Int. J. Dev. Biol. 40: 291-295 (1996)

Zinc finger proteins in early *Xenopus* development

THOMAS HOLLEMANN, ERIC BELLEFROID, REIMER STICK and TOMAS PIELER*

Georg-August-Universität Göttingen, Institut für Biochemie und Molekulare Zellbiologie, Göttingen, Germany

ABSTRACT The C_2H_2-type zinc finger motif defines a large super family of specific DNA and specific RNA binding proteins. Individual members of this protein family have been demonstrated to carry important regulatory functions in embryogenesis. We have isolated a large collection of C_2H_2-type zinc finger proteins from *Xenopus laevis*. Some of these proteins are highly conserved in evolution and found to be differentially expressed during embryonic development of the central nervous system. We also summarize our recent findings on the biochemical characterization of RNA and DNA binding activities *in vitro* for other Xenopus zinc finger proteins, which fall into structurally defined, distinct subfamilies.

KEY WORDS: *Xenopus laevis, zinc finger*

Introduction

The zinc finger is one of several conserved nucleic acid binding motifs in eukaryotes. The length of a typical zinc finger unit is 28 amino acids; it carries several invariant residues, in particular two zinc coordinating pairs of cysteines and histidines (C_2H_2) (Miller *et al.*, 1985). Other highly conserved, hydrophobic residues help to fold the unit into a compact ß-ß-α structure (Fig. 1). Hypervariable residues have been found to play important roles in sequence specific DNA recognition, which occurs primarily via base specific contacts in the major groove of the DNA double helix with amino acids of the α-helical element in the C_2H_2 type zinc finger module (Pavletich and Pabo, 1991; Pavletich and Pabo, 1993; Fairall *et al.*, 1993); this situation is representative for a principle mode of DNA recognition, which is also used by other DNA binding modules, including the universal helix-turn-helix motif. However, several marked characteristics make zinc finger proteins (ZFPs) clearly distinct from other conserved families of nucleic acid binding proteins: 1) zinc fingers are generally found in multiple copies (up to 37) per protein; these can be organized in tandem array, forming a single or multiple clusters, or they can be dispersed throughout the protein molecule. 2) ZFPs can be DNA binding, or they can exhibit specific RNA binding activity, or both, DNA and RNA binding activities. 3) ZFPs define an extremely large superfamily of nucleic acid binding proteins with several hundred, perhaps thousands of different members of vertebrates (reviewed in El-Baradi and Pieler, 1991; Pieler and Bellefroid, 1994).

Zinc finger protein structure and evolution

Our structural analysis of ZFP encoding cDNA and genomic clones from *Xenopus laevis* has revealed that common, conserved sequence elements outside of the zinc finger clusters define structurally distinct ZFP subgroups. These conserved modules have been termed FAX-(Finger-Associated-boXes) or FAR-(Finger-Associated-Repeat) domain (Knöchel *et al.*, 1989; Klocke *et al.*, 1994). Others have identified similar finger associated motifs, such as the evolutionarily conserved KRAB- or ZIN-domains (Bellefroid *et al.*, 1991; Numoto *et al.*, 1993). The exact biological function for these modules remains to be determined; in the case of the KRAB domain, it has been reported more recently that it can inhibit reporter gene transcription in transient transfection assays (Witzgall *et al.*, 1994; Margolin, 1994). However, if such KRAB-ZFPs function as bona fide transcription repressors, or if the negative regulatory effect is mediated by a more indirect molecular mechanism, such as an altered chromatin structure, remains to be determined.

