). A homology domain shared between *Drosophila* optomotor-blind and mouse Brachyury is involved in DNA binding. *Biochem. Biophys. Res. Commun.* **186**, 918-925.
- Pollock, R. and Treisman, R. (1990). A sensitive method for the determination of protein-DNA binding specificities. *Nucleic Acids Res.* **18**, 6197-6204.
- Rodriguez-Esteban, C., Tsukui, T., Yonei, S., Magallon, J., Tamura, K. and Izpisua Belmonte, J. C. (1999). The T-box genes *Tbx4* and *Tbx5* regulate limb outgrowth and identity. *Nature* **398**, 814-818.
- Rosa, F. M. (1989). *Mix.1*, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell* **57**, 965-974.
- Ruiz i Altaba, A. and Jessell, T. M. (1992). *Pintallavis*, a gene expressed in the organizer and midline cells of frog embryos: involvement in the development of the neural axis. *Development* **116**, 81-93.
- Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A., Carlton, M. B., Pearce, J. J., Barton, S. C., Surani, M. A., Ryan, K., Nehls, M. C. et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* **404**, 95-99.
- Ryan, K., Garrett, N., Mitchell, A. and Gurdon, J. B. (1996). Eomesodermin, a key early gene in *Xenopus* mesoderm differentiation. *Cell* **87**, 989-1000.
- Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. and De Robertis, E. M. (1994). *Xenopus* chordin: a novel dorsaling factor activated by organizer-specific homeobox genes. *Cell* **79**, 779-790.
- Schulte-Merker, S., Ho, R. K., Herrmann, B. G. and Nüsslein-Volhard, C. (1992). The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* **116**, 1021-1032.
- Schulte-Merker, S., van Eeden, F. J., Halpern, M. E., Kimmel, C. B. and Nüsslein-Volhard, C. (1994). *no tail (ntl)* is the zebrafish homologue of the mouse T (Brachyury) gene. *Development* **120**, 1009-1015.
- Sinha, S., Abraham, S., Gronostajski, R. M. and Campbell, C. E. (2000). Differential DNA binding and transcription modulation by three T-box proteins, T, TBX1 and TBX2. *Gene* **258**, 15-29.
- Slack, J. M. W. (1984). Regional biosynthetic markers in the early amphibian embryo. *J. Embryol. Exp. Morphol.* **80**, 289-319.
- Smith, J. (1999). T-box genes: what they do and how they do it. *Trends Genet.* **15**, 154-158.
- Smith, J. C. (1993). Purifying and assaying mesoderm-inducing factors from vertebrate embryos. In *Cellular Interactions in Development – A Practical Approach* (ed. D. Hartley), pp. 181-204. Oxford: Oxford University Press.
- Smith, J. C. and Slack, J. M. W. (1983). Dorsalization and neural induction: properties of the organizer in *Xenopus laevis*. *J. Embryol. Exp. Morphol.* **78**, 299-317.
- Smith, J. C., Price, B. M., Green, J. B. A., Weigel, D. and Herrmann, B. G. (1991). Expression of a *Xenopus* homolog of *Brachyury* (T) is an immediate-early response to mesoderm induction. *Cell* **67**, 79-87.
- Smith, W. C. and Harland, R. M. (1991). Injected *Xwnt-8* RNA acts early in *Xenopus* embryos to promote formation of a vegetal dorsaling center. *Cell* **67**, 753-765.
- Stennard, F., Carnac, G. and Gurdon, J. B. (1996). The *Xenopus* T-box gene, *Antipodean*, encodes a vegetally localised maternal mRNA and can trigger mesoderm formation. *Development* **122**, 4179-4188.
- Stennard, F., Zorn, A. M., Ryan, K., Garrett, N. and Gurdon, J. B. (1999). Differential expression of VegT and Antipodean protein isoforms in *Xenopus*. *Mech. Dev.* **86**, 87-98.
- Tada, M. and Smith, J. C. (2000). *Xwnt11* is a target of *Xenopus* Brachyury: regulation of gastrulation movements via Dishevelled but not through the canonical Wnt pathway. *Development* **127**, 2227-2238.
- Tada, M., O'Reilly, M.-A. J. and Smith, J. C. (1997). Analysis of competence and of *Brachyury* autoinduction by use of hormone-inducible *Xbra*. *Development* **124**, 2225-2234.
- Tada, M., Casey, E. S., Fairclough, L. and Smith, J. C. (1998). *Bix1*, a direct target of *Xenopus* T-box genes, causes formation of ventral mesoderm and endoderm. *Development* **125**, 3997-4006.
- Takeuchi, J. K., Koshiba-Takeuchi, K., Matsumoto, K., Vogel-Hopker, A., Naitoh-Matsuo, M., Ogura, K., Takahashi, N., Yasuda, K. and Ogura, T. (1999). *Tbx5* and *Tbx4* genes determine the wing/leg identity of limb buds. *Nature* **398**, 810-814.
- Treisman, J., M., G., E., V., Harris, E. and Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**, 553-562.
- Wilkinson, D. G., Bhatt, S. and Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657-659.
- Zhang, J. and King, M. L. (1996). *Xenopus* VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. *Development* **122**, 4119-4129.
- Zhang, J., Houston, D. W., King, M. L., Payne, C., Wylie, C. and Heasman, J. (1998). The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell* **94**, 515-524.