A detailed analysis of *Xenopus* FAX-ZFP gene organization has revealed that these transcription units are organized in large clusters as tandem repeats (Nietfeld *et al.*, 1993). Furthermore, within such FAX-ZFP clusters, the FAX domain encoding portion exhibits a striking degree of structural conservation, not only in terms of predicted protein sequence, but also in respect to exon/intron boundaries and, most strikingly, in respect to length and sequence of the introns. In contrast, zinc finger clusters connected to the FAX domain are generally encoded by a single exon and they differ largely in the number of zinc finger units encoded and also in primary sequence. Corresponding structural features have been described for human KRAB-ZFP clusters (Bellefroid *et al.*, 1993). A systematic search for homologs of one specific member of the human KRAB-ZFP family in other species has revealed that evolutionary conservation is restricted solely to higher primates (Bellefroid *et al.*, 1995). On the basis of these findings, we have proposed that the vast majority of ZFP encoding genes in vertebrate systems are the result of gene amplification events which have occurred late in evolution. The biological function for these ZFPs, i.e. for the vast majority of

***Address for reprints:** Georg-August-Universität Göttingen, Institut für Biochemie und Molekulare Zellbiologie, Humboldtallee 23, D-37073 Göttingen, Germany. FAX: 49-551.395960.

0214-6282/96/$03.00

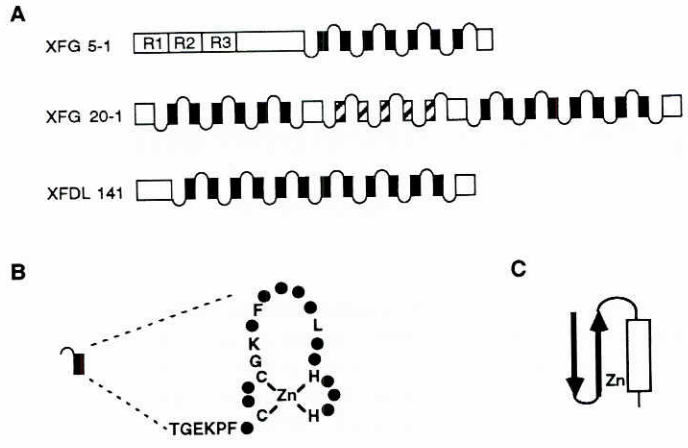

Fig. 1. Structural features of *Xenopus* zinc finger proteins. (A) *Schematic structural organization of various Xenopus zinc finger proteins discussed in this article. Zinc finger modules are represented by filled or crosshatched bars, non-finger protein elements by open bars. R1 to R3 in XFG 5-1 refers to the non finger, tandem repeat elements of the FAR domain.* **(B)** *Consensus sequence of Xenopus C2H2 ZFP repeat units folded around a zinc ion, as derived from the sequence compilation of Nietfeld et al. (1989).* **(C)** *Schematic representation of common secondary-tertiary structure features in zinc finger modules; two antiparallel ß-strands (arrows) and one helical element (rectangle) are indicated.*

ZFPs in vertebrates, remains highly enigmatic (Pieler and Bellefroid, 1994).

Very much in contrast to these clustered, late evolving subfamilies of zinc finger proteins, a structurally distinct subset of ZFPs has been highly conserved in terms of structure and also in terms of specific aspects of their biological function. One such ZFP is the eukaryotic transcription factor IIIA (TFIIIA), which serves housekeeping functions both, as a specific DNA binding protein in 5S ribosomal RNA transcription and as a specific RNA binding protein in 5S RNA storage and nucleo-cytoplasmic transport (as described in the article by Rudt *et al.*). Structural and functional homologs for TFIIIA have been described for a broad variety of organisms, ranging from yeast to man. The majority of other highly conserved ZFPs have functions clearly distinct from a role in housekeeping and they are, from the point of view of a developmental biologist, more attractive. Many of these regulatory ZFPs fall into only a few structurally conserved subfamilies, sharing highly homologous zinc finger clusters (3-5 units) with similar DNA binding characteristics (reviewed in Pieler and Bellefroid, 1994). The embryonic expression of individual members of such subfamilies has been found to be differentially regulated, providing a good indication for a regulatory role in embryonic differentiation events. Thus, the aim of our ZFP related research in the past has been twofold: 1) the identification and structural/functional characterization of *Xenopus* ZFPs with a regulatory role in early *Xenopus* embryogenesis, and 2) elucidation of the biological function for the multitude of moderately conserved, clustered ZFP encoding genes in *Xenopus*.

Xenopus MyTI

The zinc finger protein MyTI has originally been isolated by virtue of its binding activity to the upstream region of a glia specific gene, the myelin proteolipid protein (PLP) gene from human (Kim and Hudson, 1992). We have characterized the *Xenopus* homolog, XMyTI. The XMyTI protein contains five unusual finger motifs with C2HC as putative zinc coordinating residues and these are organized in two widely separated clusters (Fig. 2). In addition to the *Xenopus* sequence, we have also characterized corresponding proteins from rat and *Caenorhabditis elegans*. The predicted proteins exhibit a high degree of sequence identity within the zinc finger clusters (80-100% on the amino acid level), sequence elements outside of the zinc finger encoding region are largely divergent. The *C. elegans* MyTI protein lacks the entire second zinc finger cluster (E. Bellefroid and T. Pieler, manuscript in preparation).

The expression of XMyTI in the course of *Xenopus* embryogenesis has been analyzed by whole mount *in situ* hybridization. Embryonic XMyTI gene expression is neural specific and it is first detected in late gastrulae (stage 12). RNA signals were obtained in the area of the olfactory placodes and within the developing neural plate in a belt of eight longitudinal bands which are symmetrically located laterally to the dorsal midline. All signals are specific to the sensorial layer of the neuroectoderm. Later on in embryonic development XMyTI expression is found in prosencephalon, mesencephalon and rhombencephalon, as well as in the spinal cord, retina and cranial nerves. Neural crest cells are also strongly labeled (Fig. 3). In summary, XMyTI is neural specific and, most interestingly, it is one of the earliest neural markers described so far. These expression characteristics suggest that the XMyTI protein might play an early regulatory role in the patterning of the central nervous system (E. Bellefroid and T. Pieler, in preparation).

Xenopus X-sal

The *Drosophila* gene *spalt (sal)* has been characterized by genetic means as a "region specific homeotic" gene. It acts in

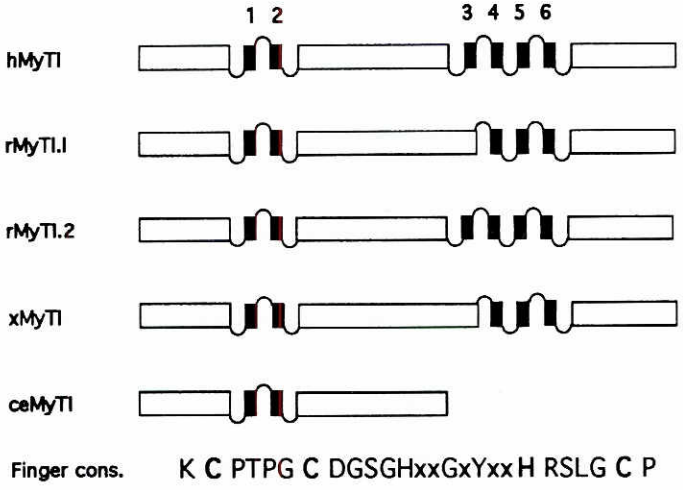

Fig. 2. Comparative schematic structural organization of the human, rat, *Xenopus* and *C. elegans* MyTI proteins. *Zinc finger modules are indicated by filled rectangles. The C2H2 type zinc finger consensus sequence is derived from the corresponding human, rat, Xenopus and C. elegans protein sequence (Kim and Hudson, 1992; Bellefroid and Pieler, in preparation).*

the terminal regions of the embryo, specifying pattern elements in both the head and tail regions (Jürgens, 1988). sal encodes a zinc finger protein of novel structure composed of three widely spaced "double zinc finger" motifs of internally conserved sequences and a single zinc finger motif of different sequence (Kühnlein *et al.*, 1994).

We have cloned a *Xenopus* gene *(X-sal)* related to the *sal* gene of *Drosophila*. The *Xenopus* X-SAL protein shows the same overall molecular architecture as the *Drosophila* SAL protein. The double zinc finger motifs as well as the single finger sequence are highly conserved between the two species. Identical zinc finger motifs were also characterized from human DNA by PCR. The remaining sequence of the *X-sal* gene, however, does not show obvious similarities to the *Drosophila* sequence and analysis of the genomic organization revealed that the two genes differ in their exon / intron organization.

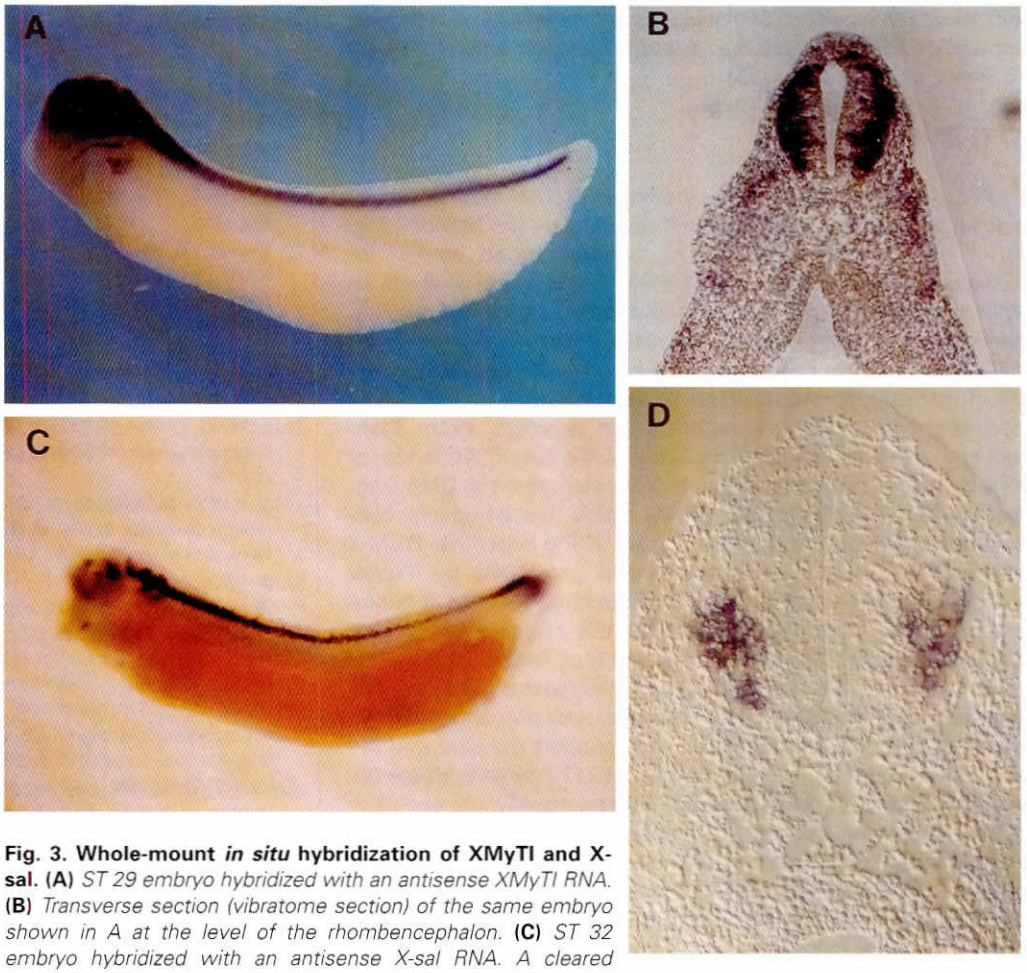

Fig. 3. Whole-mount *in situ* hybridization of XMyTI and X-sal. (A) *ST 29 embryo hybridized with an antisense XMyTI RNA.* **(B)** *Transverse section (vibratome section) of the same embryo shown in A at the level of the rhombencephalon.* **(C)** *ST 32 embryo hybridized with an antisense X-sal RNA. A cleared embryo is shown.* **(D)** *Transverse section (plastic embedded) of the same embryo shown in C at the level of the rhombencephalon.*

The temporal and spatial expression pattern of *X-sal* has been analysed by RNase protection assays and by whole mount *in situ* hybridization. *X-sal* transcripts are absent from oocytes and cleavage stage embryos. The gene is first transcribed during gastrulation. Early in neurulation RNA staining is detected in the posterior neural plate. In the course of neurulation the staining spreads anteriorly up to the prospective anterior boundary of the hindbrain and at the end of neurulation it becomes restricted to two narrow stripes parallel to the dorsal midline. Later in development (ST 30) *X-sal* is expressed in the midbrain, hindbrain, and spinal cord in a highly regionalized manner (Fig. 3). In the spinal cord, for example, expression is confined to the ventro-lateral columns of motor neurons. Outside the CNS expression is found in the VIIth cranial ganglion. In conclusion, the highly regionalized expression suggests that X-SAL protein might function in cell specification in the developing *Xenopus* nervous system.

XFDL141

An alternative experimental approach with the aim of identifying regulator proteins of development is to search for genes which are differentially expressed in regions of the embryo with particular importance for embryonic differentiation, such as the Spemann organizer. A dorsal blastopore lip specific cDNA library has been generated by Eddy M. De Robertis and his colleagues and it has lead to the identification of the homeobox gene goosecoid as an important regulator for the specification of the anterior-posterior body axis (Blumberg *et al.*, 1991; Cho *et al.*, 1991). In collaboration with the same group, we have identified ZFP encoding sequences over-represented in the blastopore lip specific cDNA library. One of these, termed XFDL141, has been characterized in detail. XFDL141 is a typical member of the C2H2 class, encoding twelve zinc finger modules in tandem repeat (Fig. 1). Analysis of the temporal pattern of XFDL141 expression in *Xenopus* embryogenesis by Northern blotting reveals that expression commences at the zygotic onset of transcription with the mid blastula transition. XFDL141 is in fact one of the earliest genes to be zygotically activated. RNA levels decline in gastrula and disappear at the gastrula-neurula transition. Whole mount *in situ* hybridization identifies XFDL 141 transcripts within the entire animal hemisphere of blastula stage embryos. At the small yolk plug stage (stage 13) a weak signal is restricted to a region that might correspond to the anterior rim of the developing neural plate. these unique expression characteristics indicate that XFDL141 could function as a very early element in the regulation of *Xenopus* development. Future experi-

mentation will be aimed at elucidating biochemical and biological aspects of XFDL141 function in early *Xenopus* embryogenesis (T. Pieler *et al.*, in preparation).

XFG 20-1

XFG 20-1 is a C2H2 type zinc finger protein that has its 19 zinc finger modules organized in three clusters (or "hands", Fig. 1). It has been isolated in our original, broad screen for *Xenopus* ZFPs (Nietfeld *et al.*, 1989). Northern blot analysis on RNA preparations from embryos and various adult tissues suggests that it is expressed ubiquitously. Structural homologs for XFG 20-1 have not been described for other species and analysis of its genomic organization in *Xenopus* reveals that all 19 zinc finger modules are organized in a single exon. More recently, we have identified a putative DNA target sequence for XFG 20-1 by affinity selection of *Xenopus* genomic DNA fragments on a XFG 20-1 protein affinity matrix. The 54 bp long consensus recognition sequence is part of the repetitive REM-1 element that had been characterized previously by Walter Knöchel and his coworkers. It is preferentially located at telomeric regions or near the centromeres, but it is also found to be dispersed within the context of various transcription units (Hummel *et al.*, 1984). These findings suggest a functional link between long repetitive DNA elements and zinc finger protein function. In more general terms, they support the idea that the non-conserved nature of the majority of ZFPs in vertebrates (as pointed out above) could correlate with an equally late evolution of their nucleic acid target structures, as it is indeed the case for the complex superfamily of diverse repetitive DNA sequence elements in higher eukaryotes. One of the possibilities that we will test experimentally is if XFG 20-1 binding to the repetitive REM-1 element influences chromatin structure (Schäfer *et al.*, 1994).

XFG 5-1

XFG 5-1 is one out of several structurally related *Xenopus* ZFPs, which exhibit specific RNA homopolymer binding activity with highest affinity for p(U) (Köster *et al.*, 1991). These proteins share a unique repeat motif at the amino-terminus (Fig. 1), which is referred to as the FAR (Finger Associated Repeat) domain (Klocke *et al.*, 1994). For different members of this subfamily it has been demonstrated that the p(U) binding activity resides within the zinc finger cluster of such proteins (Köster *et al.*, 1991; Klocke *et al.*, 1994). In *Xenopus* oocytes, the XFG 5-1 protein is found in both, nuclear and cytoplasmic compartments. In the nucleus, but not in the cytoplasm, XFG 5-1 is phosphorylated (Van Wijk *et al.*, 1992). cDNA sequence analysis identifies XFG 5-1 isoforms generated by alternative splicing (Köster *et al.*, 1993). These isoforms differ in the number of FAR specific, amino-terminal repeat units but maintain identical zinc finger clusters. The exact biological function of XFG 5-1 remains to be elucidated; as a ubiquitous RNA binding protein it will be involved in activities distinct from cell specific transcription regulation. Identification of a specific RNA target for XFG 5-1, as well as identification of proteins interacting with the FAR domain will be emphasis of our future experimentation.

References

BELLEFROID, E.J., MARINE, J-C., MATERA, A.G., DESAI, T., BOURGUIGNON, C., HEALEY, K., BRAY-WARD, P., IHLE, J., MARTIAL, J.A. and WARD, D.C. (1995). Emergence of the ZNF 91 KRAB-zinc finger gene family after the divergence between the emthropoid and posimian primate lineages. *Proc. Natl. Acad. Sci. USA 92:* 10752-10761.

BELLEFROID, E.J., MARINE, J-C., RIED, T., LOCOCQ, P.J., RIVIÈRE, M., AMEMIYA, C., POUCELET, D.A., COULIE, P.G., DE JONG, P., SZPIRER, C., WARD, D.C. and MARTIAL, J.A. (1993). Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO J. 12:* 1363-1374.

BELLEFROID, E.J., PONCELET, D.A., LECOCQ, P.J., REVELANT, O. and MARTIAL, J. (1991). The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc. Natl. Acad. Sci. USA 88:* 3608-3612.

BLUMBERG, B., WRIGHT, C.V.E., DE ROBERTIS, E.M. and CHO, K.W.Y. (1991). Organizer specific homeobox genes in *Xenopus laevis* embryos. *Science 253:* 194-196.

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 goosecoid. *Cell 67:* 1111-1120.

EL-BARADI, T. and PIELER, T. (1991). Zinc finger proteins: what we know and what we would like to know. *Mech. Dev. 35:* 155-169.

FAIRRALL, L., SCHWABE, J.W.R., CHAPMAN, L., FINCH, J.T. and RHODES, D. (1993). The crystal structure of a two zinc-finger peptide reveals an extension to the rules of zinc-finger/DNA recognition. *Nature 366:* 483-487.

HUMMEL, S., MEYERHOF, W., KORGE, E. and KNÖCHEL, W. (1984). Characterization of highly and moderately repetitive 500-bp EcoRI fragments from *Xenopus laevis* DNA. *Nucleic Acids Res. 12:* 4921-4938.

JÜRGENS, G. (1988). Head and tail development of the *Drosophila* embryo involves spalt, a novel homeotic gene. *EMBO J. 7:* 189-196.

KIM, J.G. and HUDSON, L.D. (1992). Novel member of the zinc finger superfamily: a C2H2 finger that recognizes a Glia-specific gene. *Mol. Cell. Biol. 12:* 5632-5639.

KLOCKE, B., KÖSTER, M., HILLE, S., BOUWMEESTER, T., BÖHM, S., PIELER, T. and KNÖCHEL, W. (1994). The FAR domain defines a new *Xenopus laevis* zinc finger protein subfamily with specific RNA homopolymer binding activity. *Biochim. Biophys. Acta 1217:* 81-89

KNÖCHEL, W., PÖTING, A., KÖSTER, M., EL-BARADI, T., NIETFELD, W., BOUWMEESTER, T. and PIELER, T. (1989). Evolutionary conserved modules associated with Zn fingers in *Xenopus laevis. Proc. Natl. Acad. Sci. USA 86:* 6097-6100.

KÖSTER, M., HILLE, S., PIELER, T. and KNÖCHEL, W. (1993). Gene structure and alternative splicing of XFG 5-1, a *Xenopus laevis* zinc finger protein with RNA homopolymer binding activity. *Mol. Biol. Reprod. 18:* 197-207

KÖSTER, M., KÜHN, U., BOUWMEESTER, T., NIETFELD, W., EL-BARADI, T., KNÖCHEL, E. and PIELER, T. (1991). Structure, expression and *in vitro* functional characterization of a novel RNA binding Zinc finger protein from *Xenopus. EMBO J. 10:* 3087-3093.

KÜHNLEIN, R.P., FROMMER, G., FRIEDRICH, M., GONZALEZ-GAITAN, M., WEBER, A., WAGNER-BERNHOLZ, J.F., GEHRING, W.J., JÄCKLE, H. and SCHUH, R. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the *Drosophila* embryo. *EMBO J. 13:* 168-179.

MARGOLIN, J.F., FRIEDMAN, J.R., MEYER, W.K-H., VISSING, H., THIESEN, H-J. and RAUSCHER, F.J. (1994). Krüppel-associated boxes are potent transcriptional repressor domains. *Proc. Natl. Acad. Sci. USA 91:* 4509-4513.

MILLER, J., McLACHLAN, A.D. and KLUG, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J. 4:* 1609-1614.

NIETFELD, W., CONRAD, S., VAN WIJK, I., GILTAY, R., BOUWMEESTER, T., KNÖCHEL, W. and PIELER, T. (1993). Evidence for a clustered genomic organization of Fax-zinc finger protein encoding transcription units in *Xenopus laevis. J. Mol. Biol. 230:* 400-412.

NIETFELD, W., EL-BARADI, T., MENTZEL, H., PIELER, T., KÖSTER, M., PÖTING, A. and KNÖCHEL, W. (1989). Second order repeats in *Xenopus laevis* finger proteins. *J. Mol. Biol. 208:* 639-659.

NUMOTO, M., NIWA, O., KAPLAN, J., WONG, K-K., MERRELL, K., KAMIYA, K., YANAGIHARA, K. and CALAME, K. (1993). Transcriptional repressor ZF 5 identifies a new conserved domain in zinc finger proteins. *Nucleic Acids Res. 21:* 3767-3775.

PAVLETICH, N.P. and PABO, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif 268-DNA complex at 2.1 A. *Science 252:* 809-817.

PAVLETICH, N.P. and PABO, C.O. (1993). Crystal structure of a five-finger Gli-DNA complex: New perspectives on zinc fingers. *Science 261:* 1701-1707.

PIELER, T. and BELLEFROID, E. (1994). Perspectives on zinc finger protein function - an update. *Mol. Biol. Reprod. 20:* 1-8.

SCHÄFER, U., RAUSCH, O., BOUWMEESTER, T. and PIELER, T. (1994). Sequence specific recognition of a repetitive DNA element by a C2H2 zinc finger protein in *Xenopus*. *Eur. J. Biochem. 226:* 567-576.

VAN WIJK, I., BURFEIND, J. and PIELER, T. (1992). Nuclear transport and phosphorylation of the RNA binding *Xenopus* zinc finger protein XFG 5-1. *Mech. Dev. 39:* 63-72.

WITZGALL, R., O'LEARY, E., LEAF, A., ÖNALDI, D. and BONVENTRE, J.V. (1994). The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. USA 91:* 4514-4518.